

Individual differences in visual (mis)perception: a multivariate statistical approach

Présentée le 28 mai 2021

Faculté des sciences de la vie
Laboratoire de psychophysique
Programme doctoral en neurosciences

pour l'obtention du grade de Docteur ès Sciences

par

Aline Françoise CRETENOU

Acceptée sur proposition du jury

Prof. M. D. C. Sandi Perez, présidente du jury
Prof. M. Herzog, directeur de thèse
Dr J. Bosten, rapporteuse
Prof. P. Ghisletta, rapporteur
Prof. F. Hummel, rapporteur

"All models are wrong, but some are useful."

George Box

Acknowledgments

First of all, I would like to express my gratitude to my supervisor Michael for guiding me throughout this journey. Thank you for giving me the opportunity to explore my favorite research questions in an open-minded and competition-free lab. I sincerely enjoyed sharing scientific as well as non-scientific discussions with you.

Many thanks to all past and present LPSY members for the good moments we shared. A special thanks to Leïla, who became so much more than a labmate. Your eternal support and sunny mood are invaluable. I admire you and cherish our friendship. To Oh-Hyeon, my Korean friend. Like Ryan the lion, you brought so much joy in my daily life. Thank you for all the trips, coffee breaks, dinners, and chats we shared. To Ophélie, for the many discussions about our future as women. Your strength is exemplary. To Lukasz. I learned so much from you, thank you. To Marc. Thank you for conveying so much enthusiasm and positive waves. And to all the others, thank you for making my years as a PhD student enjoyable.

To Janir, my soulmate. Your love and support mean the world to me. You are the piece that completes me and makes me a better and stronger person. I am so grateful to now call you my husband. Obrigad pá tud ke bo tá representá pá mi!

A mon grand frère Julien. Thank you for being a role model in my life. Your support and advices are priceless. I do not say it enough, but I love you.

A mes parents, Marie-Christine et Gérald. Votre soutien et amour inconditionnels sont une réelle bénédiction. Vous êtes la lumière qui me guide dans cette aventure incroyable qu'est la vie. Merci pour tout, simplement ! Je vous aime profondément.

Contents

Abstract	1
Résumé	3
1. Preface.....	5
2. Introduction	7
2.1. Visual illusions.....	7
2.1.1. What is a visual illusion?	7
2.1.2. Taxonomies of visual illusions	9
2.2. Individual differences.....	10
2.2.1. Vision in general.....	11
2.2.2. Visual illusions	12
2.3. Perceptual learning and its hyperspecificity.....	16
2.4. Aims of the present work.....	17
3. General methods.....	19
3.1. Adjustment procedure.....	19
3.2. Data analysis	20
3.2.1. Test-retest reliability	20
3.2.2. Illusion magnitudes	21
3.2.3. Data transformation and outlier removal.....	21
3.2.4. Between-condition correlations	21
3.2.5. Factor analysis.....	22
4. Results	23
4.1. Are factors for visual illusions hyperspecific?.....	23
4.1.1. Study 1: Luminance-unspecific	23
4.1.2. Study 1: Orientation-unspecific	26
4.1.3. Study 1: Feature-unspecific.....	28
4.1.4. Study 2: Context-unspecific	30
4.1.5. Study 3: Factors for illusions combine independently.....	32
4.2. Are individual differences in the perception of visual illusions stable?	35
4.2.1. Study 4: Stable individual differences across eyes.....	36
4.2.2. Study 4: Stable individual differences across time	37
4.2.3. Study 4: Stable individual differences across measurement methods.....	41
4.3. Is there a common factor for vision in special populations?.....	44
4.3.1. Study 5: Healthy aging	45
4.3.2. Study 6: Patients with schizophrenia	49
4.3.3. Study 7: Action video game players	55
5. Discussion.....	61

5.1. Summary	61
5.2. A multifactorial structure for visual illusions does not preclude from a more general – albeit weak – common factor.....	62
5.3. Implications for vision research	63
5.4. Analogy with the classic model of vision	64
5.5. Islands of significant correlations within an ocean of non-correlations.....	66
5.6. The notion of a common factor is ill-posed	67
5.7. Batteries of tasks are powerful	69
5.8. Limitations	70
5.8.1. Power	70
5.8.2. Experimental flaws.....	70
5.8.3. Non-perceptual factors	71
5.8.4. Strengths and weaknesses of the adjustment procedure	71
5.8.5. Non-zero control bias in illusion magnitudes	72
5.8.6. Cross-sectional design.....	73
5.8.7. Tasks other than visual.....	73
5.9. Future work.....	73
5.10. Conclusions	74
6. References.....	75
Appendix A	93
Appendix B	117
Appendix C.....	133
Appendix D	153
Appendix E.....	195
Appendix F.....	227
Appendix G	269
Appendix H	289
Appendix I.....	305
Curriculum Vitae.....	325

Abstract

Common factors are omnipresent in everyday life, e.g., people who do well in one cognitive test are likely to perform well in other cognitive tests as well, and vice versa. In vision, however, there seems to be a multitude of specific factors rather than a strong and unique common factor. For example, only weak correlations were observed between the susceptibility to different visual illusions, suggesting that the structure underlying visual illusions is multifactorial.

In this thesis, I first examined the multidimensionality of the structure underlying visual illusions. To this aim, the susceptibility to various visual illusions was measured. In addition, subjects were tested with variants of the same illusion, which differed in spatial features, luminance, orientation, or contextual conditions. In line with previous work, only weak correlations were observed between the susceptibility to different visual illusions (i.e., between-illusion). An individual showing a strong susceptibility to one visual illusion does not necessarily show a strong susceptibility to other visual illusions. In contrast, there were strong correlations between the susceptibility to variants of the same illusion (i.e., within-illusion). Hence, factors seem to be illusion-specific but not feature-specific. In addition, the susceptibility to combinations of two illusions, which I call merged illusions, was measured to investigate how factors for illusions interact. The susceptibility to a merged illusion was strongly predicted from the susceptibility to the two illusions of which it was made, suggesting that factors for illusions combine independently to form complex illusory percepts.

Second, I showed that individual differences in the perception of visual illusions are robust across the eye(s) to which the stimulus is presented, over a month, and when measured with an adjustment procedure or a method of constant stimuli.

Last, I investigated whether a strong visual factor emerges in healthy elderly and patients with schizophrenia, which may be expected from the general decline in perceptual abilities usually reported in these two populations compared to healthy young adults. Similarly, a strong visual factor may emerge in action video gamers, who, contrary to healthy elderly and patients with schizophrenia, often show enhanced perceptual performance compared to non-video gamers. Hence, healthy elderly, patients with schizophrenia, and action video gamers were tested with a battery of visual tasks, such as a contrast detection and orientation discrimination task. As in control groups, between-task correlations were weak in general, which argues against the emergence of a strong common factor for vision in these populations.

While similar tasks are usually assumed to rely on similar neural mechanisms, the performances in different visual tasks were only weakly related to each other, i.e., performance does not generalize across visual tasks. The results presented in this thesis highlight the relevance of an individual differences approach to unravel the multidimensionality of the visual structure.

Keywords: visual perception, visual illusions, individual differences, common factors, structure, variability, reliability, correlational analyses

Résumé

Les facteurs communs sont omniprésents dans la vie de tous les jours. Ainsi, les personnes qui présentent un bon score à un test cognitif sont également susceptibles de réaliser une bonne performance à d'autres tests cognitifs, et inversement. Dans le domaine visuel, au contraire, il semble y avoir une multitude de facteurs spécifiques plutôt qu'un unique et fort facteur commun. Par exemple, seules de faibles corrélations ont été observées entre la susceptibilité à diverses illusions d'optique, suggérant que la structure sous-jacente aux illusions d'optique est multifactorielle.

Dans cette thèse, j'ai commencé par examiner l'étendue de l'espace sous-jacent aux illusions d'optique. Pour ce faire, la susceptibilité à diverses illusions d'optique a été mesurée. Les sujets ont également été testés avec des variantes d'une même illusion qui différaient en termes de caractéristiques spatiales, de luminance, d'orientation, ou de contexte. Comme attendu au vu de précédents travaux, seules de faibles corrélations ont été observées entre la susceptibilité à différentes illusions d'optique (c.-à-d. inter-illusion). En d'autres termes, un individu présentant une forte susceptibilité à une illusion d'optique donnée ne montre pas forcément une forte susceptibilité à d'autres illusions d'optique. En revanche, différentes variantes d'une même illusion (c.-à-d. intra-illusion) étaient fortement inter-corrélées. Ainsi, les facteurs semblent être spécifiques aux illusions, mais pas à leurs caractéristiques intrinsèques. De plus, la susceptibilité à des combinaisons de deux illusions, appelées illusions combinées, a été mesurée afin d'examiner l'interaction entre facteurs. La susceptibilité à une illusion combinée était fortement prédite à partir de la susceptibilité aux deux illusions qui la formaient, ce qui suggère que les facteurs sous-jacents aux illusions d'optique se combinent indépendamment pour former des percepts illusoire complexes.

Dans un deuxième temps, j'ai pu montrer que les différences individuelles dans la perception des illusions d'optique sont constantes à travers l'œil ou les yeux au(x)quel(s) est présenté le stimulus, sur une période d'un mois, et lorsqu'elles sont mesurées avec une procédure d'ajustement ou une méthode des stimuli constants.

Finalement, j'ai cherché à découvrir si un fort facteur visuel émerge chez les personnes âgées en bonne santé et les patients atteints de schizophrénie, ce que l'on pourrait attendre compte tenu du déclin général des capacités perceptives habituellement observé dans ces deux populations en comparaison avec de jeunes sujets sains. De façon similaire, un fort facteur visuel pourrait apparaître chez les joueurs de jeux vidéo d'action, qui, au contraire, présentent souvent de meilleures performances perceptives comparés à des non-joueurs. Ainsi, des personnes âgées en bonne santé, des patients atteints de schizophrénie, et des joueurs de jeux vidéo d'action ont été testés avec une batterie de tâches visuelles, telles que la détection de contraste et la discrimination d'orientation. Les corrélations entre différentes tâches étaient en général faibles dans ces trois populations, au même titre que dans les groupes contrôle. Ces résultats plaident donc contre l'émergence d'un fort facteur visuel dans ces populations.

Alors que des tâches similaires sont généralement supposées reposer sur des mécanismes neuronaux similaires, les performances à diverses tâches visuelles ne sont en réalité que peu reliées les unes aux autres. En d'autres termes, une performance n'est pas généralisable à plusieurs tâches visuelles. Les résultats présentés dans cette thèse mettent ainsi en évidence la pertinence d'une approche des différences individuelles pour élucider la multidimensionnalité de la structure visuelle.

Mots-clés : perception visuelle, illusions d'optique, différences individuelles, facteurs communs, structure, variabilité, fiabilité, analyses corrélationnelles

1. Preface

In this thesis, I present the results of a series of projects conducted over the course of my doctoral studies, in which I investigated individual differences in visual perception. A list of studies is included below and annotated with my personal contributions. At the time of writing, six articles were published, two are submitted (under review), and two projects are ongoing. Published articles and preprints of the submitted manuscripts are provided in the Appendix with permission from the copyright holders.

1. **Cretenoud, A. F.**, Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, 19(14):12, 1–21. (Appendix A)
Designed and conducted two experiments, analyzed and interpreted the data, wrote the manuscript.
2. **Cretenoud, A. F.**, Grzeczowski, L., Bertamini, M., & Herzog, M. H. (2020). Individual differences in the Müller-Lyer and Ponzo illusions are stable across different contexts. *Journal of Vision*, 20(6):4, 1–14. (Appendix B)
Designed and conducted the experiment, analyzed and interpreted the data, wrote the manuscript.
3. **Cretenoud, A. F.**, Francis, G., & Herzog, M. H. (2020). When illusions merge. *Journal of Vision*, 20(8):12, 1–15. (Appendix C)
Designed and conducted one experiment, analyzed and interpreted the data, wrote the manuscript.
4. **Cretenoud, A. F.**, Grzeczowski, L., Kunchulia, M., & Herzog, M. H. (submitted). Individual differences in the perception of visual illusions are stable across eyes, time, and measurement methods. (Appendix D)
Designed two experiments, conducted one experiment, analyzed and interpreted the data, wrote the manuscript.
5. Shaqiri, A., Pilz, K. S., **Cretenoud, A. F.**, Neumann, K., Clarke, A., Kunchulia, M., & Herzog, M. H. (2019). No evidence for a common factor underlying visual abilities in healthy older people. *Developmental Psychology*, 55(8), 1775–1787. (Appendix E)
Participated in the data analysis and writing of the manuscript.

6. **Cretenoud, A. F.**, Favrod, O., Ramos da Cruz, J., Shaqiri, A., Garobbio, S., Gordillo, D., Berdzenishvili, E., Brand, A., Roinishvili, M., Chkonia, E., & Herzog, M. H. (ongoing). Exploring the heterogeneity of visual perception in schizophrenia.
Analyzed and interpreted the data, wrote the manuscript.
7. **Cretenoud, A. F.**, Barakat, A., Milliet, A., Choung, O.-H., Bertamini, M., Constantin, C., & Herzog, M. H. (submitted). How do visual skills relate to action video game performance? (Appendix F)
Designed the experiment, analyzed and interpreted the data, wrote the manuscript.
8. Rashal, E., **Cretenoud, A. F.**, & Herzog, M. H. (2020). Perceptual grouping leads to objecthood effects in the Ebbinghaus illusion. *Journal of Vision*, 20(8):11, 1–15. (Appendix G)
Participated in the design of the experiments, conducted the experiments, and participated in the data analysis and interpretation.
9. Jastrzębowska, M. A., **Cretenoud, A. F.**, Ozkirli, A., Draganski, B., Herzog, M. H. (ongoing). Is there a neural common factor in illusion magnitude? (Appendix H)
Participated in the design of the experiment, analysis and interpretation of the behavioral data.
10. Bertamini, M., **Cretenoud, A. F.**, & Herzog, M. H. (2019). Exploring the Extent in the Visual Field of the Honeycomb and Extinction Illusions. *i-Perception*, 10(4), 1–19. (Appendix I)
Analyzed and interpreted the data, participated in the writing of the manuscript.

2. Introduction

2.1. Visual illusions

2.1.1. What is a visual illusion?

Illusions, may they be visual, haptic, or auditory, are known and have been studied for centuries. Perceptual illusions refer to incorrect perceptions, i.e., cases, in which perception does not match reality (Todorović, 2014, 2020). There are two distinct cases, in which perception is illusory. First, perception is illusory when two objects appear different, while they are physically equal (Figure 1A). Second, an illusory percept may also arise when two objects are different but perceived equally (Figure 1B).

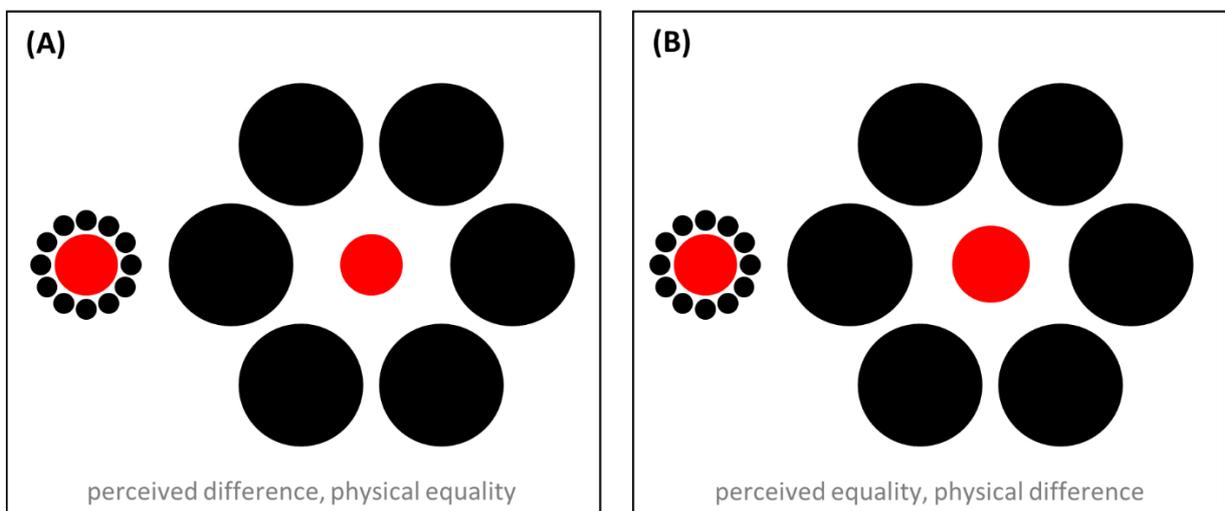


Figure 1. Classic representation of the Ebbinghaus illusion. Targets are shown in red, flankers in black. (A) The left target, which is surrounded by small flankers, looks larger compared to the right target, which is surrounded by large flankers. However, both targets are the same physical size. (B) In the same context, two targets of different sizes look the same or similar. Inspired from Todorović (2014).

The Ebbinghaus illusion is a well-known example of a visual illusion, in which a target looks smaller when it is surrounded by large flankers or looks bigger when it is surrounded by small flankers (Figure 1). Two accounts were suggested to explain the Ebbinghaus phenomenon: a cognitive size contrast mechanism or low-level contour interactions. The former poses that the sizes of the target and surrounding flankers are directly compared (e.g., Coren & Enns, 1993; Coren & Miller, 1974; Vuk & Podlesek, 2005), while the latter proposes that a stretching or contraction of the target is perceived following an attraction or repulsion, respectively, of the contours of the target by the contours of the flankers (e.g., Jaeger & Klahs, 2015; Schwarzkopf & Rees, 2013; Sherman & Chouinard, 2016; Todorović & Jovanović, 2018; Weintraub & Schneck, 1986). The strength of the illusion is a function of several

factors, such as the number of surrounding flankers (e.g., Jaeger & Klahs, 2015), the target-flankers distance (Massaro & Anderson, 1971), or the similarity in shape between the target and surrounding flankers (Coren & Miller, 1974), rendering the two accounts hard to disentangle.

Similarly, the magnitude of the Ebbinghaus illusion may be affected by grouping effects. In a side project (Appendix G), I manipulated the objecthood, i.e., the degree to which an object is a cohesive perceptual entity, to examine the effect of grouping on the illusion. The Ebbinghaus illusion was tested with square targets and flankers, which were made of corners or sides, i.e., gaps of varying sizes were introduced in the square shapes. As expected, objecthood was reduced with increasing gap sizes, i.e., the comparison process was less efficient when larger gaps were introduced in the objects, resulting in decreased illusion magnitudes. In addition, participants were asked to subjectively report the perceived shape of all objects tested. Interestingly, the perceived shape of the objects did not correlate with the illusion magnitudes, suggesting a dissociation between conscious (i.e., reports of perceived shape) and unconscious (i.e., illusion magnitudes) percepts.

While visual illusions are usually considered as relevant, if not essential, to vision studies (e.g., Rose, 2018; Shapiro, 2018; Shapiro & Todorović, 2017; Todorović, 2018), the usefulness of visual illusions as a research field was sometimes questioned. For instance, Braddick (1972, 2018) nourished the debate by defining visual illusions as an irrelevant category of science. The author did not dismiss each illusion to be a valuable subject of research in its own, i.e., a visual illusion may be informative about visual processing. He rather argued about the absence of a clear benefit in considering visual illusions as a whole, because the mechanism underlying one specific illusion does not necessarily underlie other illusions. According to Braddick, the fact that objects are misperceived is not a unifying principle. Instead, the mechanisms of perception, which act both in the veridical and non-veridical (i.e., illusory) world, were suggested to be a more relevant subject of study. However, as highlighted by Todorović (2018), many more fields than just visual illusions should be invalidated for the same reason. For example, a treatment against one disorder may be inefficient against another disorder. Still, both disorders are studied in a larger field of research commonly called medicine.

Despite this debate, illusions are usually used as a tool to explore the complex nature of the brain constructs and functions. In addition to be a common subject of fascination in young as well as older generations, illusions have, as other visual phenomena, provided key insights in neuroscience. Likewise, visual illusions may help investigating cultural differences (e.g., Caparos et al., 2012), age and gender differences (e.g., Doherty et al., 2010; Phillips et al., 2004), as well as mental disorders such as autism (e.g., Burghoorn et al., 2020; Chouinard et al., 2016) or schizophrenia (e.g., Bressan & Kramer, 2013; Grzeczowski et al., 2018). While plenty of theoretical and practical approaches are used in the

field of visual illusions, Shapiro (2018) suggested that different visual illusions may have more in common than expected. The application of sophisticated models may reveal relationships that have no *a priori* reason to exist, between an illusion of contrast and another of size, for example.

2.1.2. Taxonomies of visual illusions

While similar illusions are usually assumed to rely on similar neural mechanisms, vision scientists have attempted to classify visual illusions according to their communalities, such as brightness or spatial features (e.g., Binet, 1895; Coren et al., 1976; Ninio, 2014; Piaget, 1961).

Some taxonomies distinguished between innate and acquired illusions. For example, Binet (1895) observed weaker Müller-Lyer illusion magnitudes in older compared to younger children. The author therefore suggested that the magnitude of innate illusions, such as the Müller-Lyer illusion, decreases with age, whereas the magnitude of acquired illusions increases with age. A weight illusion was suggested as an acquired illusion, since adults were observed to be more susceptible than children (Dresslar, 1894).

Piaget (1961, 1963) developed a concept called *centration*, according to which an object in the center of the visual field is overestimated in size compared to surrounding objects. The author suggested that children make more eye movements when they get older, i.e., their visual exploration is more systematic with multiple *centrations*. Hence, the susceptibility to some illusions, which were called primary or *type I* illusions, decreases with age. In contrast, the perception of depth becomes more accurate when children grow, giving rise to increasing susceptibility to other illusions, which were called secondary or *type II* illusions. The Müller-Lyer and Ponzo illusions were first thought to belong to the type I (Binet, 1895) and II (Leibowitz & Heisel, 1958; but see Pollack, 1964) illusions, respectively.

However, the difference between both illusion types may lie in the complexity of the stimulus rather than in the illusion itself. For example, the Ponzo illusion magnitude was stronger when embedded in a photograph (with perspective cues) than when embedded in an abstract background (Leibowitz et al., 1969). Wagner (1977) tested children with real-world (photograph) and abstract (geometrical) Ponzo illusions and observed increasing and decreasing illusion magnitudes with age, respectively. Therefore, he interpreted the two-factor theory of Piaget as depending on the richness of context in which illusions are embedded and the participants' age (i.e., abstract and real-world illusions are primary and secondary types of illusions, respectively).

2.2. Individual differences

Traditionally, scientific studies test several individuals under the same conditions and compare the average performance across conditions, i.e., the differences between individuals are usually considered as noise in the measurements. However, it is valuable to study the pattern of variations between individuals (i.e., the individual differences) and to examine systematic differences (Mollon et al., 2017; Peterzell, 2016; Wilmer, 2008), as they may give insights into the underlying visual processing. Hence, rather than focusing on the performances *per se*, the study of individual differences investigates how much individuals vary in performance and whether this variability is reliable across tasks (or across conditions).

When studying individual differences, scientists typically rely on correlational analyses to examine whether different tasks rely on a similar mechanism. A common underlying source of variance is expected to come out as strong correlations and can be estimated from statistical techniques such as a principal component analysis (PCA) or factor analysis (FA).

As proposed in Ward et al. (2017), the visual structure may be described as a mono-, bi-, or multifactorial entity. First, in the monolithic view, an individual who performs better in one visual task compared to other individuals is assumed to also perform better in other visual tasks, suggesting that there is a strong common factor underlying vision. Such common factors are ubiquitous in everyday life. For example, it is widely held that there is a common factor, *g*, for intelligence (Spearman, 1904a). This factor is not measurable *per se* but is inferred from several indicator variables, such as the Wechsler scale (Wechsler, 2003). Note, however, that intelligence was recently described as a general factor as well as a collection of more specific abilities (Mackintosh, 2011). In metacognition and somatosensation, there are common factors between different modalities, for example, between touch and audition, which highly correlate relying on the same genetic factor (Frenzel et al., 2012). Faivre et al. (2017) showed that participants with high performance in one metacognitive modality are likely to show high performance in other metacognitive modalities. Similarly, there seems to be a strong common factor for cognition in healthy aging, i.e., cognitive abilities are reliably affected with age (e.g., Kiely & Anstey, 2017; Lindenberger & Ghisletta, 2009). While age-related changes of five cognitive functions, such as perceptual speed and reasoning skills, were indeed reported to strongly correlate, sensory functioning, e.g., vision and hearing, was suggested to significantly predict the individual differences in cognitive aging (Baltes & Lindenberger, 1997; Ghisletta & Lindenberger, 2005; Lindenberger & Baltes, 1994). In analogy, a strong common factor for vision may be expected.

Second, vision may be divided according to pairs of anatomically distinct structures or pathways, such as dorsal versus ventral (e.g., Goodale & Milner, 1992), or magnocellular versus parvocellular streams (e.g., Livingstone & Hubel, 1988). In such a bifactorial structure, strong correlations are expected between tasks tapping into one stream, while two tasks tapping into distinct streams may only weakly correlate.

Last, the visual structure may be highly multifactorial, i.e., there may be a broad range of specific visual skills, which only weakly relate to each other. Importantly, the three alternatives (i.e., a mono-, bi-, or multifactorial structure) are not mutually exclusive, e.g., there may be several specific factors, which complement a more general common factor.

2.2.1. Vision in general

A few studies supported a monolithic alternative for visual perception. For example, Halpern et al. (1999) tested 20 subjects with seven basic visual tasks, such as orientation discrimination and contrast sensitivity. A measure of overall visual ability, computed as a composite score across all tasks, significantly differed across subjects, indicating strong individual differences. A PCA revealed that a non-negligible proportion of the variance in the tasks (30%) was explained by the first component, even though not all between-task correlations were strong. The authors suggested that such a common factor for visual performance may be related to structural differences across individuals, e.g., in the lateral geniculate nucleus (LGN) or primary visual cortex (V1) volume.

Other studies emphasized the second alternative, which poses that the visual structure is fractionated according to pairs of anatomical pathways. For example, a few – but not all (e.g., Goodbourn et al., 2012) – studies reported two common factors, which are consistent with the activity of magnocellular and parvocellular systems (Dobkins et al., 2000; Peterzell & Teller, 1996; Simpson & McFadden, 2005; Ward et al., 2017). Similarly, two factors for visual stability and visual ability were suggested to explain change detection performance (Andermane et al., 2019).

However, most studies rather suggested a sparse factorial structure for individual differences in vision, arguing against the existence of a strong common factor (for reviews, see Bosten et al., 2017; Peterzell, 2016; Tulver, 2019). For example, Cappe and colleagues (2014) only observed weak correlations between the performances of 40 subjects in six basic visual tasks, such as visual acuity and contrast detection. In addition, the authors reported a first component of a PCA explaining only 34% of the variance in the data (note, however, that this is more than the 30% of variance explained by the first principal component in Halpern et al., 1999). Hence, they suggested that the structure underlying vision is multifactorial rather than monolithic.

Even studies that had a narrowly defined hypothesis by including several tasks that tap into a specific functional ability or theoretical construct of perception have often not succeeded in finding evidence to support the existence of a stable factor in perception (Tulver, 2019). For example, Bosten and Mollon (2010) claimed that a single mechanism is unlikely to account for simultaneous contrast across different dimensions (e.g., luminance, color, and orientation). Likewise, no strong common factor was reported in oculomotor tasks (Bargary et al., 2017), bistable perception (Brascamp et al., 2018, 2019; Cao et al., 2018; see also Wexler et al., 2015), local-global processing (Chamberlain et al., 2017; Milne & Szczerbinski, 2009), and face recognition (Verhallen et al., 2017; see also Čepulić et al., 2018).

Many studies suggested that the visual structure is best represented by several specific factors. For example, a complex factor structure was suggested to underlie contrast sensitivity (Peterzell, 2016; Peterzell, Scheffrin, et al., 2000). Indeed, only weak correlations were observed between three ranges of scotopic sensitivities (0.2-0.4 c/deg, 0.4-1.2 c/deg, and 1.2-3.0 c/deg), while sensitivities strongly correlated within each range. Similarly, individual differences in hue scaling (e.g., Emery et al., 2017a, 2017b), color matching (Webster & MacLeod, 1988), luminance contrast sensitivity (Dobkins et al., 2000; Peterzell, Chang, et al., 2000), stereopsis (Peterzell et al., 2017), and in the effects of priors (Tulver et al., 2019) were suggested to rely on several specific factors.

While there seems to be only weak evidence for a strong and unique common factor in vision, a weak general and several more specific factors may coexist. For example, Bosten and colleagues (2017) tested more than a thousand subjects with 25 visual and auditory measures. A unique common factor accounted for about 20% of the total variance in the dataset, while 57% of the total variance was explained by eight more specific (and rotated) factors, such as a factor for stereo acuity and another related to oculomotor speed. As between-measure correlations were almost uniformly positive, the authors both suggested evidence for a general – albeit weak – factor underlying perception and for factors specific for independent perceptual skills.

2.2.2. Visual illusions

As in vision in general, several studies of the structure underlying visual illusions showed mixed results. A couple of studies suggested that there is a strong common factor for visual illusions. For example, Thurstone (1944) found a factor underlying geometric illusions, which however, did not show strong loadings on other classes of illusions such as brightness or size-weight illusions. Similarly, a factorial study of 70 perceptual tests revealed a single factor for illusions (Roff, 1953). Since performance in the simple estimation of line length correlated with illusion magnitudes, the author concluded that the factor was underlying length perception tasks in general (i.e., not only illusion magnitudes).

Other studies proposed that there are several subclasses of illusions. For example, Coren and colleagues (1976) tested 45 illusion configurations and computed a first-order factor analysis, which resulted in five factors for illusions of shape and direction, size contrast, overestimation, underestimation, and frame of reference illusions. In addition, the authors reported a second-order factor solution, which was computed from the intercorrelations between the five first-order factors. Two second-order factors were suggested to underlie illusions of extent and illusions of shape or direction.

Similarly, Taylor (1974) computed a factor analysis on 21 measures of visual illusions, which revealed three distinct factors. The first factor was mainly driven by several variants of the same illusion, namely the Poggendorff illusion, while the other two factors were mainly driven by different illusions (distortions of parallelism and length judgments). Note that ten illusions were not strongly accounted for by any of these factors. In a follow-up study including 18 illusion measures (Taylor, 1976), a four-factor solution was found with the first three factors that were similar to the ones found in the first study. Note, however, that some high loadings did not replicate between the two studies and that the average between-illusion correlation in both studies was low ($r = 0.16$). When adding 12 perceptual, cognitive, and temperament measures to the factor analysis, the illusions still did not group into a general factor but were split into four factors. Contrary to Thurstone (1944) and Roff (1953), Taylor therefore showed that illusions do not cluster on a single dimension when embedded into a heterogeneous collection of tasks, suggesting that illusions themselves are heterogeneous perceptual tasks. Similarly, Robinson (1968) suggested that illusions are too heterogeneous to be explained by a single mechanism. The author made a clear distinction between illusions, which involve a misperception of length or size, and distortions, which imply perceptual bending of lines.

More recently, Grzeczowski and colleagues (2017) measured the susceptibility to visual illusions in 144 participants. Three spatial illusions, namely the Ebbinghaus, Müller-Lyer, and Ponzo hallway (also called corridor) illusions, and three non-spatial illusions, namely, the contrast, White, and Tilt illusions (Figure 2A), were tested with an adjustment procedure, in which participants had to reach the point of subjective equality in size, contrast, or orientation, by moving the computer mouse on the horizontal axis. For example, participants had to adjust the size of the right target in the Ebbinghaus illusion so that they perceived it to be equal in size to the left target.

While two trials of the same illusion significantly correlated (i.e., suggesting reliable measurements), between-illusion correlations were weak and non-significant, except between the Ebbinghaus and Ponzo hallway illusions (Figure 2B). The authors reported a first component of a PCA accounting for about 24% of the variability in the data, suggesting that there is no strong common factor underlying

visual illusions. A rank analysis, which examines the relationship between the individual ranks to each visual illusion, did not significantly differ from chance, suggesting that an individual with a strong susceptibility to one visual illusion does not necessarily show strong susceptibilities to other visual illusions.

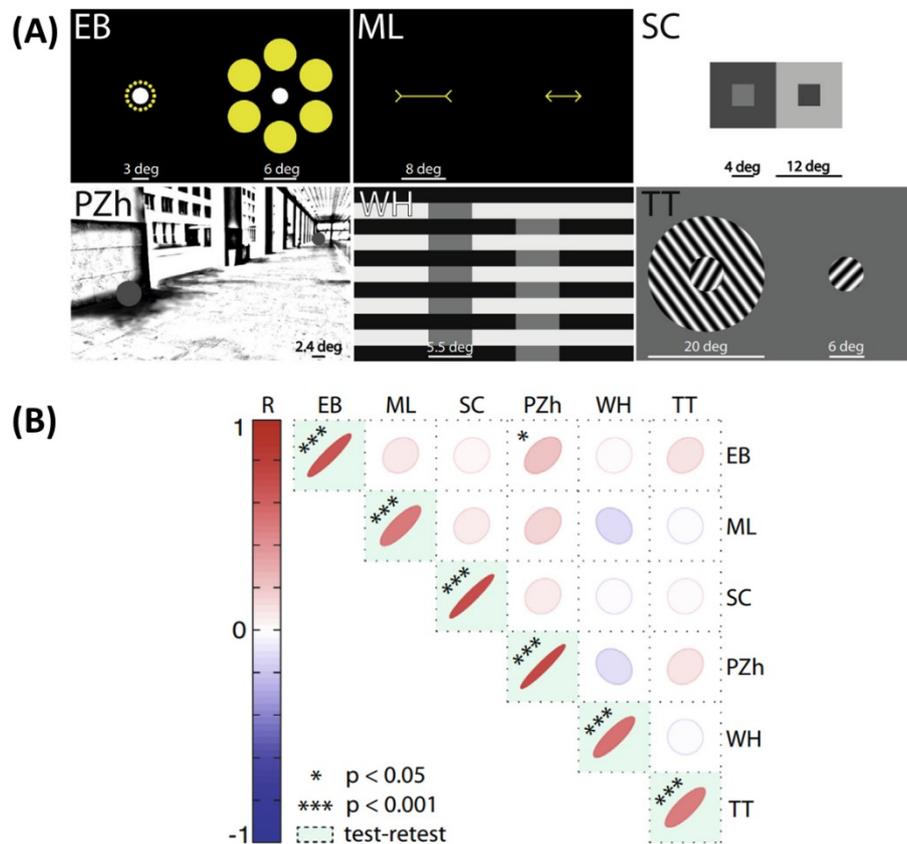


Figure 2. (A) Grzeczowski et al. (2017) used an adjustment procedure to measure the susceptibility to six visual illusions, namely the Ebbinghaus (EB), Müller-Lyer (ML), contrast (CS), Ponzo hallway (PZh; also called corridor), White (WH), and Tilt (TT) illusions. (B) Pairwise correlations plotted as iso-probability contours of the joint histograms with narrower ellipses indicating stronger correlations. The color scale from blue to red reflects effect sizes from $r = -1$ to $r = 1$. Only weak and non-significant between-illusion correlations were observed, except between the Ebbinghaus and Ponzo hallway illusions (triangle). However, strong test-retest reliabilities were reported (diagonal). A principal component analysis (PCA) revealed a first component explaining about 24% of the variability, which suggests that there is no strong common factor underlying these six illusions. Adapted from Grzeczowski et al. (2017) and reprinted with permission from Elsevier.

In a second experiment, the authors examined the relationship between the susceptibility to visual illusions and personality traits, by testing two variants of the Ebbinghaus illusion, a Müller-Lyer illusion, and four variants of the Ponzo illusion. As in the first experiment, weak between-illusion correlations were observed, except between the Ebbinghaus and Ponzo hallway illusions and between different variants of the same illusion (e.g., between the two variants of the Ebbinghaus illusion). Significant correlations were also observed between cognitive disorganization, mental imagery and most variants of the Ponzo illusion, suggesting that there is an association between the susceptibility to the Ponzo

illusion and personality traits. Note, however, that these significant associations did not survive Bonferroni correction.

Similarly, Axelrod et al. (2017) measured the susceptibility to a contrast and four spatial illusions and observed only weak between-illusion correlations, except between the Ebbinghaus and Müller-Lyer illusions. The authors computed a PCA, which resulted in a first component explaining 36% of the variance in the data and mainly capturing variability in the spatial illusion magnitudes. In addition, they correlated behavioral and neuroanatomical data and observed a strong correlation between spatial illusion magnitudes and the grey matter density in the parahippocampal cortex (but not in other brain regions). Hence, the authors suggested that the perception of visual illusions is strongly related to visuospatial and scene-integration processing, which happen in the parahippocampal cortex and hippocampus.

In the same line of research, individual variability in the perception of visual illusions was recently suggested to be related to idiosyncrasies in the visual cortex (Chen et al., 2020; Moutsiana et al., 2016; Schwarzkopf et al., 2011; Schwarzkopf & Rees, 2013). For example, the Ebbinghaus and Ponzo hallway illusion magnitudes correlated negatively with V1 surface area (Schwarzkopf et al., 2011; Schwarzkopf & Rees, 2013), i.e., individuals with larger V1 surface area showed in general weaker susceptibilities to these two illusions. Note, however, that the authors did not find any significant association between the Ebbinghaus and Ponzo hallway illusions, unlike in Grzeczkowski et al. (2017). In a side project (Appendix H), I tried replicating the association between V1 surface area and illusion magnitudes. However, only weak correlations were observed between the size of early visual areas and illusion magnitudes.

While there seems to be no strong common factor for visual illusions in healthy young adults, it may be that special populations, e.g., individuals suffering from pathologies or healthy elderly, show altered but reliable susceptibilities to visual illusions. For example, patients with schizophrenia were reported to have an increased and decreased susceptibility to the Müller-Lyer and Ponzo illusion, respectively, but similar susceptibility to the Ebbinghaus illusion compared to healthy controls (e.g., Kantrowitz et al., 2009; for a review, see King et al., 2017). However, when testing a battery of several visual illusions, only weak between-illusion correlations were observed in patients with schizophrenia (Grzeczkowski et al., 2018), as in healthy controls. These results suggest that there is no strong common factor for visual illusions in patients with schizophrenia, i.e., patients do not present a reliable pattern of illusory misperception. Note that the authors did not observe any significant differences in terms of susceptibility to visual illusions between patients with schizophrenia and healthy controls, except for a contrast illusion, where patients were more susceptible than healthy controls.

Enhanced abilities in discriminating local details were reported in individuals with autism spectrum disorder (e.g., Mottron et al., 2006), which suggests that these individuals may show specific patterns of illusion susceptibilities. Chouinard et al. (2016) tested 13 illusions in a large group of adults with various levels of autistic traits. A PCA suggested a five-factor solution accounting for more than 57% of the variability in the data, while only 15% of the variability was explained by the first factor. Importantly, autistic traits, as measured with the autism spectrum quotient (AQ), significantly (but negatively) correlated with only one factor, which highly loaded on the susceptibility to the Shepard's tabletops and square-diamond illusions. Because only two illusions were affected (and because these two illusions do not seem to rely on local-global processing), it seems that there is no strong common factor for the perception of visual illusions in relationship with autistic-like traits.

2.3. Perceptual learning and its hyperspecificity

A multifactorial structure is not unlikely to underlie visual illusions, as was reported for perceptual learning, i.e., the ability to improve perception with training. Indeed, one of the hallmarks of perceptual learning is its hyperspecificity. For example, when a vernier offset discrimination task is trained, performance improves. However, learning transfers only for the same stimulus rotated up to 10° (Ball & Sekuler, 1987; Fahle & Morgan, 1996; Grzeczowski, Cretenoud, et al., 2017; Schoups et al., 1995; Spang et al., 2010). In addition, perceptual learning was also shown to be specific to the contrast (Sowden et al., 2002; Yu et al., 2004), motion direction (Ball & Sekuler, 1982, 1987), and spatial frequency (Berardi & Fiorentini, 1987).

In a project conducted during my Master thesis and later published during my doctoral studies (Grzeczowski, Cretenoud, et al., 2017; not included in the preface), I showed that specificity also holds true beyond decision making. It is usually assumed that perceptual learning occurs within sensory areas or when decisions are made. Hence, further processes, such as motor ones, are thought to be independent of the learning process. However, while performance in a bisection task with an adjustment procedure improved with training, there was no transfer of learning to the same task when tested with a button press condition. The contrary was also true, suggesting that perceptual learning is even specific to the motor response type. In another project conducted during my Master thesis and later published during my doctoral studies (Grzeczowski et al., 2019; not included in the preface), I observed that perceptual improvements only partially transfer from an adjustment to a button press condition when using a double training procedure, in which two stimuli with distinct features are trained in parallel (double training has been shown to release the feature-specificity of perceptual learning; for a review see Herzog et al., 2017). These results suggest that perceptual and motor signals are encoded together.

2.4. Aims of the present work

Rather than a strong common factor, the structure underlying visual illusions seems to be multifactorial. If perceptual learning plays a role in the perception of visual illusions, i.e., if our personal experience and environment shape the way we perceive illusions, then weak correlations may even be observed between illusions that look very similar to each other, e.g., two variants of the same illusion differing in orientation only. Here, in a first series of studies (Studies 1 to 3), I systematically examined whether factors for illusions are hyperspecific, as in perceptual learning. Next, I examined whether individual differences in the perception of visual illusions are stable regardless of viewing condition, repeated measures across a month, and how they are measured (adjustment vs. binary responses; Study 4). Last, I wondered whether a strong common factor for vision in general emerges in special populations, e.g., with the general decline observed when aging and in clinical populations or with the enhanced visual abilities usually observed in video game players (Studies 5 to 7).

3. General methods

In this section, I give an overview of the methods used in the different studies. Note, however, that each study has its specificities, which are later mentioned in the respective sections. For detailed descriptions, please refer to the papers or preprints (Appendix A to Appendix I).

3.1. Adjustment procedure

To measure the susceptibility to visual illusions, an adjustment procedure was used. Participants had to adjust the size, length, position, orientation, or shade of grey, of a target to match the size, length, position, orientation, or shade of grey, of a reference on the screen, respectively. For example, to measure the susceptibility to the Ebbinghaus illusion, participants had either to adjust the size of the target surrounded by large flankers so that they perceived it to be the same size as the target surrounded by small flankers, or to adjust the size of the target surrounded by small flankers so that they perceived it to be the same size as the target surrounded by large flankers. In other terms, participants were asked to adjust a target until reaching the point of subjective equality (PSE) with a reference. The target was adjusted by moving the mouse on the horizontal or vertical axis, except in one experiment of Study 1 (luminance-unspecific), where participants used the left and right arrow keys on a keyboard to make the adjustments. Participants validated their adjustments by clicking on the left button of the computer mouse.

The distance to the screen was approximately 60 cm. Prior to the experiments, the color look-up tables of the monitor were linearized and calibrated. Participants were tested in a quiet experimental room with controlled light conditions.

The experimenter first explained the task to the participants, who completed some warming up trials to familiarize with the task. Each condition was then presented twice in a random order, except in one experiment of Study 1 (luminance-unspecific), where each condition was presented four times in a random order. No feedback was provided and there was no time constraint. The initial size of the target was pseudo randomly chosen at each trial (i.e., so that the target did not overlap with other elements on the screen and stayed within the screen limits). The experimenter stayed in the experimental room during the whole experiment. Participants were asked to perform the task relying on their percepts only.

3.2. Data analysis

Statistical results of hypothesis testing were usually reported both without correction for multiple comparisons and with the very conservative Bonferroni correction, since I sometimes aimed at null results and sometimes not. For example, in the case of null results, non-significant correlations are expected, even without correcting the alpha level for multiple comparisons. On the contrary, a correlation coefficient that is significant after correcting for multiple comparisons speaks in favor of a true effect. The alpha level for statistical significance was set to 0.05 as usually.

According to Cohen (1988), correlation coefficients of 0.1, 0.3, and 0.5 are considered as small, medium, and large effect sizes, respectively. According to a meta-analysis by Gignac and Szodorai (2016), effect sizes of 0.1, 0.2, and 0.3 are however considered as small, medium, and large, respectively.

3.2.1. Test-retest reliability

When applicable, I assessed test-retest reliability, i.e., the within-individual variation across trials, by computing an intraclass correlation (ICC) between the several measures of each variable (or condition). First introduced by Fisher (1992) as an extension of the Pearson correlation coefficient, the concept of ICC was later developed as a measure of reliability within a class of data (Bartko, 1966; Shrout & Fleiss, 1979) rather than between different classes (the correlation between two different classes of data is usually computed as a Pearson correlation). Intraclass correlations are based on the analysis of variance (ANOVA) and therefore assume normally distributed data. In short, an ICC is computed as a ratio between the variance of interest (e.g., between-individual variance) and the total variance, i.e., the variance of interest and unwanted variance (e.g., within-individual variance or instrumental variation). The larger the ICC coefficient, the more reliable the data.

Several types of ICCs were developed to fit different experimental situations (e.g., Koo & Li, 2016; Liljequist et al., 2019). Here, intraclass correlations of type (3,1) or $ICC_{3,1}$ were computed. The first subscript refers to a two-way mixed effects model (i.e., model 3), in which a random sampling of participants is assumed, while biases are assumed to be fixed (i.e., the only measure of interest is the variable extracted, which was measured several times to assess test-retest reliability). The second subscript indicates the type of selection used, i.e., each data point either represents a single measurement (i.e., type 1) or an average across several measurements (i.e., type k).

The test-retest reliabilities were in general significant with medium to large effect sizes (see Tables 1, 2, 3, 4, and 5A), suggesting reliable measurements. Hence, the several trials of each condition were

averaged for further analysis. Note, however, that for the sake of time, variables were not always measured several times in Studies 5 to 7.

3.2.2. Illusion magnitudes

As a measure of illusion magnitude, the reference of each condition was subtracted from the mean adjustment. Hence, the illusion magnitude is expressed as a difference (or sometimes as a proportion, i.e., the difference was further divided by the reference) compared to the reference with positive and negative values indicating over- and underadjustments, respectively.

3.2.3. Data transformation and outlier removal

The normality assumption was tested by computing a Shapiro-Wilk test for each condition (or each variable).

In Studies 1 to 4, most distributions were normal (hence, data was not standardized). Outliers were detected using a modified z-score, which is more robust than the commonly used z-score, because it makes use of the median and median absolute deviation (*MAD*) instead of the mean and standard deviation (*SD*), respectively. Absolute modified z-scores larger than 3.5 were considered as outliers and removed for further analysis, as suggested by Iglewicz and Hoaglin (1993).

In Studies 5 to 7, most distributions violated the normality assumption. Hence, each distribution was rescaled to approximate a normal distribution. First, the data distribution was shifted to positive values only. Second, as above, modified z-scores were computed, and outliers removed according to a 3.5 criterion. Third, the λ exponent of a Tukey power transformation was optimized to maximize normality according to the Shapiro-Wilk test. Fourth, including the previously removed outliers, data was transformed using the Tukey transformation with the optimized λ parameter. Last, data was standardized, and outliers removed based on modified z-scores, as above.

3.2.4. Between-condition correlations

Pearson correlations were computed between each pair of conditions (or pair of variables in Studies 5 to 7). The observed between-condition correlations were underestimated because of measurement errors (Spearman, 1904b; Wang, 2010), which are reflected in the test-retest reliabilities. To account for these measurement errors, which put an upper limit on the between-condition correlations, disattenuated correlations were also computed as:

$$r_{xy'} = \frac{r_{xy}}{\sqrt{r_{xx}r_{yy}}}$$

where r_{xy} , is the disattenuated relationship between x and y, r_{xx} and r_{yy} are the test-retest reliabilities of the x and y conditions, and r_{xy} is the attenuated (i.e., non-disattenuated) correlation coefficient between x and y (e.g., Osborne, 2003; Wang, 2010). Note that disattenuated correlations were not computed in Studies 5 to 7 because not all variables were tested more than once (i.e., test-retest reliabilities were not available for all variables).

3.2.5. Factor analysis

To explore the structure underlying a dataset, exploratory factor analysis (EFA) was computed using the guidelines outlined in Preacher et al. (2013). Factors were either extracted with a common factor analysis, which identifies factors that reflect the common variance only, or with a principal component analysis, which identifies factors that summarize most of the original information (i.e., common and unique variances). Oblique rotations allow factors to correlate, while uncorrelated factors result from orthogonal rotations. Since there was no reason to preclude correlated factors from the datasets, an oblique rotation (e.g., promax) was used (except in Study 5). Note that both oblique and orthogonal rotations produce very similar results when the factors are uncorrelated (e.g., Costello & Osborne, 2005).

4. Results

4.1. Are factors for visual illusions hyperspecific?

Only weak correlations were previously observed between the susceptibility to different visual illusions, i.e., a high susceptibility to one illusion does not inevitably imply a high susceptibility to other illusions (Grzeczowski, Clarke, et al., 2017). In three studies (Studies 1 to 3), I investigated whether illusions are hyperspecific, as in perceptual learning.

4.1.1. Study 1: Luminance-unspecific¹

Hamburger and colleagues (2007) previously tested ten visual illusions under different luminance conditions. The authors observed high correlations between the different luminance conditions for each illusion, suggesting that factors for illusions are not luminance-specific. However, the authors did not explore the relationship between different illusions. Here, I specifically wondered whether previous results, i.e., weak between-illusion correlations (e.g., Grzeczowski, Clarke, et al., 2017), could be replicated. As strong within-illusion correlations were observed in the dataset previously published by Hamburger and colleagues (2007), it is unlikely that methodological flaws prevent strong effects from showing up (i.e., if between-illusion correlations are strong, they will likely show up here).

Twenty-one participants were tested with ten visual illusions (Figure 3): bisection (BS), Delboeuf (DB), Ebbinghaus (EH), Hering (HN), horizontal-vertical (HV), Judd (JD), Müller-Lyer (ML), Poggendorff (PD), Ponzo (PZ), and Zöllner (ZN). Each illusion was tested with four different luminance conditions (Figure 4): a high luminance contrast condition (50%; Lum), a low luminance contrast condition (10%; LumLow), a “red-green” [L – M] isoluminant condition (Iso) and a “blue-yellow” [S – (L + M)] isoluminant condition (IsoS). Participants had to reach subjective equality in length (BS, HV, JD, ML, PZ), size (DB, EB), curvature (HN), position (PD), or orientation (ZN) by pressing the right and left arrow keys on the keyboard. Each illusion was presented under four different luminance conditions, four trials each (160 trials in total). The conditions were randomly intermixed.

Attenuated (upper triangle) and disattenuated (lower triangle) correlation coefficients are reported in Table 1. Within each illusion, the different luminance conditions were strongly correlated ($r > 0.3$ for all within-illusion correlations), as already observed in Hamburger et al. (2007). In contrast, only 27% (193 out of 720) and 55% (395 out of 720) of the attenuated and disattenuated between-illusion

¹ This section is based on Cretenoud, A. F., Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, 19(14):12, 1–21. (Appendix A; Experiment 2)

correlations showed $r > 0.3$, respectively. Hence, within-illusion correlations were strong, while between-illusion correlations were weaker. However, the bisection illusion magnitudes strongly correlated with the horizontal-vertical illusion magnitudes. In fact, the horizontal-vertical illusion is a bisection illusion rotated by 90° .

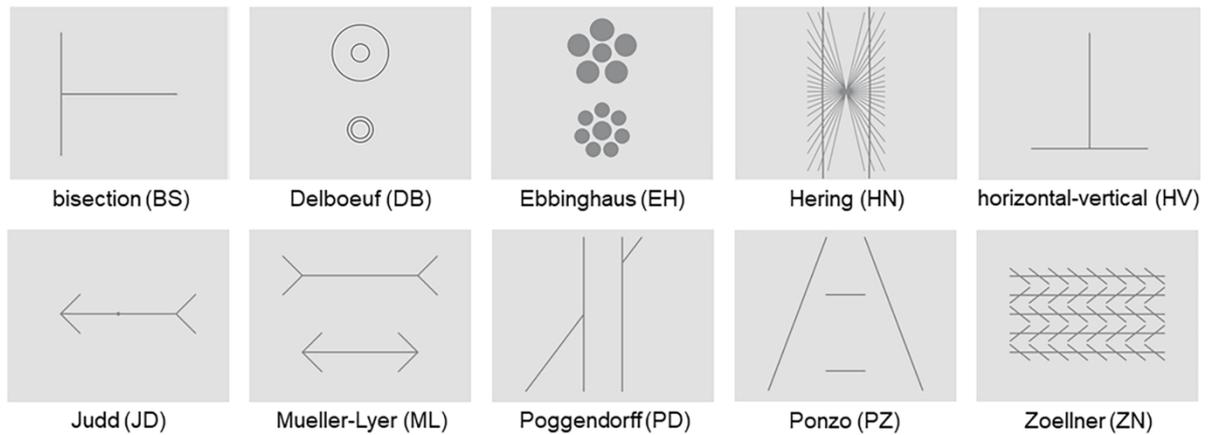


Figure 3. Study 1: Luminance-unspecific. Participants were asked to reach subjective equality in length (BS, HV, JD, ML, PZ), size (DB, EH), curvature (HN), position (PD), or orientation (ZN) with an adjustment procedure. For a detailed description of the different adjustments, please see Hamburger et al. (2007). Reprinted from Cretenoud et al. (2019).

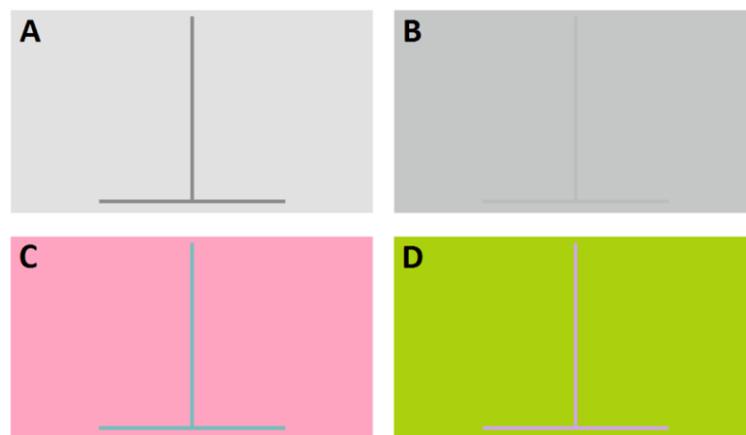


Figure 4. Study 1: Luminance-unspecific. Example of a horizontal-vertical stimulus with two luminance conditions: (A) 50% (Lum) and (B) 10% (LumLow) contrast and two isoluminant color conditions: (C) L – M (Iso) and (D) S – (L + M) (IsoS) according to DKL color space (Derrington et al., 1984). Please note that the colors may here vary in luminance due to reproduction. Reprinted from Cretenoud et al. (2019).

An EFA identified nine main factors, which explained $\sim 90\%$ of the variance. Strikingly, each factor mainly loaded on one illusion except for the bisection and horizontal-vertical illusions, which highly loaded on the same factor. In comparison, the first factor accounted for about 27% of the total variance in an unrotated factor solution.

The strong correlations between the four luminance conditions and the illusion-specific factors in the EFA indicate that, contrary to a widely held belief (e.g., Livingstone & Hubel, 1987), visual illusions do not break down under conditions of isoluminance (i.e., factors for illusions are not luminance-specific) and are therefore not primarily processed by the magnocellular system (also after controlling for subjective isoluminance, see Hamburger et al., 2007).

Table 1. Study 1: Luminance-unspecific. Diagonal: Test-retest reliability expressed as an intraclass correlation coefficient ($ICC_{3,1}$) for each condition. Triangle: Attenuated (upper) and disattenuated (lower) correlation coefficients between each pair of conditions (Pearson's r). Italics and bold font indicate significant results without and with Bonferroni correction, respectively. The color scale from blue to red reflects effect sizes from $r = -1$ to $r = 1$ (white corresponds to $r = 0$). Within-illusion correlations are strong while between-illusion correlations are in general weaker (except between the BS and HV illusions). Reprinted from Cretenoud et al. (2019).

	BS Lum	BS LumLow	BS Iso	BS IsoS	DB Lum	DB LumLow	DB Iso	DB IsoS	EH Lum	EH LumLow	EH Iso	EH IsoS	HN Lum	HN LumLow	HN Iso	HN IsoS	HV Lum	HV LumLow	HV Iso	HV IsoS	JD Lum	JD LumLow	JD Iso	JD IsoS	ML Lum	ML LumLow	ML Iso	ML IsoS	PD Lum	PD LumLow	PD Iso	PD IsoS	PZ Lum	PZ LumLow	PZ Iso	PZ IsoS	ZN Lum	ZN LumLow	ZN Iso	ZN IsoS
BS Lum	.67	.90	.92	.87	.00	.02	-.08	-.06	.10	.03	.09	-.02	.01	.01	-.24	.16	.80	.78	.79	.77	.39	.13	.27	.30	.05	.08	.11	.03	.44	.32	.32	.32	-.19	-.41	-.47	-.39	.11	.06	-.07	-.13
BS LumLow	1	.42	.91	.96	.25	.29	.19	.16	.06	.00	.04	.07	.20	.11	-.07	.32	.81	.82	.85	.84	.27	-.02	.33	.34	.22	.24	.21	.15	.39	.31	.30	.31	-.11	-.50	-.56	-.42	.26	.20	.12	.09
BS Iso	1	1	.70	.92	.06	.14	.06	.05	.14	.04	.16	.13	.08	.10	-.08	.28	.88	.87	.88	.84	.41	.20	.44	.41	.29	.30	.31	.28	.46	.32	.33	.35	-.14	-.39	-.46	-.32	.26	.31	.06	.03
BS IsoS	1	1	1	.55	.09	.13	.03	.00	.04	-.01	.06	.03	.13	.04	-.13	.25	.81	.80	.82	.82	.36	.07	.36	.37	.23	.22	.24	.17	.31	.27	.29	.27	-.10	-.43	-.50	-.34	.21	.19	.10	.03
DB Lum	.00	.48	.09	.15	.67	.96	.93	.89	-.21	-.33	-.08	.20	.15	.16	.16	.34	.07	.04	.19	.13	-.48	-.43	.04	-.10	.21	.28	.08	.10	.03	-.09	-.07	-.01	.11	-.14	-.06	-.08	.10	.20	.29	.46
DB LumLow	.02	.53	.19	.21	1	.71	.98	.93	-.10	-.19	.01	.28	.23	.20	.21	.40	.13	.11	.22	.15	-.43	-.40	.17	-.10	.35	.40	.22	.24	.05	-.08	-.09	-.03	.10	-.15	-.08	-.04	.27	.38	.43	.57
DB Iso	-.10	.34	.08	.04	1	1	.79	.97	-.05	-.15	.02	.32	.25	.24	.27	.41	.07	.05	.16	.10	-.49	-.42	.11	-.14	.36	.39	.22	.25	.05	-.08	-.10	.01	.06	-.16	-.06	-.07	.25	.40	.41	.58
DB IsoS	-.08	.27	.07	.00	1	1	1	.79	.04	-.11	.09	.33	.18	.20	.13	.35	.09	.02	.11	.06	-.43	-.37	.12	-.15	.33	.31	.16	.19	.02	-.10	-.15	-.06	-.02	-.14	-.03	-.07	.18	.39	.38	.56
EH Lum	.20	.14	.27	.10	-.43	-.19	-.10	.07	.37	.89	.74	.74	.24	.14	-.02	-.10	.26	.25	.10	.24	.48	.46	.52	.31	-.12	-.28	-.24	-.11	.18	.24	.02	-.02	-.26	-.01	-.08	.03	.57	.56	.45	.37
EH LumLow	.05	.00	.07	-.01	-.52	-.29	-.22	-.16	1	.61	.60	.64	.33	.21	.11	-.05	.14	.22	.04	.19	.36	.33	.35	.19	-.07	-.17	-.13	.02	.11	.18	.01	-.05	-.13	.05	.01	.08	.62	.53	.47	.23
EH Iso	.13	.07	.24	.10	-.12	.02	.03	.12	1	.94	.68	.78	.35	.30	.16	-.05	.30	.30	.19	.19	.35	.35	.53	.21	.03	-.02	-.06	-.01	.18	.25	.19	.06	-.02	.09	.15	.27	.46	.53	.45	.44
EH IsoS	-.05	.24	.33	.09	.51	.70	.75	.79	1	1	1	.23	.44	.37	.37	.13	.28	.33	.27	.32	.09	.17	.48	.12	.02	.00	-.13	.05	.03	.09	-.02	-.04	-.09	-.08	-.10	.04	.63	.69	.70	.54
HN Lum	.01	.46	.14	.26	.26	.39	.41	.29	.57	.60	.61	1	.48	.83	.75	.72	.13	.24	.24	.26	-.01	-.11	.25	.13	.16	.17	.16	.24	.21	.37	.40	.33	.10	-.19	-.02	-.01	.50	.40	.50	.45
HN LumLow	.02	.25	.18	.08	.28	.35	.40	.33	.33	.38	.54	1	1	.47	.77	.82	.24	.29	.31	.28	.10	.02	.24	.24	.14	.21	.25	.39	.14	.13	.27	.22	.10	-.12	-.18	-.02	.38	.33	.33	.29
HN Iso	-.50	-.18	-.15	-.30	.32	.42	.50	.25	-.05	.23	.32	1	1	1	.36	.65	-.04	.14	.19	.10	-.20	-.11	.11	.07	.08	.19	.16	.31	.02	.07	.20	.18	.07	-.06	.14	.10	.33	.28	.35	.27
HN IsoS	.24	.63	.42	.42	.52	.60	.58	.50	-.21	-.09	-.08	.34	1	1	1	.64	.28	.29	.36	.35	-.01	-.13	.22	.15	.32	.36	.44	.51	.16	.12	.23	.25	-.05	-.33	-.10	-.20	.29	.30	.29	.29
HV Lum	1	1	1	1	.10	.18	.09	.12	.51	.21	.43	.70	.23	.41	-.08	.42	.72	.92	.92	.90	.45	.19	.42	.40	.20	.25	.24	.22	.33	.16	.22	.24	-.14	-.47	-.47	-.43	.36	.37	.18	.10
HV LumLow	1	1	1	1	.06	.16	.07	.03	.49	.34	.44	.83	.41	.51	.29	.43	1	.69	.95	.91	.32	.09	.29	.32	.18	.26	.19	.21	.32	.17	.25	.26	.01	-.33	-.38	-.30	.41	.38	.21	.07
HV Iso	1	1	1	1	.32	.35	.24	.17	.22	.07	.31	.78	.46	.62	.43	.61	1	1	.55	.93	.25	.04	.32	.37	.11	.23	.18	.19	.32	.16	.27	.30	-.05	-.40	-.40	-.36	.33	.31	.18	.07
HV IsoS	1	1	1	1	.19	.20	.13	.08	.45	.28	.26	.79	.43	.48	.19	.51	1	1	1	.75	.30	.09	.29	.38	.14	.20	.16	.19	.41	.29	.33	.41	-.12	-.46	-.51	-.46	.39	.39	.22	.07
JD Lum	.68	.61	.71	.70	-.85	-.73	-.79	-.70	1	.67	.61	.28	-.01	.20	-.49	-.01	.76	.55	.48	.50	.48	.81	.70	.78	-.06	-.18	.04	.07	.26	.23	.17	.03	-.24	-.12	-.11	-.04	.25	.11	-.06	-.02
JD LumLow	.26	-.05	.38	.14	-.84	-.75	-.75	-.67	1	.67	.68	.58	-.26	.04	-.30	-.25	.35	.16	.09	.16	1	.39	.68	.75	-.01	-.17	.04	.19	.29	.30	.22	.14	-.36	-.01	.00	.02	.21	.18	-.05	-.03
JD Iso	.53	.80	.83	.76	.08	.31	.20	.22	1	.70	1	1	.58	.54	.28	.45	.78	.56	.68	.52	1	1	.40	.72	.18	.07	.19	.28	.27	.26	.17	.04	-.28	-.20	-.13	.06	.54	.52	.41	.46
JD IsoS	.56	.79	.75	.76	-.19	-.17	-.24	-.26	.79	.37	.39	.39	.30	.54	.17	.28	.71	.59	.75	.66	1	1	1	.44	.01	-.09	.06	.18	.46	.39	.40	.27	-.15	-.05	.02	.01	.24	.21	-.05	.08
ML Lum	.06	.40	.41	.36	.29	.49	.47	.43	-.22	-.11	.04	.06	.27	.23	.15	.47	.27	.25	.18	.18	-.11	-.02	.32	.02	.74	.92	.90	.86	.27	.18	.18	.27	.18	-.20	-.18	-.13	.22	.41	.05	.13
ML LumLow	.12	.47	.44	.37	.43	.59	.55	.43	-.57	-.27	-.04	.00	.31	.38	.39	.55	.36	.38	.39	.29	-.32	-.34	.13	-.17	1	.65	.90	.84	.24	.06	.15	.23	.30	-.23	-.17	-.16	.24	.34	.05	.07
ML Iso	.15	.37	.43	.36	.11	.30	.28	.21	-.45	-.19	-.08	-.31	.27	.42	.30	.63	.33	.26	.27	.21	.06	.07	.35	.11	1	1	.77	.93	.24	.08	.16	.26	.12	-.25	-.12	-.18	.15	.26	-.08	-.05
ML IsoS	.04	.29	.42	.28	.15	.35	.35	.27	-.23	.02	-.02	.13	.43	.70	.63	.80	.32	.32	.32	.27	.13	.38	.55	.34	1	1	1	.65	.18	.04	.10	.21	.10	-.20	-.07	-.16	.30	.39	.07	.02
PD Lum	.75	.84	.75	.57	.06	.07	.08	.03	.41	.20	.30	.09	.42	.29	.04	.27	.53	.53	.60	.65	.51	.65	.58	.96	.44	.41	.38	.32	.52	.89	.86	.86	-.04	-.12	-.14	-.08	.20	.32	-.15	.09
PD LumLow	.58	.69	.55	.53	-.16	-.14	-.12	-.17	.58	.33	.44	.27	.77	.28	.16	.22	.27	.30	.31	.48	.48	.70	.61	.86	.30	.11	.13	.07	1	.47	.93	.87	-.10	-.09	-.13	.01	.21	.32	.00	.20
PD Iso	.51	.61	.52	.51	-.11	-.13	-.14	-.22	.04	-.01	.30	-.06	.75	.52	.43	.38	.33	.39	.47	.50	.32	.45	.35	.78	.27	.24	.24	1	1	.59	.91	.05	.00	.03	.10	.11	.25	-.08	.11	
PD IsoS	.57	.70	.61	.53	-.02	-.06	.01	-.10	-.04	-.09	-.10	-.12	.69	.46	.43	.46	.42	.45	.59	.68	.05	.33	.10	.59	.45	.42	.43	1	1	1	.48	-.04	-.16	-.15	-.16	.12	.27	-.12	.06	
PZ Lum	-.34	-.26	-.26	-.21	.21	.18	.09	-.03	-.63	-.25	-.04	-.27	.22	.21	.16	-.10	-.25	.01	-.09	-.21	-.52	-.84	-.65	-.34	.32	.56	.21	.17	-.08	-.22	.09	-.08	.46	.63	.49	.53	.12	.06	.00	.08
PZ LumLow	-.75	-.1	-.70	-.85	-.25	-.27	-.27	-.23	-.03	.10	.17	-.24	-.42	-.26	-.15	-.62	-.82	-.59	-.80	-.80	-.26	-.02	-.47	-.12	-.34	-.43	-.43	-.37	-.25	-.19	.00	-.35	1	.45	.87	.90	-.09	.04	-.06	.06
PZ Iso																																								

4.1.2. Study 1: Orientation-unspecific²

The strong correlations between the horizontal-vertical and bisection (which in fact is a horizontal-vertical illusion rotated by 90°) illusion magnitudes suggest that individual differences in the perception of visual illusions are stable across changes in orientation, which I further tested here. Five illusions were tested in twenty participants: horizontal-vertical (HV), Müller-Lyer (ML), Poggendorff (PD), Ponzo (PZ), and Zöllner (ZN). Each illusion was presented twice under four different orientations: -60°, -15°, 30°, and 75° (Figure 5).

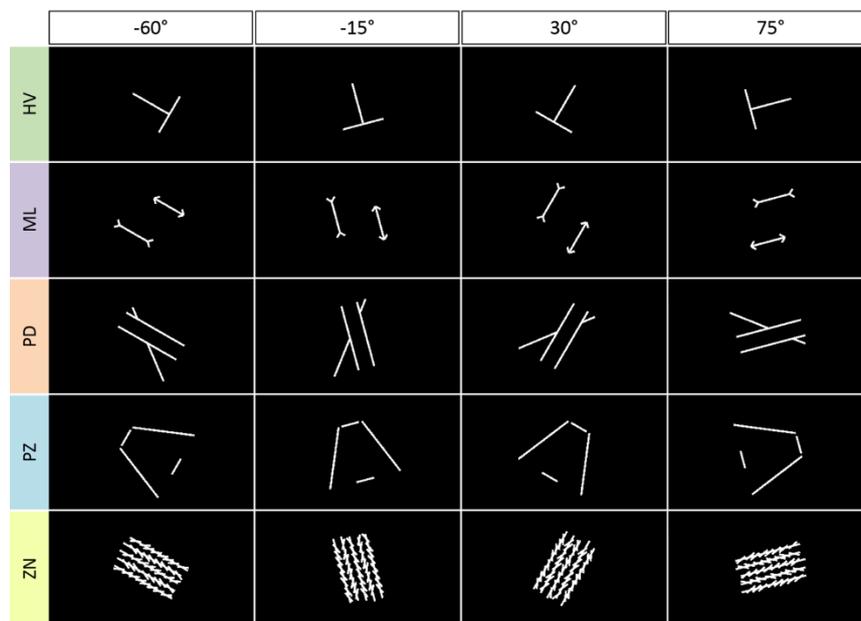


Figure 5. Study 1: Orientation-unspecific. Five illusions were tested with four orientations. By row: horizontal-vertical (HV), Müller-Lyer (ML), Poggendorff (PD), Ponzo (PZ), and Zöllner (ZN) illusion. By column: -60°, -15°, 30°, and 75°. Reprinted from Cretenoud et al. (2019).

Table 2 shows attenuated (upper triangle) and disattenuated (lower triangle) correlation coefficients. Interestingly, 97% (29 out of 30) and 100% of the attenuated and disattenuated within-illusion correlations showed $r > 0.3$, respectively. In contrast, only 16% (26 out of 160) and 32% (51 out of 160) of the attenuated and disattenuated between-illusion correlations showed $r > 0.3$, respectively.

² This section is based on Cretenoud, A. F., Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, 19(14):12, 1–21. (Appendix A; Experiment 3)

Table 2. Study 1: Orientation-unspecific. Diagonal: Test-retest reliability expressed as an intraclass correlation coefficient ($ICC_{3,1}$) for each condition. Triangle: Attenuated (upper) and disattenuated (lower) correlation coefficients between each pair of conditions (Pearson's r). Italics and bold font indicate significant results without and with Bonferroni correction, respectively. The color scale from blue to red reflects effect sizes from $r = -1$ to $r = 1$ (white corresponds to $r = 0$). Reprinted from Cretenoud et al. (2019).

	HV-60	HV-15	HV30	HV75	ML-60	ML-15	ML30	ML75	PD-60	PD-15	PD30	PD75	PZ-60	PZ-15	PZ30	PZ75	ZN-60	ZN-15	ZN30	ZN75
HV-60	.74	.43	.50	.71	-.13	.12	.05	-.35	.08	-.08	-.18	-.17	.07	.04	.15	.06	-.08	.07	-.12	-.04
HV-15	.55	.86	.61	.73	-.06	.05	-.25	-.31	-.35	-.31	-.09	-.28	-.19	-.13	-.15	.01	-.06	.15	.04	.22
HV30	.72	.80	.66	.71	.10	.00	-.13	-.33	-.45	-.31	-.40	-.26	.02	.04	.21	.51	.03	.16	.05	-.05
HV75	.85	.81	.90	.94	.13	.23	.06	-.29	-.22	-.30	-.28	-.38	.01	.16	.11	0.11	-.03	.10	-.06	.05
ML-60	-.17	-.07	.14	.14	.83	.72	.60	.52	-.18	-.17	-.30	-.29	.12	.27	.11	.01	.01	-.01	.07	-.08
ML-15	.15	.06	.00	.26	.85	.85	.79	.39	-.09	-.11	-.17	-.24	.25	.38	.24	-.02	.23	.18	.18	-.02
ML30	.06	-.28	-.16	.06	.69	.89	.91	.55	-.01	-.06	-.10	-.25	.11	.15	.17	-.11	.33	.16	.08	-.08
ML75	-.46	-.37	-.45	-.34	.63	.47	.64	.81	.02	.15	.13	.08	-.11	-.06	.03	-.37	.33	.23	.29	.41
PD-60	.10	-.40	-.59	-.25	-.21	-.11	-.01	.03	.87	.45	.55	.58	.21	.11	-.01	-.26	.23	.08	.18	-.04
PD-15	-.12	-.45	-.51	-.42	-.25	-.16	-.08	.23	.65	.55	.71	.70	-.13	-.33	-.03	.00	.47	.16	.19	.01
PD30	-.22	-.10	-.52	-.31	-.36	-.20	-.12	.16	.64	1	.85	.78	-.26	-.35	-.43	-.30	.40	.31	.32	.25
PD75	-.21	-.33	-.34	-.41	-.34	-.28	-.28	.09	.66	1	.90	.88	.04	-.21	-.22	.05	.41	.20	.34	.06
PZ-60	.09	-.25	.04	.01	.16	.32	.14	-.14	.26	-.21	-.33	.05	.73	.54	.52	.47	.04	-.03	.04	-.22
PZ-15	.06	-.16	.06	.20	.35	.48	.18	-.08	.14	-.53	-.45	-.27	.74	.72	.50	.19	-.16	-.24	-.15	.04
PZ30	.27	-.24	.38	.18	.18	.40	.27	.05	-.02	-.06	-.70	-.36	.93	.89	.44	.57	.03	-.04	.07	-.25
PZ75	.10	.02	.95	.18	.02	-.03	-.17	-.63	-.42	.00	-.49	.09	.83	.35	1	.43	-.06	-.18	-.07	-.56
ZN-60	-.14	-.10	.06	-.04	.02	.40	.56	.58	.39	1	.69	.69	.07	-.30	.06	-.13	.40	.74	.69	.31
ZN-15	.12	.23	.27	.14	-.02	.27	.24	.37	.13	.31	.47	.31	-.04	-.41	-.09	-.40	1	.49	.92	.44
ZN30	-.15	.05	.06	-.07	.09	.22	.10	.36	.21	.29	.39	.41	.05	-.20	.12	-.12	1	1	.79	.31
ZN75	-.05	.28	-.07	.07	-.11	-.02	-.11	.55	-.05	.01	.33	.08	-.31	.06	-.46	-1	.59	.76	.43	.68

Four factors were identified by inspecting the EFA scree plot. The four-factor rotated solution accounted for ~61% of the variance (RF1: 17%, RF2: 17%, RF3: 15%, RF4: 12%). The first factor was mainly composed of the Poggendorff and Zöllner conditions, while the second factor mainly loaded on the horizontal-vertical and Poggendorff conditions. The third and fourth factors were respectively dominated by loadings from the Müller-Lyer and Ponzo conditions. Each illusion mainly loaded on one factor, suggesting that factors for illusions are not orientation-specific. Instead, these results suggest that more or less each illusion makes up its own factor, with the exception of the Poggendorff illusion, which highly loaded on two factors. Note that the first factor of an unrotated factor solution only accounted for about 22% of the total variance.

4.1.3. Study 1: Feature-unspecific³

Factors for visual illusions are neither luminance- nor orientation-specific. In other terms, there are strong correlations between different variants of an illusion, which are presented with different conditions of luminance or orientation. Here, I examined to what extent the susceptibility to a single visual illusion differs as a function of other features, such as size, color, shape, and texture.

Eighty participants were tested with 19 variants of the Ebbinghaus illusion, which varied in size, color, shape, texture, or dynamics (static versus moving), and a control condition (20 conditions in total; Figure 6). There were two trials per condition.

To control for the control condition variability, which showed up as a significant effect in the control condition (illusion magnitude: $2.595\% \pm 0.399\%$, $t[79] = 6.508$, $p < 0.001$, $d = 0.728$), partial correlations were computed for each pair of variants. Attenuated (upper triangle) and disattenuated (lower triangle) partial correlation coefficients are reported in Table 3. Strong between-variant (i.e., within-illusion) effects were observed. Indeed, out of 171 attenuated correlations, only 26 showed $r < 0.3$. Similarly, only seven disattenuated correlations showed $r < 0.3$.

An EFA suggested a one-factor model explaining ~44% of the variance. All conditions except the control condition highly loaded on the unique factor, suggesting that there is one mechanism underlying the Ebbinghaus illusion that is not feature-specific. A similar result was previously found for the Müller-Lyer illusion (Coren et al., 1976). A factor analysis showed that 45 measures of different illusions were best represented by a 2-factor model, with one factor mainly loading on several variants of the Müller-Lyer illusion.

³ This section is based on Cretenoud, A. F., Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, 19(14):12, 1–21. (Appendix A; Experiment 1)

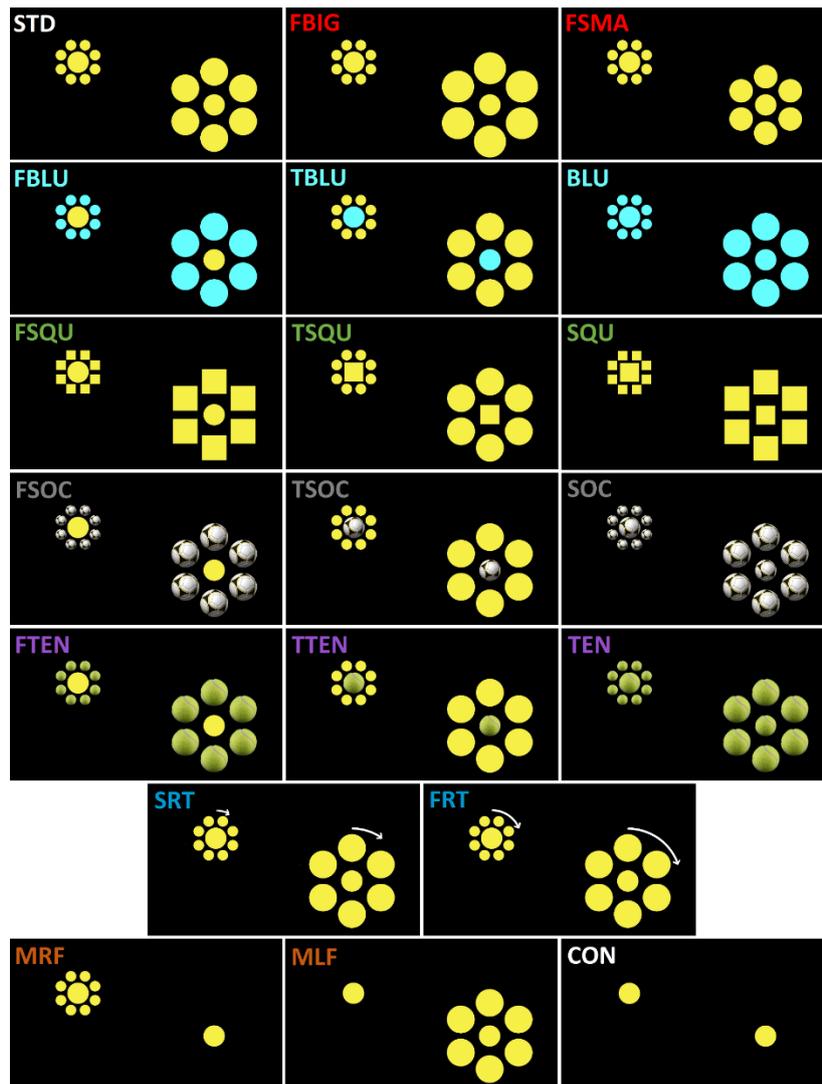


Figure 6. Study 1: Feature-unspecific. Participants were tested with 19 variants of the Ebbinghaus illusion and a control condition. The conditions of the illusion varied from the standard one (STD) as follows: bigger right flankers (FBIG), smaller right flankers (FSMA), blue flankers (FBLU), blue targets (TBLU), both blue flankers and targets (BLU), square flankers (FSQU), square targets (TSQU), both square flankers and targets (SQU), soccer ball flankers (FSOC), soccer ball targets (TSOC), both soccer ball flankers and targets (SOC), tennis ball flankers (FTEN), tennis ball targets (TTEN), both tennis ball flankers and targets (TEN), slow clockwise rotation of the flankers (SRT), fast clockwise rotation of the flankers (FRT), missing right flankers (MRF), missing left flankers (MLF) and missing all flankers (i.e., control condition; CON). The acronyms of similar conditions are presented in the same color. Reprinted from Cretenoud et al. (2019).

Table 3. Study 1: Feature-unspecific. Diagonal: Test-retest reliability expressed as an intraclass correlation coefficient (ICC_{3,1}) for each variant. All ICCs were significant. Triangle: Attenuated (upper) and disattenuated (lower) partial correlation coefficients between each pair of variants (Pearson's *r*), controlling for the control condition variability. Italics and bold font indicate significant results without and with Bonferroni correction, respectively. All correlation coefficients were positive. The color scale from white to red reflects effect sizes from *r* = 0 to *r* = 1. The acronyms of similar variants are presented in the same color, as in Figure 6. Reprinted from Cretienoud et al. (2019).

	STD	FBIG	FSMA	FBLU	TBLU	BLU	FSQU	TSQU	SQU	FSOC	TSOC	SOC	FTEN	TTEN	TEN	SRT	FRT	MRF	MLF
STD	.59	.59	.52	.39	.56	.40	.41	.38	.39	.29	.34	.50	.43	.48	.29	.48	.44	.30	.23
FBIG	.93	.72	.50	.60	.63	.49	.50	.33	.59	.45	.37	.57	.55	.35	.43	.63	.56	.33	.36
FSMA	.96	.84	.50	.46	.54	.30	.34	.38	.57	.36	.42	.62	.52	.30	.49	.52	.49	.27	.38
FBLU	.65	.86	.81	.67	.63	.60	.36	.28	.45	.52	.39	.45	.58	.35	.30	.45	.48	.30	.34
TBLU	.98	.97	1	1	.60	.52	.53	.44	.53	.50	.52	.61	.59	.47	.43	.55	.55	.29	.35
BLU	.67	.73	.53	.95	.86	.63	.42	.30	.40	.49	.38	.41	.46	.43	.36	.47	.47	.24	.30
FSQU	1	1	.84	.71	1	.91	.44	.46	.39	.43	.42	.43	.51	.61	.41	.49	.47	.21	.14
TSQU	.72	.48	.74	.39	.76	.48	1	.59	.46	.36	.32	.46	.39	.45	.31	.33	.27	.25	.29
SQU	.74	1	1	.79	1	.72	.95	.89	.50	.45	.46	.65	.55	.41	.58	.51	.39	.40	.41
FSOC	.46	.63	.61	.77	.79	.75	.90	.56	.77	.68	.47	.40	.54	.29	.40	.58	.49	.18	.30
TSOC	.57	.51	.74	.56	.84	.57	.87	.48	.81	.71	.65	.58	.58	.47	.45	.48	.46	.37	.22
SOC	.89	.89	1	.70	1	.66	.95	.79	1	.63	.95	.60	.57	.50	.54	.54	.50	.33	.25
FTEN	.79	.86	1	.94	1	.76	1	.63	1	.87	.97	.99	.60	.51	.53	.55	.53	.27	.31
TTEN	.86	.52	.55	.54	.81	.71	1	.78	.76	.43	.77	.87	.90	.58	.49	.40	.49	.31	.16
TEN	.61	.69	1	.44	.78	.63	.85	.44	1	.68	.76	1	.97	.93	.60	.53	.42	.23	.40
SRT	.82	.96	.97	.69	.93	.76	1	.53	.94	.91	.76	.92	.96	.66	1	.61	.76	.10	.37
QRT	.76	.83	.88	.72	.90	.74	1	.37	.68	.75	.69	.81	.87	.82	.70	1	.65	.15	.24
MRF	.56	.49	.51	.44	.46	.37	.32	.35	.76	.24	.59	.54	.40	.51	.26	.10	.15	.57	.16
MLF	.38	.52	.69	.51	.55	.47	.23	.46	.73	.44	.31	.39	.49	.22	.71	.59	.35	.22	.63

4.1.4. Study 2: Context-unspecific⁴

Here, I investigated whether individual differences in the perception of visual illusions are stable across contexts, i.e., whether factors for illusions are context-specific. Seventy-six participants were tested with two illusions. On one hand, the inward and outward Müller-Lyer (ML) illusions were tested. On the other hand, the Ponzo (PZ) illusion was presented either with the upper line only (up) or with the lower line only (down). Each of the four conditions was presented with 3 contexts (poor-, moderate-, and rich-context), making up 12 variants in total (Figure 7). Each variant was tested twice. In the rich-context variants, real-world pictures were used. Poor-context and moderate-context variants were drawn based on the rich-context backgrounds, so that perspective lines matched the perspective of the real-world pictures.

⁴ This section is based on Cretienoud, A. F., Grzeczowski, L., Bertamini, M., & Herzog, M. H. (2020). Individual differences in the Müller-Lyer and Ponzo illusions are stable across different contexts. *Journal of Vision*, 20(6):4, 1–14. (Appendix B)

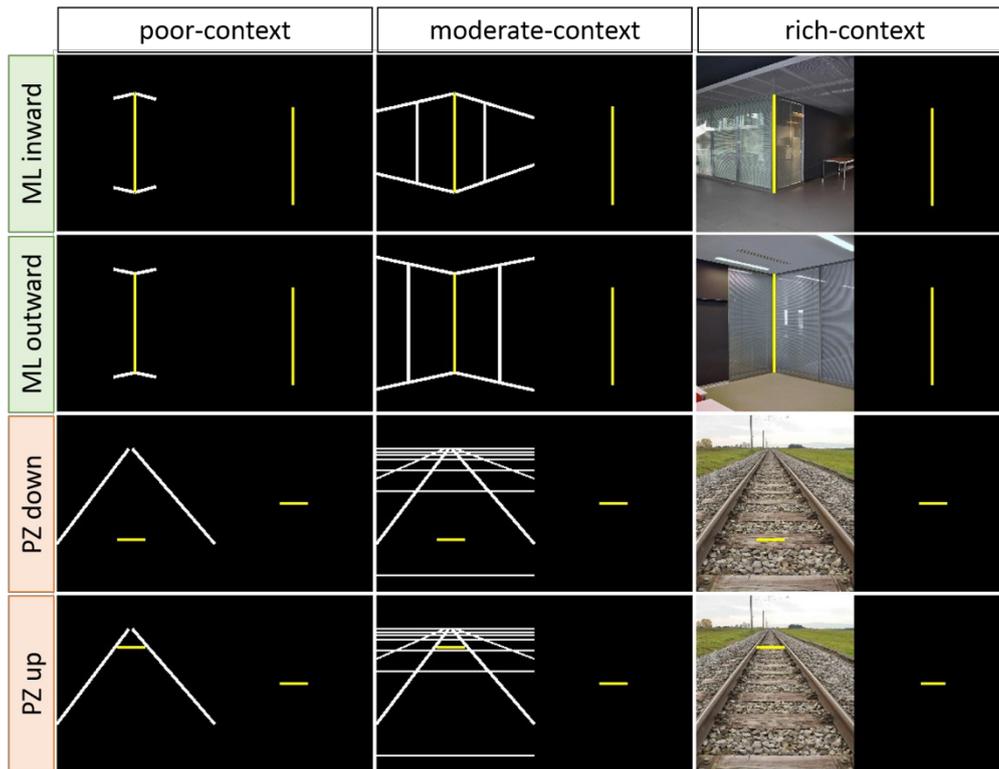


Figure 7. Study 2. Participants had to adjust the length of the yellow adjustable line (right half screen) to match the length of the yellow reference line (left half screen). The inward and outward Müller-Lyer illusions were presented in distinct trials. Similarly, the Ponzo illusion was presented either with the upper line only (up) or with the lower line only (down). Each condition was tested with 3 different contexts, making up 12 variants in total. By row: Müller-Lyer (green) inward and outward conditions, Ponzo (orange) down and up conditions. By column: poor-, moderate-, and rich-context variants. Reprinted from Cretienoud, Grzeczkowski, et al. (2020).

Attenuated (upper triangle) and disattenuated (lower triangle) correlations were computed between each pair of variants (Table 4). Strong correlations were observed between different variants of the same illusion (within-illusion) but only weak correlations were observed across illusions (between-illusion, i.e., between a Müller-Lyer and a Ponzo variant), which suggests that factors are illusion-specific but not context-specific. Note, however, that 35 out of 36 attenuated between-illusion correlations were positive. They significantly differed from zero ($t[35] = 11.146, p < 0.001, d = 1.858$), suggesting that a small proportion of the variance underlying the Müller-Lyer and Ponzo illusions is accounted for by a common – albeit weak – factor.

Table 4. Study 2. Diagonal: Test-retest reliability expressed as a Pearson correlation coefficient for each variant. All correlations were significant. Triangle: Attenuated (upper) and disattenuated (lower) correlation coefficients between each pair of variants (Pearson’s r). Italics and bold font indicate significant results without and with Bonferroni correction, respectively. The color scale from blue to red reflects effect sizes from $r = -1$ to $r = 1$ (white corresponds to $r = 0$). Within-illusion correlations are strong while between-illusion correlations are in general weaker. ML: Müller-Lyer illusion; PZ: Ponzo illusion; 1: poor-context; 2: moderate-context; 3: rich-context. Adapted from Cretenoud, Grzeczowski, et al. (2020).

	ML1 inward	ML1 outward	ML2 inward	ML2 outward	ML3 inward	ML3 outward	PZ1 down	PZ1 up	PZ2 down	PZ2 up	PZ3 down	PZ3 up
ML1 inward	.51	.48	.45	.40	.57	.41	.18	.32	.10	.12	.21	.31
ML1 outward	.97	.48	.24	.57	.46	.40	-.04	.18	.06	.18	.05	.13
ML2 inward	.93	.51	.45	.33	.56	.44	.21	.24	.07	.18	.19	.12
ML2 outward	.66	.98	.59	.72	.42	.53	.13	.24	.11	.26	.07	.40
ML3 inward	.95	.79	.99	.59	.71	.67	.11	.27	.16	.20	.16	.29
ML3 outward	.69	.69	.78	.74	.96	.70	.06	.12	.18	.15	.12	.33
PZ1 down	.37	-.09	.47	.23	.19	.11	.44	.49	.38	.48	.39	.34
PZ1 up	.59	.35	.47	.38	.42	.19	.98	.57	.33	.60	.47	.31
PZ2 down	.20	.11	.14	.19	.27	.31	.82	.63	.49	.38	.43	.30
PZ2 up	.21	.35	.36	.40	.32	.24	.96	1.00	.71	.58	.42	.40
PZ3 down	.37	.10	.35	.10	.24	.18	.74	.78	.76	.69	.64	.28
PZ3 up	.54	.24	.23	.60	.44	.50	.65	.52	.55	.66	.45	.62

A PCA revealed two components that accounted for ~54% of the total variance in the data (~35% for the first component). The first and second factors mainly loaded on the Müller-Lyer and Ponzo variants, respectively, with an intercorrelation of $r = 0.324$. A mixed effects model did not reveal any significant interaction between the context and participants’ age, arguing against the two-factor model of Piaget (1961, 1963; see also Wagner, 1977).

4.1.5. Study 3: Factors for illusions combine independently⁵

The two previous experiments suggest that the visual space of illusions includes several illusion-specific factors. Here, I specifically examined how factors for the vertical-horizontal, Müller-Lyer, and Ponzo illusions relate to each other by measuring the susceptibility to each illusion separately and to combinations of two illusions, which I call merged illusions.

⁵ This section is based on Cretenoud, A. F., Francis, G., & Herzog, M. H. (2020). When illusions merge. *Journal of Vision*, 20(8):12, 1–15. (Appendix C)

Ninety-eight participants were tested with seven illusions (Figure 8): the vertical-horizontal (VH; also called horizontal-vertical), Müller-Lyer (ML), and Ponzo (PZ) illusions, as well as the vertical-horizontal illusion congruently (VH-ML con.) and incongruently (VH-ML inc.) merged with the Müller-Lyer illusion, and the Ponzo illusion congruently (PZ-ML con.) and incongruently (PZ-ML inc.) merged with the Müller-Lyer illusion. Two reference-dependent conditions were tested for each illusion, making up 14 conditions. For example, either the horizontal (VH hor.) or vertical (VH ver.) segment of the VH illusion was adjusted. Each condition was tested twice.

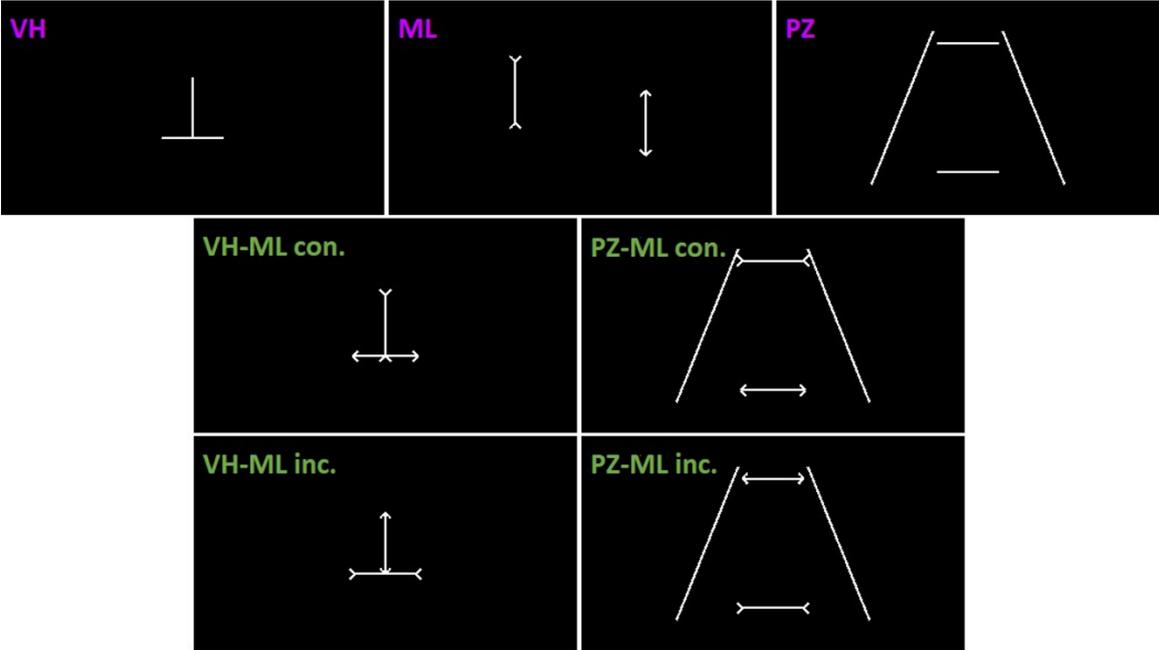


Figure 8. Study 3. Seven illusions were tested: the vertical-horizontal (VH), Müller-Lyer (ML), and Ponzo (PZ) illusions, as well as the vertical-horizontal illusion congruently (VH-ML con.) and incongruently (VH-ML inc.) merged with the Müller-Lyer illusion, and the Ponzo illusion congruently (PZ-ML con.) and incongruently (PZ-ML inc.) merged with the Müller-Lyer illusion. Purple labels indicate non-merged illusions; green labels indicate merged illusions. Reprinted from Cretenoud, Francis, et al. (2020).

Significant pairwise correlations (Table 5A) were mostly observed between a merged condition and the non-merged conditions of which it was made. For example, the horizontal congruent and incongruent VH-ML conditions (VH-ML con. hor. and VH-ML inc. hor.) significantly correlated with the horizontal VH condition (VH hor.). Interestingly, weaker correlations were observed between the merged conditions and the Müller-Lyer illusion (ML in. and ML out. conditions), suggesting that one illusion is stronger than the other when two illusions are merged.

Note that the 28 between-illusion correlations (VH versus ML, VH versus PZ, ML versus PZ, VH versus PZ-ML, and PZ versus VH-ML conditions) showed an averaged effect that was significantly different from zero ($t[27] = 4.029, p < 0.001, d = 0.761$), suggesting that a small proportion of the variance underlying the VH, ML, and PZ illusions is accounted for by a weak common factor.

Table 5. Study 3. (A) Diagonal: Test-retest reliability expressed as an intraclass correlation coefficient ($ICC_{3,1}$) for each condition. All ICCs were significant. Triangle: Correlations between each pair of conditions expressed as correlation coefficients (Pearson's r). A color scale from blue to red reflects effect sizes from $r = -1$ to $r = 1$ (white corresponds to $r = 0$). Italics and bold font indicate significant results without and with Bonferroni correction, respectively. (B) Standardized path coefficients ($*p < 0.05$; $**p < 0.01$; $***p < 0.001$) from a path model and communalities (i.e., the variance explained r^2) of each merged condition. Grey shading indicates the non-merged conditions that made up every merged condition (i.e., the expected predictors). Purple labels indicate non-merged conditions; green labels indicate merged conditions. Adapted from Cretenoud, Francis, et al. (2020).

(A)	VH hor.	VH ver.	ML in.	ML out.	PZ down	PZ up	VH-ML con. hor.	VH-ML con. ver.	VH-ML inc. hor.	VH-ML inc. ver.	PZ-ML con. down	PZ-ML con. up	PZ-ML inc. down	PZ-ML inc. up
VH hor.	.50	-.36	.11	.16	.18	.29	.62	-.28	.65	-.19	.27	.12	.25	.22
VH ver.		.32	-.04	.03	-.05	.00	-.33	.66	-.19	.54	-.12	.10	.06	-.12
ML in.			.39	.02	.15	.07	.23	.02	.20	.16	.24	.12	.02	.14
ML out.				.31	.15	.05	.19	.30	.21	.10	.21	.25	.41	.01
PZ down					.65	-.36	.14	-.00	.20	-.10	.59	-.34	.47	-.19
PZ up						.43	.06	.01	.21	.07	-.27	.58	-.20	.45
VH-ML con. hor.							.31	-.28	.55	-.18	.40	.00	.13	.26
VH-ML con. ver.								.53	-.18	.58	-.13	.28	.22	-.26
VH-ML inc. hor.									.34	-.08	.30	.06	.25	.24
VH-ML inc. ver.										.30	-.02	.07	-.06	.01
PZ-ML con. down											.57	-.39	.36	.14
PZ-ML con. up												.49	.10	.10
PZ-ML inc. down													.62	-.32
PZ-ML inc. up														.50

(B)	VH hor.	VH ver.	ML in.	ML out.	PZ down	PZ up	r^2
VH-ML con. hor.	.601***	-.113	.175*	.110	-.069	-.151	.452
VH-ML con. ver.	-.122	.606***	.053	.301***	.004	.024	.524
VH-ML inc. hor.	.606***	.037	.122	.093	.073	.050	.455
VH-ML inc. ver.	-.022	.530***	.194*	.100	-.105	.016	.341
PZ-ML con. down	.194*	-.029	.165*	.120	.451***	-.177	.437
PZ-ML con. up	-.003	.083	.119	.260**	-.221*	.484***	.446
PZ-ML inc. down	.235*	.143	-.037	.332***	.328***	-.166	.389
PZ-ML inc. up	.080	-.092	.104	-.008	-.079	.391***	.233

A path model was computed to determine whether the illusion magnitudes of the merged conditions (VH-ML con. hor., VH-ML con. ver., VH-ML inc. hor., VH-ML inc. ver., PZ-ML con. down, PZ-ML con. up, PZ-ML inc. down, PZ-ML inc. up) can be predicted from the illusion magnitudes of the non-merged conditions (VH hor., VH ver., ML in., ML out., PZ down, PZ up). In a path model, each predictor is regressed onto each outcome (Beaujean, 2014). Importantly, the path model considers the covariances between predictors (i.e., non-merged conditions).

The illusion magnitude of a merged condition was strongly predicted from the illusion magnitudes of the two non-merged conditions of which it was made (Table 5B). For example, adjusting the vertical segment of the incongruent VH-ML illusion (VH-ML inc. ver.) refers to both the vertical VH (VH ver.) and the inward ML (ML in.) conditions. Indeed, both the vertical VH (VH ver.) and the inward ML (ML in.) conditions showed significant standardized path coefficients (VH ver.: 0.530, $p < 0.001$; ML in.: 0.194, $p = 0.021$) to the vertical incongruent VH-ML condition (VH-ML inc. ver.).

An EFA suggested a four-factor model explaining 56% of the total variance (RF1: 16%; RF2: 14%; RF3: 14%; RF4: 12%). The first and third factors highly loaded on the different conditions including the horizontal and vertical VH illusion, respectively, while the second and fourth factors highly loaded on the conditions including the PZ and ML illusions, respectively (Table 6).

Table 6. Study 3. Factor loadings from an exploratory factor analysis (EFA) after promax rotation. A color scale from blue (negative loadings) to red (positive loadings) is shown. Purple labels indicate non-merged conditions; green labels indicate merged conditions. Bold numbers indicate loadings that are expected to be high under the hypothesis that the factors relate to conditions including horizontal VH, PZ, vertical VH, and ML, respectively. Reprinted from Cretienoud, Francis, et al. (2020).

	RF1	RF2	RF3	RF4
VH hor.	0.628	0.158	-0.308	0.226
VH ver.	-0.150	0.009	0.699	0.048
ML in.	0.353	-0.054	0.157	-0.032
ML out.	0.182	0.131	0.141	0.460
PZ down	0.181	-0.542	0.000	0.340
PZ up	0.337	0.692	0.046	-0.119
VH-ML con. hor.	0.651	-0.041	-0.218	0.076
VH-ML con. ver.	-0.171	0.153	0.770	0.366
VH-ML inc. hor.	0.649	0.066	-0.120	0.183
VH-ML inc. ver.	0.131	-0.024	0.774	-0.106
PZ-ML con. down	0.547	-0.621	0.090	0.106
PZ-ML con. up	0.021	0.831	0.064	0.330
PZ-ML inc. down	0.008	-0.073	-0.037	0.856
PZ-ML inc. up	0.550	0.152	0.038	-0.422

4.2. Are individual differences in the perception of visual illusions stable?

Vision scientists have tried to classify illusions for more than a century with mixed results, e.g., there are two classes of illusions, namely illusions of linear extent and distortions, versus there is a higher multifactorial space with – more or less – each illusion making up its own factor. In a fourth study (Study 4), I examined how likely it is that these discrepancies arise from instability in the perception of visual illusions across eyes, time, and measurement methods. In short, the results suggest that none of these aspects significantly affect the existing individual differences in susceptibility to an illusion.

In this study, both traditional null hypothesis significance testing (NHST) and Bayesian statistics were reported. Unlike NHST (where absence of proof is not proof of absence), inferences about null results are allowed with Bayesian statistics. A Bayes factor (BF₁₀, later referred to as BF) smaller than 0.33 indicates evidence in favor of the null hypothesis, while a BF larger than 3 indicates substantial support for the alternative hypothesis (Jeffreys, 1961). BFs lying between 0.33 and 3 are considered as inconclusive.

4.2.1. Study 4: Stable individual differences across eyes⁶

Monocular viewing condition was shown to significantly decrease the magnitude of an actual, real-world Ponzo illusion compared to binocular viewing condition, because of a reduced perception of depth cues following the elimination of stereopsis (Leibowitz et al., 1969). Here, I examined whether several two-dimensional illusory percepts are reliable interocularly, i.e., from one eye to the other, and from a monocular to a binocular viewing condition. I also investigated the role of visual acuity in the perception of visual illusions.

Monocular (left and right) and binocular visual acuities were measured using the Freiburg visual acuity test (Bach, 1996) in 15 participants, which were then tested with seven illusions (Figure 9): two variants of the Ebbinghaus (EB and EB2), a Müller-Lyer (ML), and four variants of the Ponzo (PZ; a wider variant, PZw; a grid variant, PZg; and a corridor variant, PZc) illusion. Each illusion was tested twice monocularly (both left and right) and binocularly (i.e., there were three eye conditions).

Mixed effects models were computed. The fixed effects were eye condition (monocular left, monocular right, binocular), visual acuity, and sex. Random effects of participants and illusions were accounted for as random intercepts. The significance of each predictor in the model was assessed by computing likelihood ratio tests, which express the relative likelihood of the data given two competing models. The effect size was computed as a measure of explained variance with the random effect structure included.

⁶ This section is based on Cretenoud, A. F., Grzeczowski, L., Kunchulia, M., & Herzog, M. H. (submitted). Individual differences in the perception of visual illusions are stable across eyes, time, and measurement methods. (Appendix D; Experiment 1)

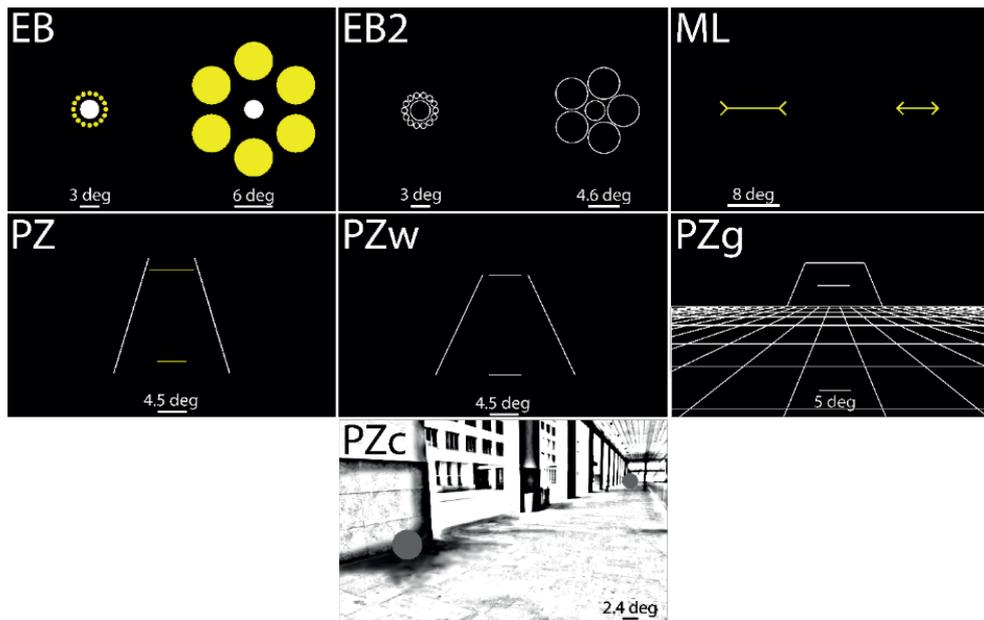


Figure 9. Study 4: Stability across eyes. Seven illusions were tested monocularly (left and right) and binocularly: two variants of the Ebbinghaus (EB and EB2), a Müller-Lyer (ML), and four variants of the Ponzo (PZ, PZw, PZg, and PZc) illusion. Reprinted from Cretenoud, Grzeczowski, et al. (submitted).

First, and most importantly, a likelihood ratio test showed that the eye condition did not significantly improve the model ($\chi^2(2) = 0.153, p = 0.926$). The eye condition was therefore removed from the model. Second, illusion magnitudes did not significantly vary with visual acuity ($\chi^2(1) = 0.140, p = 0.710$), which was therefore also removed from the model. Last, sex did not significantly improve the model ($\chi^2(1) = 0.310, p = 0.579$; see Shaqiri et al., 2018). Hence, the best model did not include any predictor but only random effects for participants and for illusions. This model accounted for 83% of the variance in the data, which suggests that a large part of the variability is accounted for by individual and between-illusion differences. To sum up, our results suggest that illusion magnitudes and individual differences in illusion magnitudes do not significantly vary with manipulation of the eye(s) to which the stimulus is presented and are not a function of visual acuity.

4.2.2. Study 4: Stable individual differences across time⁷

Bistable percepts were shown to change over time (Wexler et al., 2015). Likewise, illusion decrements, i.e., a decrease in the susceptibility to an illusion with repeated visual exposure, have been shown for a long time (e.g., Coren & Girgus, 1972; Judd, 1902; Predebon, 2006). Here, the susceptibility to several illusions was measured at different time points within a month to evaluate how stable individual differences in the perception of visual illusions are over time.

⁷ This section is based on Cretenoud, A. F., Grzeczowski, L., Kunchulia, M., & Herzog, M. H. (submitted). Individual differences in the perception of visual illusions are stable across eyes, time, and measurement methods. (Appendix D; Experiment 2)

Fourteen participants were tested with eight illusions (two trials each; Figure 10): the Ebbinghaus (EB), horizontal-vertical (HV), Müller-Lyer (ML), Poggendorff (PD), Ponzo (PZ), two variants of the contrast (CS and CS2), and White (WH) illusions. Participants were tested in 12 sessions at days 1, 2, 3, 4, 5, 8, 9, 10, 11, 12, 19, and 33.

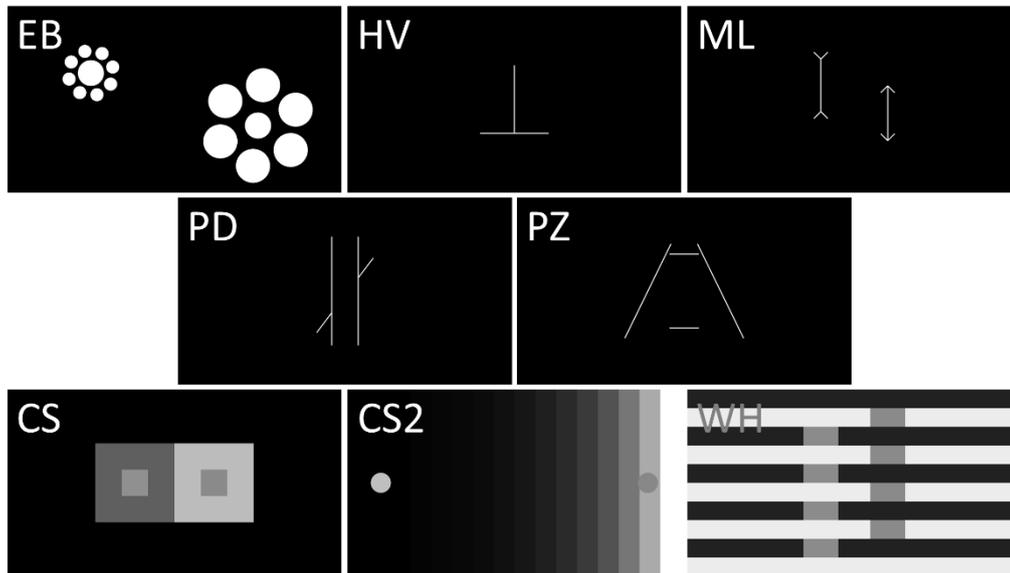


Figure 10. Study 4: Stability over time. Eight illusions were tested: Ebbinghaus (EB), horizontal-vertical (HV), Müller-Lyer (ML), Poggendorff (PD), Ponzo (PZ), two variants of the contrast (CS and CS2), and White (WH). Reprinted from Cretenoud, Grzeczowski, et al. (submitted).

To estimate the effect of time (i.e., sessions) on the illusion magnitudes, I computed a repeated measures analysis of variance (rmANOVA) with illusion and session as within-subject factors. There was a significant interaction ($F(77, 1001) = 1.565, p = 0.002, \eta_p^2 = 0.107; BF = 3 \times 10^{270}$) and a significant main effect of illusion ($F(7, 91) = 28.14, p < 0.001, \eta_p^2 = 0.684; BF = 1 \times 10^{281}$) but the main effect of session was not significant ($F(11, 143) = 0.707, p = 0.73, \eta_p^2 = 0.052; BF = 7 \times 10^{-6}$). The data from the different illusions were further subjected to separate one-way repeated measures ANOVAs with session as within-subject factor. The one-way rmANOVAs revealed a significant main effect of session in the Ebbinghaus illusion and in one variant of the contrast (CS2) illusion (EB: $F(11,143) = 1.893, p = 0.045, \eta_p^2 = 0.127, BF = 0.788$; CS2: $F(11,143) = 2.679, p = 0.004, \eta_p^2 = 0.171, BF = 7.939$) but not in the other illusions (HV: $F(11,143) = 1.687, p = 0.082, \eta_p^2 = 0.115, BF = 0.435$; ML: $F(11,143) = 1.034, p = 0.420, \eta_p^2 = 0.074, BF = 0.073$; PD: $F(11,143) = 1.452, p = 0.156, \eta_p^2 = 0.100, BF = 0.226$; PZ: $F(11,143) = 0.895, p = 0.547, \eta_p^2 = 0.064, BF = 0.050$; CS: $F(11,143) = 0.780, p = 0.659, \eta_p^2 = 0.057, BF = 0.037$; WH: $F(11,143) = 1.473, p = 0.148, \eta_p^2 = 0.102, BF = 0.240$). Hence, there were differences in the mean magnitudes of the Ebbinghaus and one variant of the contrast illusion across sessions (Figure 11). However, this does not imply that individual differences are not stable. For example, it may be that all participants undergo a similar change in the susceptibility to one illusion over time (e.g., a 20%

decrease within a certain amount of time). In this case, individual differences may remain stable, which would show up as strong and significant between-session within-illusion correlations.

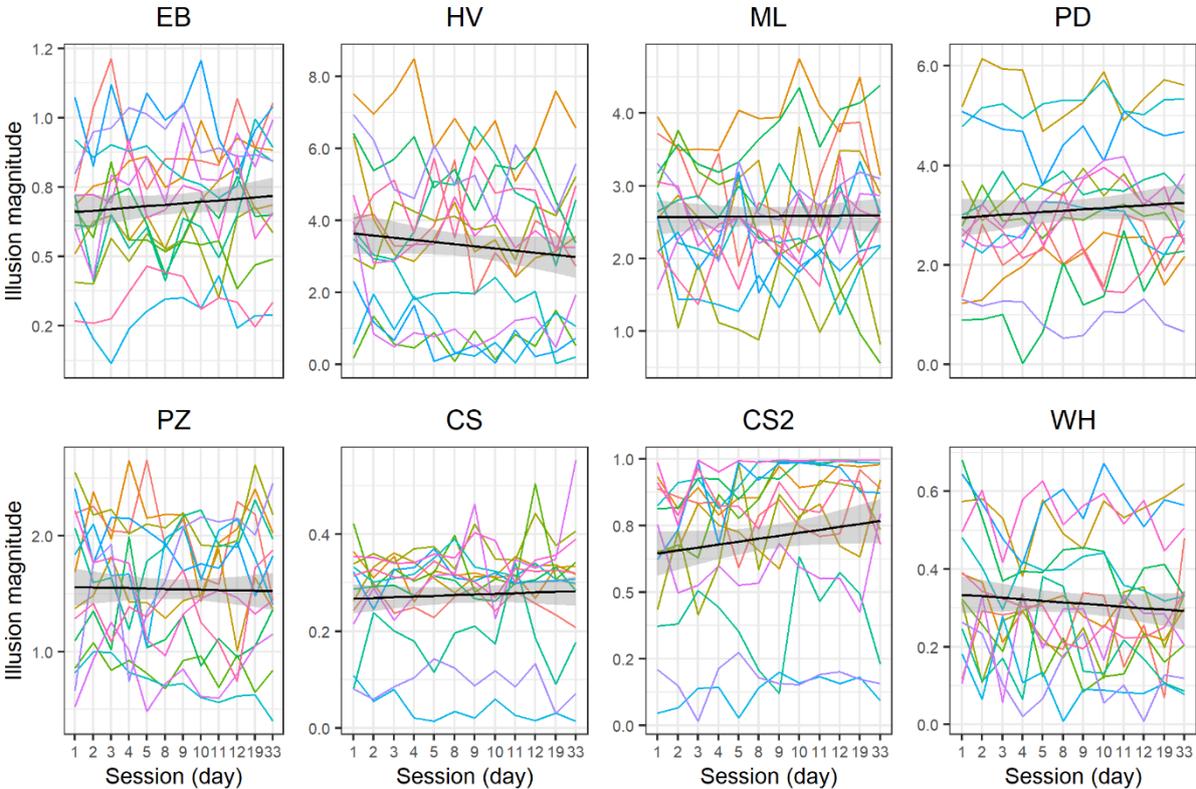


Figure 11. Study 4: Stability over time. Illusion magnitudes plotted for each illusion (sub-panels) across sessions. Colors represent individuals. A linear regression (black line) was fitted with 95% confidence interval (shaded area). Illusion magnitudes are plotted in units of arcdeg for the EB, HV, ML, PD, and PZ illusions and in normalized luminance for contrast illusions (CS, CS2, WH). Reprinted from Cretenoud, Grzeczowski, et al. (submitted).

Hence, I computed between-session correlations for each illusion (Figure 12) to examine how individual differences evolve with time (i.e., sessions). Correlations were strong in general and tended to weaken between pairs of sessions that were further apart in time (especially in the Müller-Lyer and Ponzo illusions).

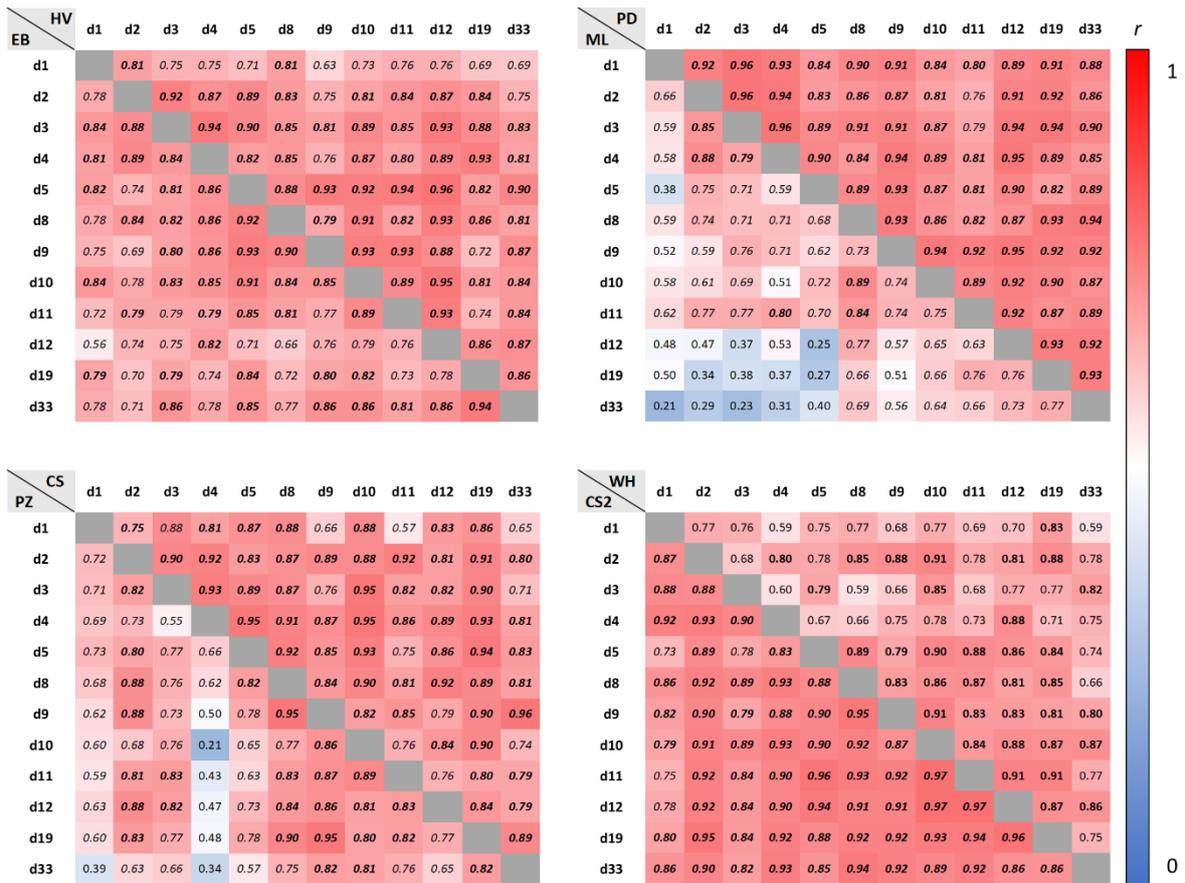


Figure 12. Study 4: Stability over time. Between-session within-illusion correlation coefficients (r). Upper left panel – lower triangle: Ebbinghaus (EB) illusion; upper left panel – upper triangle: horizontal-vertical (HV) illusion; upper right panel – lower triangle: Müller-Lyer (ML) illusion; upper right panel – upper triangle: Poggendorff (PD) illusion; lower left panel – lower triangle: Ponzó (PZ) illusion; lower left panel – upper triangle: contrast illusion (CS); lower right panel – lower triangle: contrast illusion (CS2); lower right panel – upper triangle: White (WH) illusion. Color coding is from blue ($r = 0$) to red ($r = 1$). Only positive correlations were observed. Italics and bold font indicate significant results without ($\alpha = 0.05$) and with ($\alpha = 0.05/66$) Bonferroni correction, respectively.

Then, I wondered whether such patterns of correlations may result from stationary time series, i.e., stable individual differences over time. For this purpose, I simulated the individual illusion magnitudes from normal distributions centered on the behavioral data. In other words, for each participant and each illusion, I computed the mean magnitude and standard deviation across sessions and randomly picked 12 values (for the 12 sessions) from a normal distribution centered and scaled on these values. Between-session within-illusion correlations were then computed and averaged across 10,000 simulations. Behavioral and simulated correlation coefficients are shown in Figure 13 as a function of the time-lag (i.e., the time difference in days) between each pair of sessions.

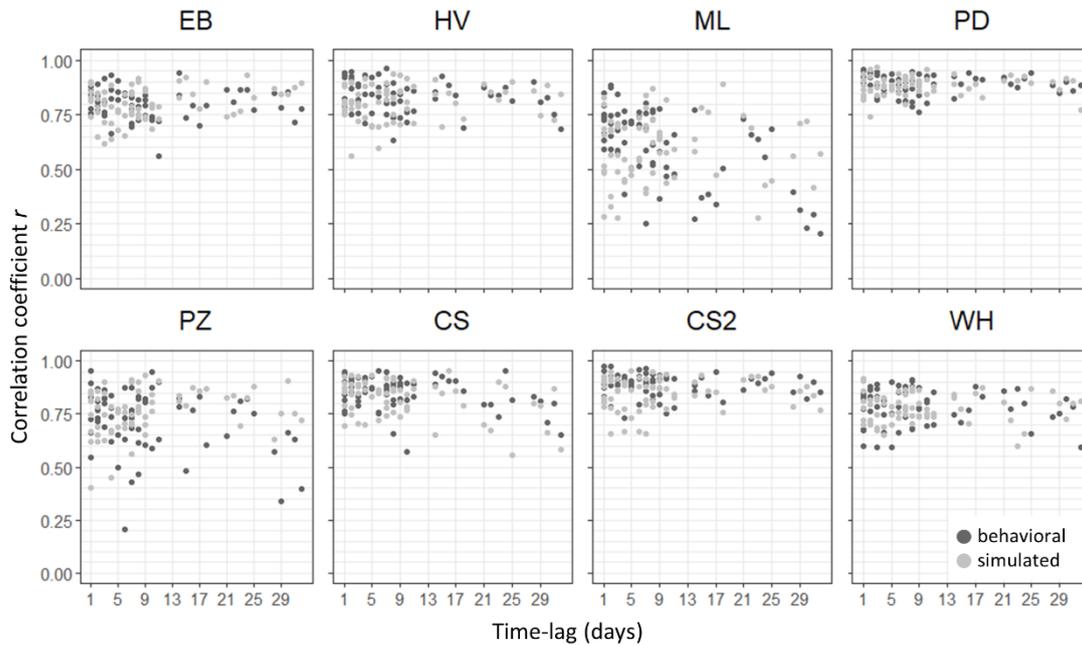


Figure 13. Study 4: Stability over time. Between-session within-illusion correlation coefficients (r) as a function of the time-lag (i.e., the time difference) between sessions in days from behavioral data (dark grey) and simulated data (light grey). Correlations from simulated data were expected to be stronger than correlations from behavioral data under the hypothesis that individual differences in visual illusions vary with time. However, there were no significant differences between simulated and behavioral correlation coefficients. Reprinted from Cretenoud, Grzeczowski, et al. (submitted).

If there were a time effect on the individual differences in the perception of visual illusions, correlations from simulated data should be stronger than correlations from behavioral data and the difference may strengthen with the time-lag (i.e., when two sessions are farther away in time). Paired two-tailed t -tests were computed between the behavioral and simulated correlation coefficients for each illusion. All of them resulted in a non-significant difference with a Bayes factor (BF) smaller than 0.33, except for the second variant of the contrast illusion, which showed inconclusive BF (CS2: $t(65) = 1.879$, $p = 0.065$, $d = 0.231$, BF = 0.705). Overall, these results suggest that individual differences in visual illusions are stable across sessions, i.e., over time.

4.2.3. Study 4: Stable individual differences across measurement methods⁸

In all previous experiments (including Grzeczowski et al., 2017), an adjustment procedure was used to measure the susceptibility to visual illusions. However, forced choice response modalities, e.g., a method of constant stimuli, are usually thought to be more reliable (e.g., Todorović & Jovanović, 2018).

⁸ This section is based on Cretenoud, A. F., Grzeczowski, L., Kunchulia, M., & Herzog, M. H. (submitted). Individual differences in the perception of visual illusions are stable across eyes, time, and measurement methods. (Appendix D; Experiment 3)

Here, I compared the individual differences when two visual illusions were tested with an adjustment procedure and a method of constant stimuli.

Twenty participants were tested with two illusions: the Ebbinghaus (EB) and Müller-Lyer (ML) illusions (see Figure 10). Both illusions were tested with an adjustment procedure, as previously, and with a method of constant stimuli (MCS). In each illusion, one element was in turn the target or reference, i.e., each illusion was tested with two reference-dependent conditions. For example, the central disk surrounded by large flankers was the target in the EB large condition, while the central disk surrounded by small flankers was the target in the EB small condition.

For the adjustment procedure, each illusion and reference-dependent condition was tested twice (2 illusions x 2 reference-dependent conditions x 2 trials = 8). In addition, each illusion was tested twice without flankers (EB) or arrows (ML), i.e., as control conditions (2 illusions x 2 trials = 4).

For the method of constant stimuli, participants had to report whether the target was larger or smaller than the reference by using the up or down key on the keyboard, respectively. The target in each illusion and reference-dependent condition was tested 20 times with eight different increments (i.e., eight sizes; 2 illusions x 2 reference-dependent conditions x 8 increments x 20 trials = 640). For each illusion and each reference-dependent condition, the incremental range was centered on the mean adjustment, i.e., the point of subjective equality (PSE), from a previous experiment (section 4.2.2) and covered three times the absolute difference between the mean adjustment from that previous experiment and the reference. If a participant shows an illusion bias similar to the average bias observed previously, 50% of all responses related to that illusion are expected to be “larger”. Similarly, 50% of the responses are expected to be “larger” if a participant has no illusory bias. In addition, control conditions were tested with eight increments equally distributed around the reference value (2 illusions x 8 increments x 20 trials = 320).

For both methods, the position of the target and reference on the screen was counterbalanced, i.e., the reference was displayed in the right half screen in half of the trials and in the left half screen in the other half of the trials. For the method of constant stimuli, a red square cue (0.5° in side, vertically centered) displayed on the left or right of the stimulus indicated which element was the target.

Stimuli were shown until a response was given, i.e., a mouse click with the adjustment procedure or a key press with the method of constant stimuli. There was no time restriction and no feedback. Psychometric curves were fitted for each condition with the method of constant stimuli to determine the PSE, i.e., the size of the target needed so that the participant perceived both the target and reference to be the same size (2 illusions x (2 reference-dependent conditions + 1 control) = 6

conditions). To this end, I summed the reports when the target was perceived as “larger” than the reference at each increment. Using a cumulative Gaussian function with a lapse rate of 0.02, the PSE was then computed as the size of the target corresponding to 50% of “larger” responses. Data was screened for invalid psychometric fits, i.e., no PSE was extracted when the μ or σ parameter of the underlying Gaussian function was at the edge or outside of the search space, which was defined as a function of the incremental range.

The non-control conditions showed significant test-retest reliabilities (i.e., intraclass correlations), except for the ML inward condition, which only showed a small and non-significant effect size. Importantly, the control conditions of both illusions showed weaker test-retest reliabilities, which suggests that there is no strong measurement bias and that the residual bias is mainly due to noise.

To compare the illusion magnitudes from the adjustment procedure and the method of constant stimuli, I first accounted for the bias in the control conditions. Hence, for each illusion and each method, the magnitude of the control condition was subtracted from the magnitudes of both reference-dependent conditions, e.g., the EB control condition was subtracted from the EB large and EB small conditions in the adjustment method, and similarly in the method of constant stimuli. Then, paired t -tests and Pearson correlations were computed between the illusion magnitudes from both methods.

Results are reported in Table 7. None of the four t -tests revealed a significant difference. According to the Bayes factors, there was substantial support for the null hypothesis in the EB small condition, i.e., both methods led to similar illusion magnitudes. The EB large, ML inward, and ML outward conditions were associated with inconclusive BFs. Correlations between the illusion magnitudes from both methods resulted in large and significant effect sizes for the EB large, EB small, and ML outward conditions (Table 7 and Figure 14). However, the ML inward condition resulted in a smaller and non-significant effect size of $r = 0.324$ with $BF = 1.021$.

Altogether, these results suggest that the individual differences in the perception of visual illusions are overall reliable when measured with an adjustment procedure or a method of constant stimuli. Similarly, Schwarzkopf and colleagues (2011) observed strong correlations between an adjustment procedure (with 8 trials per condition) and a method of constant stimuli in the estimation of susceptibility to the Ebbinghaus and Ponzo corridor illusions.

Table 7. Study 4: Stability across measurement methods. Statistics from paired *t*-tests and correlations between illusion magnitudes measured with an adjustment procedure and a method of constant stimuli, after accounting for the bias in the respective control conditions. Reprinted from Cretienoud, Grzeczowski, et al. (submitted).

		EB large	EB small	ML inward	ML outward
Paired <i>t</i>-tests	<i>t</i>	-0.953	0.617	1.648	1.662
	<i>p</i>	0.353	0.544	0.116	0.113
	BF	0.347	0.276	0.735	0.748
Correlations	<i>r</i>	0.602	0.671	0.324	0.615
	<i>p</i>	0.005	0.001	0.164	0.004
	BF	10.355	27.743	1.021	12.200

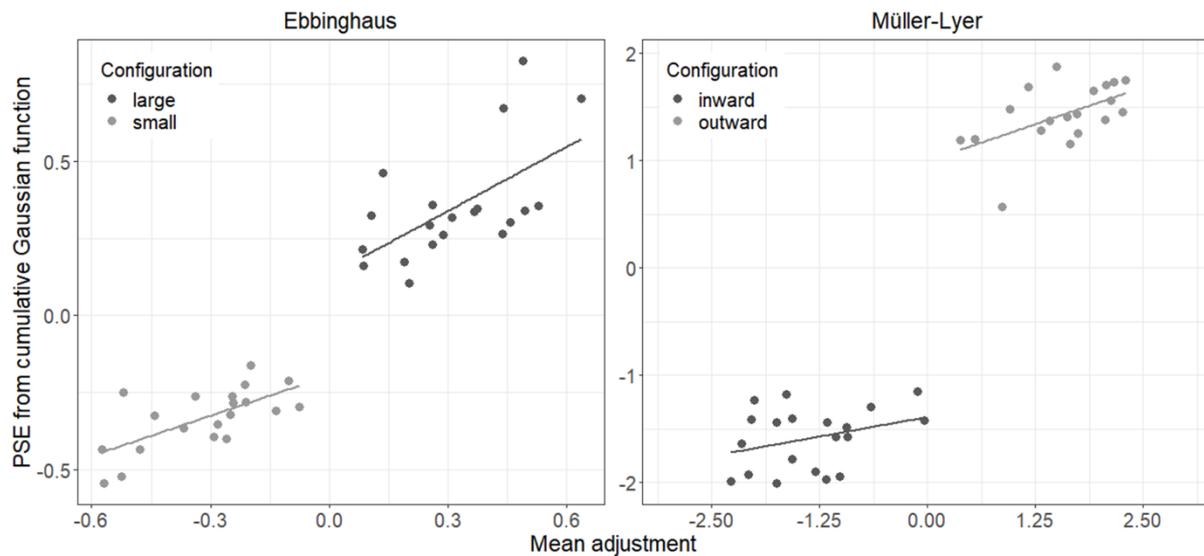


Figure 14. Study 4: Stability across measurement methods. Correlations between illusion magnitudes measured with an adjustment procedure and a method of constant stimuli (in units of arcdeg) for both reference-dependent conditions of the Ebbinghaus (left panel) and Müller-Lyer (right panel) illusions, after accounting for the bias in the respective control conditions. Reprinted from Cretienoud, Grzeczowski, et al. (submitted).

4.3. Is there a common factor for vision in special populations?

As mentioned earlier, unlike cognition, there seems to be no strong common factor for vision. However, it may be that a strong common factor for vision emerges in special populations, e.g., with the general perceptual decline observed when aging. Here, I investigated whether performances in various visual paradigms strongly correlate (i.e., whether there is a strong common factor for vision) in healthy elderly (Study 5), patients with schizophrenia (Study 6), and action video game players (Study 7).

4.3.1. Study 5: Healthy aging⁹

Aging may affect certain people more than others, which may, for instance, be due to differences in optical decline and lifestyle. For this reason, even though there is no strong common factor for vision in young people, it is possible that such a factor emerges as one ages. Here, a battery of perceptual (and a few cognitive) paradigms was tested in 104 healthy younger and 92 healthy elderly participants.

All participants performed 10 visual paradigms (from which 16 variables were extracted): vernier discrimination (duration and offset), visual backward masking (with a 5-element or 25-element grating), visual acuity, orientation discrimination, contrast sensitivity, motion direction discrimination, biological motion perception (upward or inverted, presented for 200 or 800 ms), simple reaction time, visual search (slope and reaction time), and a Simon test. Participants performed three cognitive paradigms as well (verbal fluency, digit span forward and backward, and Wisconsin Card Sorting Test) and answered three questionnaires: the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), the Geriatric Depression Scale (GDS; Yesavage et al., 1982) and the Autism-spectrum Quotient questionnaire (AQ; Baron-Cohen et al., 2001). Participants with a score lower than 26 to the MoCA were excluded from the experiment, as it indicates mild cognitive impairments.

Older participants performed significantly worse compared to young participants in all paradigms (i.e., in all variables), except for the verbal fluency paradigm. In contrast, there was no significant group difference in AQ and GDS scores. Between-variable correlations are shown for the young group in Figure 15 and for the older group in Figure 16. Correlations were weak in general in both groups, except between two variables that were extracted from the same paradigm, e.g., between the 200 ms and 800 ms inverted biological motion perception.

A measure of sampling adequacy (MSA) indicated mediocre (but acceptable, i.e., $MSA \geq 0.5$) intercorrelations among the variables in both groups (0.69 for the young group; 0.62 for the older group). EFAs were computed for each group and the optimal number of factors was determined based on the Akaike criterion (i.e., the smaller, the better). A varimax (i.e., orthogonal) rotation was used to ensure that all factors are independent of each other. The best fitting model according to the Akaike criterion was obtained with 3 factors for the young group and 4 factors for the older one (Figure 17). The models explained 35% and 29% of the variance in the data, respectively. However, these models did not reveal a clear factor structure, where each factor is associated with specific variables. Variables rather loaded on many factors with often weak loadings. In addition, the uniqueness was larger than

⁹ This section is based on Shaqiri, A., Pilz, K. S., Cretenoud, A. F., Neumann, K., Clarke, A., Kunchulia, M., & Herzog, M. H. (2019). No evidence for a common factor underlying visual abilities in healthy older people. *Developmental Psychology*, 55(8), 1775–1787. (Appendix E)

the communalities for most variables, indicating that the variance accounted for by the factors is smaller than the unexplained one.

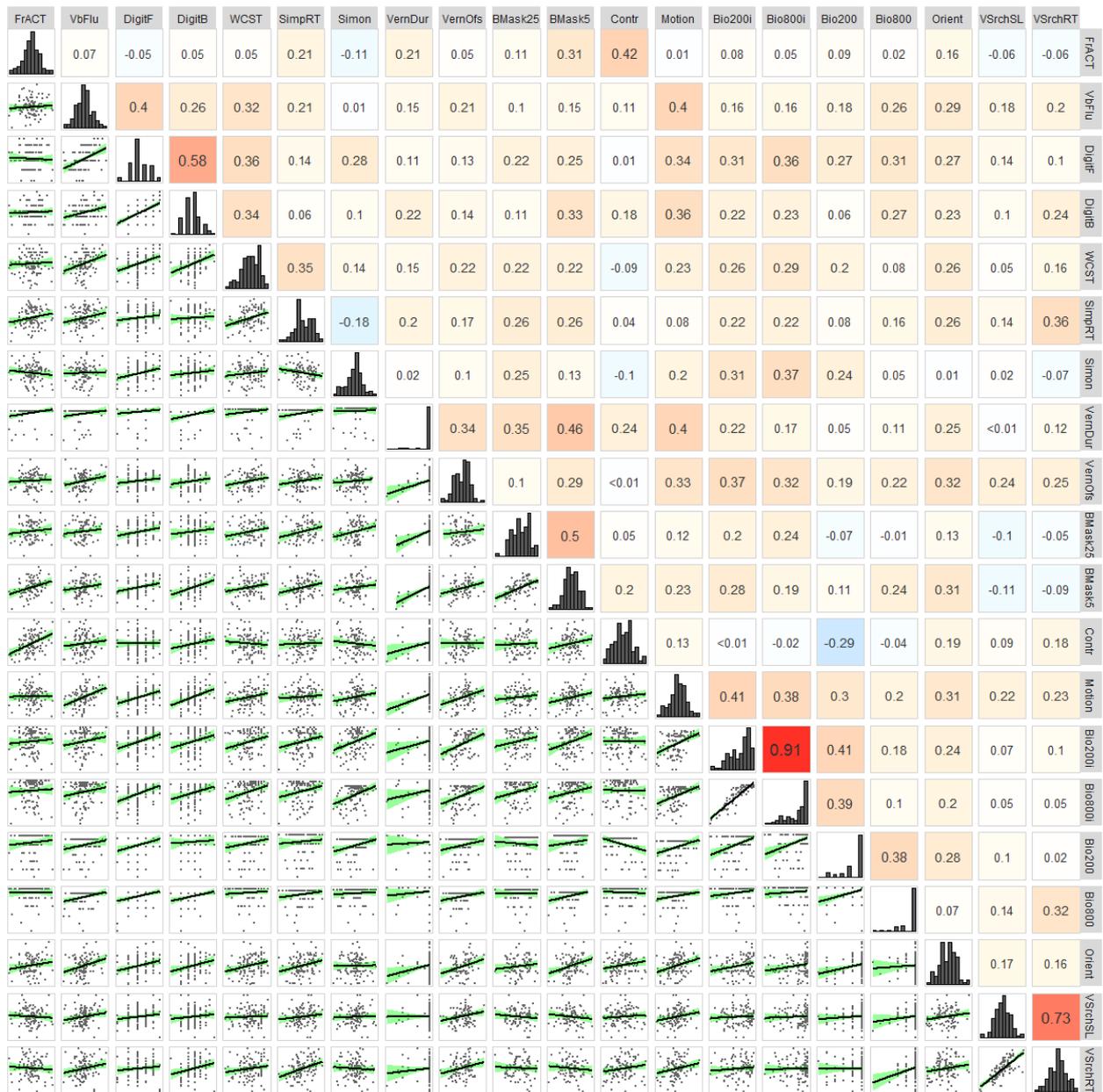


Figure 15. Study 5. Between-variable correlations for the young group. Upper triangle: Pearson correlation coefficients (r) with a color scale from blue to red reflecting effect sizes from $r = -1$ to $r = 1$ (white corresponds to $r = 0$). Lower triangle: scatter plots of the between-variable correlations with regression line and 95% confidence interval. Diagonal: histogram of the scores of each variable. Vernier duration (VernDur) and upward biological motion perception (Bio200 and Bio800) showed ceiling effects. FrACT: Freiburg visual acuity test; VbFlu: verbal fluency; DigitF and DigitB: digit span forward and backward, respectively; WCST: Wisconsin Card Sorting Test; SimpRT: simple reaction time; Simon: Simon test; VernDur and VernOfs: vernier duration and offset, respectively; BMask25 and BMask5: visual backward masking with 25- and 5-element grating, respectively; Contr: contrast sensitivity; Motion: motion direction discrimination; Bio200i and Bio800i: inverted biological motion perception for 200 ms and 800 ms, respectively; Bio200 and Bio800: upward biological motion perception for 200 ms and 800 ms, respectively; Orient: orientation discrimination; VSrchSL and VSrchRT: visual search slope and response time, respectively. Reprinted from Shaqiri et al. (2019) with permission from the American Psychological Association (APA).

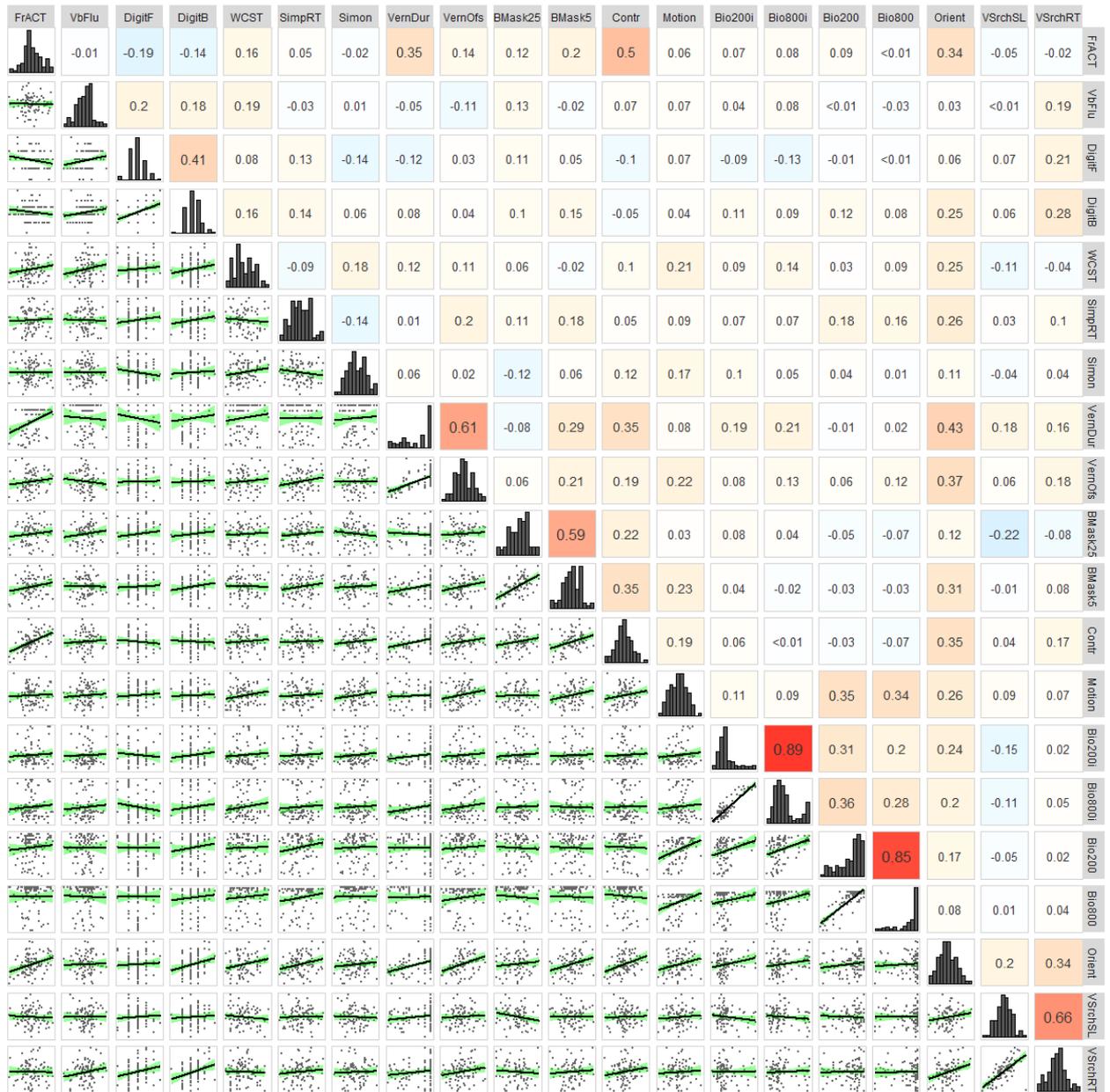


Figure 16. Study 5. Between-variable correlations for the older group. Same format as in Figure 15. Reprinted from Shaqiri et al. (2019) with permission from APA.

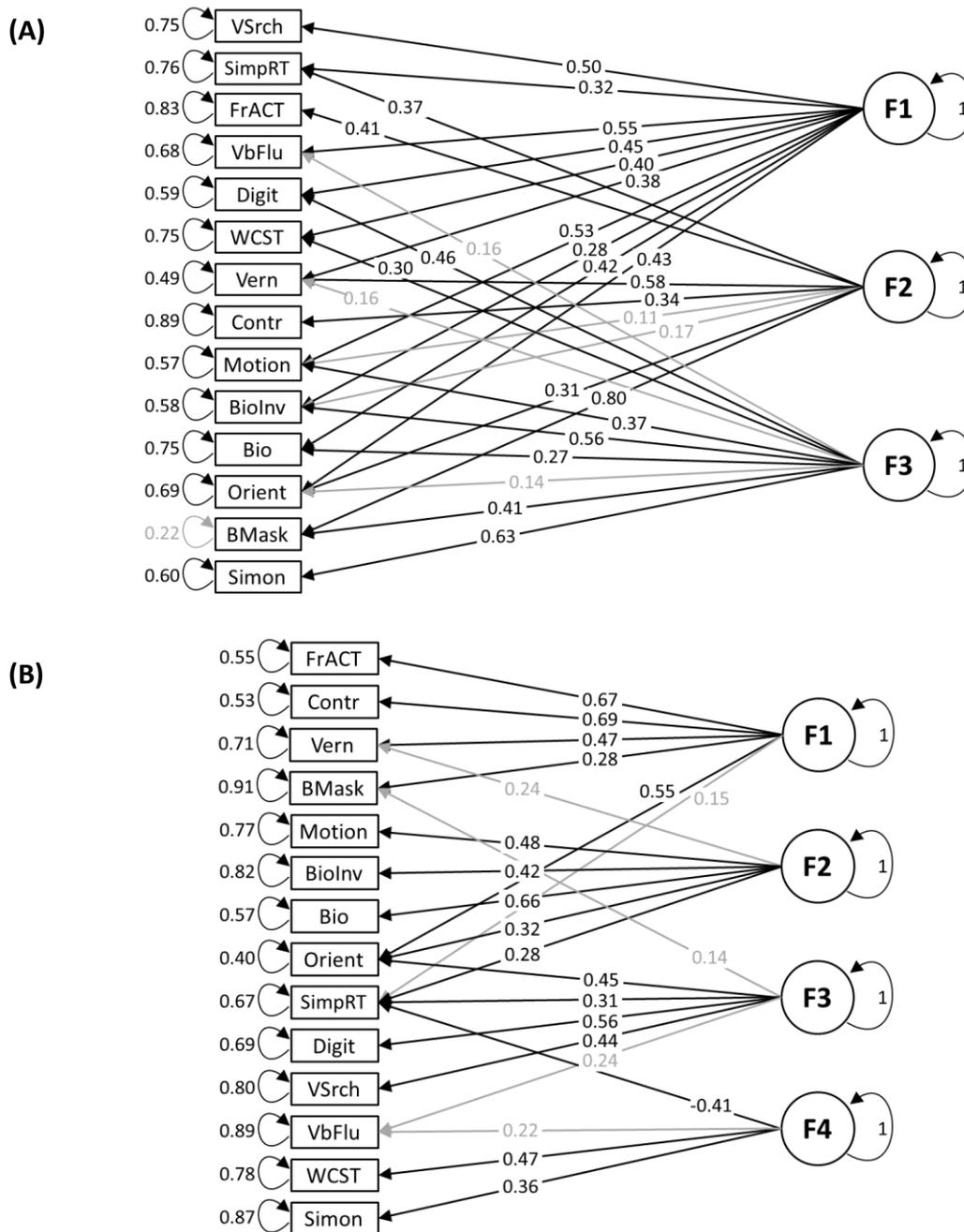


Figure 17. Study 5. Factor structure of the best fitting model according to the Akaike criterion for the (A) young (3-factor model) and (B) older (4-factor model) group. Loadings larger than 0.1 and uniqueness values ($u^2 = 1 - \text{communalities}$) are shown. Grayed-out loadings were not significantly different from zero ($p > 0.05$). Adapted from Shaqiri et al. (2019) and reprinted with permission from APA.

Importantly, the cross-sectional design of this study does not allow to draw conclusions on how perception changes with age. However, the results suggest that no strong common factor for vision emerges with healthy aging.

4.3.2. Study 6: Patients with schizophrenia

In vision, only weak between-variable correlations have been observed in healthy populations (including elderly). Strong between-variable correlations may be observed in clinical populations, such as patients with schizophrenia, where neurophysiological and neuropsychological functions are often reported to be affected (e.g., Braff et al., 2007; Chkonia et al., 2010; da Cruz, Favrod, et al., 2020; Dickinson et al., 2011; Price et al., 2006; Seidman et al., 2015). If patients with schizophrenia always perform worse than healthy controls, then a strong common factor underlying vision may be expected in patients, which I tested here.

Sixty-eight patients with schizophrenia (SCZ) and 60 healthy controls (CON) were included in this study. Group characteristics and statistics are depicted in Table 8. Out of 68 SCZ, 67 were receiving neuroleptic medication.

Table 8. Study 6. Group characteristics ($M \pm SD$) and statistical analysis of the group characteristics

	CON	SCZ	CON vs SCZ
N	60	68	
Gender (F/M)	31/29	15/53	$\chi^2(1) = 12.137, p < 0.001$
Age	37.58 \pm 8.80	38.96 \pm 8.37	$t(122.218) = -0.901, d = -0.160, p = 0.370$
Education	15.55 \pm 2.78	13.33 \pm 2.43	$t(118.180) = 4.782, d = 0.850, p < 0.001$
Illness duration (months)		14.43 \pm 8.20	
SANS		10.04 \pm 4.98	
SAPS		8.47 \pm 2.88	
Handedness (L/R)	3/57	1/67	$\chi^2(1) = 1.312, p = 0.252$
Visual acuity	1.65 \pm 0.43	1.30 \pm 0.45	$t(125.187) = 4.500, d = 0.796, p < 0.001$
In/Out patients		12/56	

Post-hoc Welch's t -tests were computed for ordinal variables. Chi-square tests were performed for categorical variables. Bold indicates significant p -values after Holm-Bonferroni correction for multiple comparisons. Negative (SANS; Andreasen, 1984a) and positive (SAPS; Andreasen, 1984b) symptoms were assessed by an experienced psychiatrist. CON = healthy controls; SCZ = patients with schizophrenia.

Participants were tested with nine paradigms, which tested perceptual, cognitive, and neurophysiological processing: occlusion (Occlusion1, Occlusion2, Occlusion3, Occlusion4, Occlusion5), free viewing (FreeViewing), delay-discounting (Discounting), coherent motion (CohMotion), pro- and anti-saccades (ProTravel, AntiTravel, ProSaccade, AntiSaccade), visual illusions (EB, EB big, EB small, ML, PZ, PZ bw, PZ grid, SC, WH, TT), Stroop (StroopCon, StroopInc), Self (SelfMatch, SelfUnmatch, FamMatch, FamUnmatch, UnfamMatch, UnfamUnmatch), and ambiguous figures (AmbFig1, AmbFig2, AmbFig3, AmbFig4). For six out of nine paradigms, more than one variable was extracted with a total of 34 variables (spelled out in brackets) considered for analysis.

First, the performances of CON were compared to the performances of SCZ in all variables. There were significant differences in the SelfMatch, SelfUnmatch, FamMatch, FamUnmatch, and AmbFig3 variables after correcting for multiple comparisons (Holm-Bonferroni). In addition to these variables, Bayesian independent samples *t*-tests suggested evidence for a significant difference ($BF < 0.33$) between CON and SCZ in the FreeViewing, SC, StroopCon, StroopInc, UnfamMatch, UnfamUnmatch, AmbFig2, and AmbFig4 variables.

Second, between-variable correlations for the CON (upper triangle) and SCZ (lower triangle) were computed (Figure 18). Most correlations were low and non-significant. Correlation coefficients *r* were transformed to *z* statistics (Fisher's *r* to *z* transformation) and compared between groups. The proportion of significant comparisons between CON and SCZ was 7.5%. Note that the aim here was not to point at specific pairwise correlations but rather at the general pattern of correlations underlying the dataset, which was similar across both groups. Importantly, the sign of variables was not flipped when higher values indicate lower performance (e.g., variables related to the Self paradigm). Hence, if there truly was a common cause, an ensemble of both positive and negative, strong correlations would be expected.

Third, latent variable models were estimated in R (R Core Team, 2018) using the *lavaan* package (Rosseel, 2012). Missing values were imputed using a full information maximum likelihood (FIML) estimator. Multiple fit indices were checked to evaluate the model fit (Hu & Bentler, 1998, 1999): χ^2 goodness-of-fit statistic, comparative fit index (CFI), root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), Akaike information criterion (AIC), and Bayesian information criterion (BIC). Note that the structural equation modeling (SEM) was used in an exploratory rather than confirmatory way. Fit indices were not expected to be very good, since I did not aim at validating a scale as usually the case with confirmatory factor analyses. Here, SEM was used as a tool to examine the probability of a single, general versus multiple, specific cause(s) underlying the different paradigms and to explore the differences between CON and SCZ.

The *a priori* hypothesis, i.e., the first model tested (Model A; Figure 19A), was that there is one latent variable for each paradigm. For instance, all five variables extracted from the occlusion paradigm, i.e., Occlusion1 to Occlusion5, were expected to strongly load on an Occlusion factor. Note that the paradigms, from which one variable only was extracted (i.e., free viewing, discounting, and coherent motion paradigms), were discarded from the SEM models because a latent factor is not supposed to be related to one variable only. Similarly, the ML, WH, SC, and TT illusions were discarded since illusion-specific factors were previously observed. Also note that the terms "latent variables" and "factors" are used interchangeably.

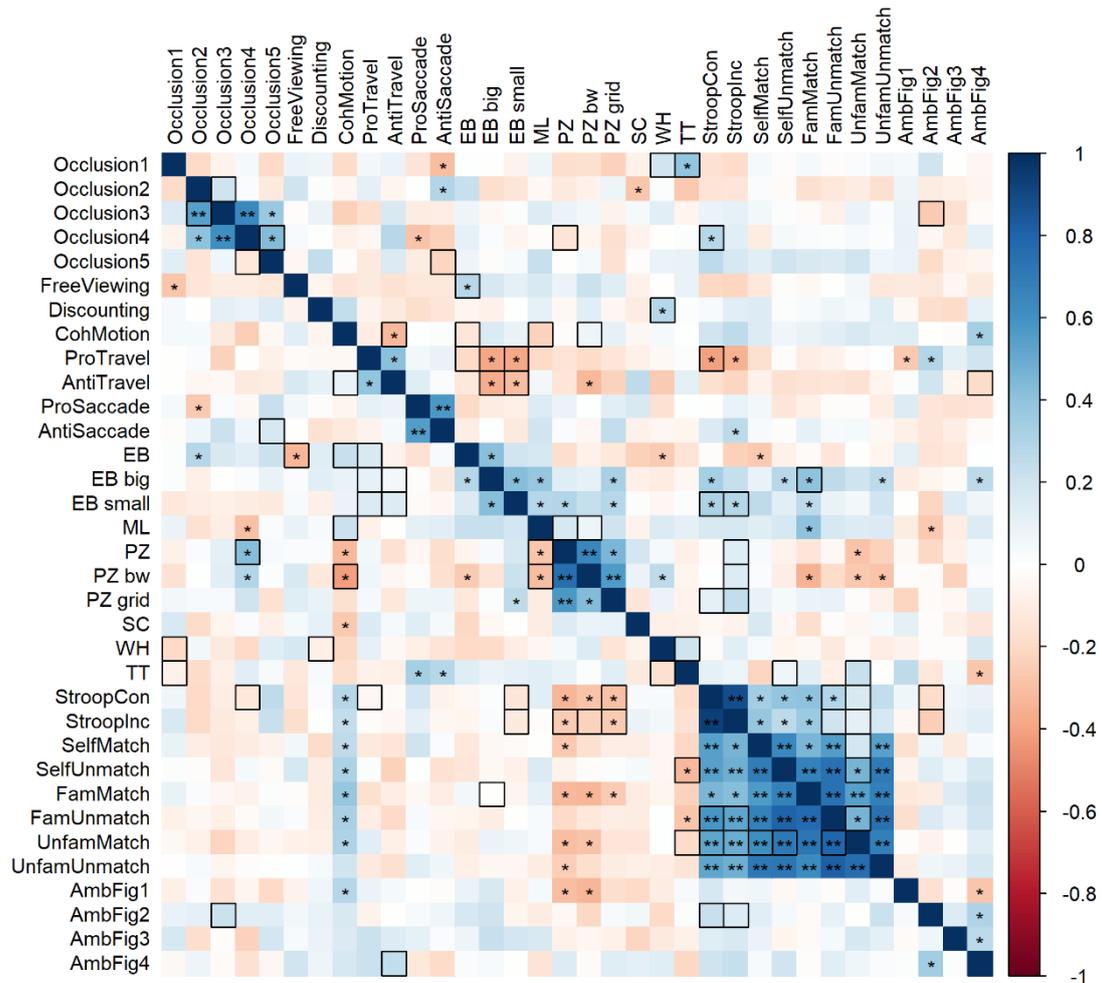


Figure 18. Study 6. Between-variable correlation coefficients r in the group of healthy controls (upper triangle; $n_{CON} = 60$) and in the group of patients with schizophrenia (lower triangle; $n_{SCZ} = 68$). The color scale from blue to red reflects effect sizes from $r = -1$ to $r = 1$ (*uncorrected p values < 0.05 ; **Bonferroni corrected p values < 0.05). Each correlation coefficient r was then transformed to a z value (Fisher's r to z transformation) to be comparable across groups. Outlined squares indicate correlations, which were significantly different between controls and patients with schizophrenia in terms of z ($p < 0.05$).

Model A (Figure 19A) therefore fitted a model with six factors, which were allowed to covary. As shown in Table 9, the model provided an overall acceptable fit to the data. Most manifest variables showed significant loadings on the corresponding latent variable. To investigate whether inter-factor correlations are important to the model, the same model was fitted with the constraint that the latent variables were assumed to be orthogonal (Model A.2). This model showed a significantly worse fit to the data compared to Model A ($p < 0.001$; see Table 9), suggesting that factors correlate, even though inter-factor correlations were non-significant, except between the Saccade and Stroop factors ($r = -0.315$, $p = 0.004$) and between the Stroop and Self factors ($r = 0.559$, $p < 0.001$).

Table 9. Study 6. Structural equation modeling: goodness-of-fit statistics and model comparison results

Model	χ^2	df	p	CFI	RMSEA [90% CI]	SRMR	AIC	BIC	$\Delta\chi^2$, df
A	421.067	244	< 0.001	0.903	0.063 [0.053, 0.073]	0.078	10589.781	10846.101	
A.2	530.391	259	< 0.001	0.852	0.076 [0.067, 0.085]	0.137	10669.105	10877.366	109.32, 15***
B	983.234	252	< 0.001	0.601	0.126 [0.118, 0.135]	0.106	11135.948	11366.636	562.17, 8***
C	405.406	235	< 0.001	0.907	0.063 [0.053, 0.073]	0.125	10592.120	10877.277	15.661, 9°
D	323.017	238	< 0.001	0.954	0.044 [0.031, 0.056]	0.068	10503.731	10779.275	98.050, 6***
D.2	451.84	306	< 0.001	0.895	0.061 [0.049, 0.073]	0.075	7422.026	77470157	

Note. Model A: specific, correlated factors; Model A.2: specific, uncorrelated factors; Model B: general, common factor only; Model C: bifactor; Model D: specific, correlated factors with two separate factors for the saccade paradigm; Model D.2: specific, correlated factors with two separate factors for the saccade paradigm and with group, gender, education, and visual acuity as covariates. CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual; AIC: Akaike information criterion; BIC: Bayesian information criterion. Changes in χ^2 ($\Delta\chi^2$) of Models A.2, B, C, and D were calculated relative to Model A (° $p < 0.1$; * $p < 0.05$; *** $p < 0.001$).

Model B tested a unique common factor by fitting a single-factor model, in which all manifest variables loaded on a general latent variable (Figure 19B). This model resulted in a bad fit, which was critically worse ($p < 0.001$) than the fit for Model A (Table 9).

Even though the specific-factor model (i.e., Model A) showed a significantly better fit than the single-factor model (i.e., Model B), there is not enough evidence to claim that there is no general common factor underlying the different paradigms. It may be that there is a weak general common factor in addition to more specific ones, as was observed in the perception of visual illusions (see section 4.1). To test this, I fitted a bifactor model (Model C; Figure 19C), in which all manifest variables loaded on a general common factor in addition to factors specific to each paradigm. The general common factor catches the variance, which is not explained by the specific factors (see for example Gignac & Kretzschmar, 2017; Yung et al., 1999). There was no significant difference between the fit of Models A and C, despite a p -value close to the 0.05 alpha level for statistical significance ($p = 0.074$; Table 9), suggesting that it is still not completely unlikely that there is a weak general common factor underlying the paradigms tested in this study.

In the bifactor model, the ProTravel (standardized loading: 0.707) and AntiTravel (standardized loading: 0.470) variables strongly loaded on the general common factor, while these two manifest variables showed weaker loadings on the specific factor for the saccade paradigm (standardized loadings of 0.250 and 0.235, respectively). In contrast, the ProSaccade and AntiSaccade variables weakly loaded on the general common factor (standardized loadings of -0.254 and -0.274, respectively), while they strongly loaded on the specific factor for the saccade paradigm (standardized loadings of 0.749 and 0.740, respectively). Hence, most of the variability in the ProTravel and AntiTravel variables was not caught by the specific factor for the saccade paradigm. Therefore, I tested

a Model D (Figure 19D), in which the factor specific to the saccade paradigm was split into two, i.e., a Saccade factor, on which the ProSaccade and AntiSaccade variables loaded, and a Travel factor, on which the ProTravel and AntiTravel variables loaded. The model fit was significantly better than the fit for Model A ($p < 0.001$; Table 9), suggesting that the general but weak common factor in Model C was mostly explaining the variance in the ProTravel and AntiTravel variables, i.e., it was a specific rather than general factor.

Lastly, I fitted Model D with covariates for group (CON versus SCZ) as well as gender, education, and visual acuity (Model D.2; similarly, an age dichotomous predictor was used in Rey-Mermet et al., 2018), since they significantly differed between CON and SCZ (Table 8). Note that multiple-group SEM were not converging due to the small sample sizes. The model fit was acceptable, and the group covariate significantly regressed onto two factors, i.e., Stroop and Self, suggesting that there are differences in the way that the corresponding manifest variables are represented in both groups. On the contrary, the covariates for gender, education, and visual acuity did not regress significantly onto any factor.

In addition to the small sample sizes, there were two main issues with the models. First, two Heywood cases were detected, i.e., the variance explained of some manifest variables was more than 100%, which is impossible. One such case was detected in the Occlusion factor. The Heywood case was solved by constraining all loadings in the Occlusion factor to be equal. However, the Heywood case in the Ebbinghaus factor, on which all three variants of the Ebbinghaus illusion loaded, i.e., EB, EB big, and EB small, could not be solved in the same way. Hence, the Ebbinghaus factor was discarded. Second, the error variances of some variables were high, i.e., the factors only explained a small proportion of the variance in some variables. Therefore, even though Model D.2 showed an acceptable fit to the data, it has low explanatory power.

To sum up, a multiple-factor model with seven factors resulted in a better fit compared to a single-factor model. The two groups did not differ significantly in the evaluation of all factors (i.e., only in some), which suggests that schizophrenia is a heterogeneous disease and that there is no strong common factor emerging with the disease. It is also important to highlight that patients with schizophrenia did not perform worse than healthy controls in all paradigms (i.e., only in some).

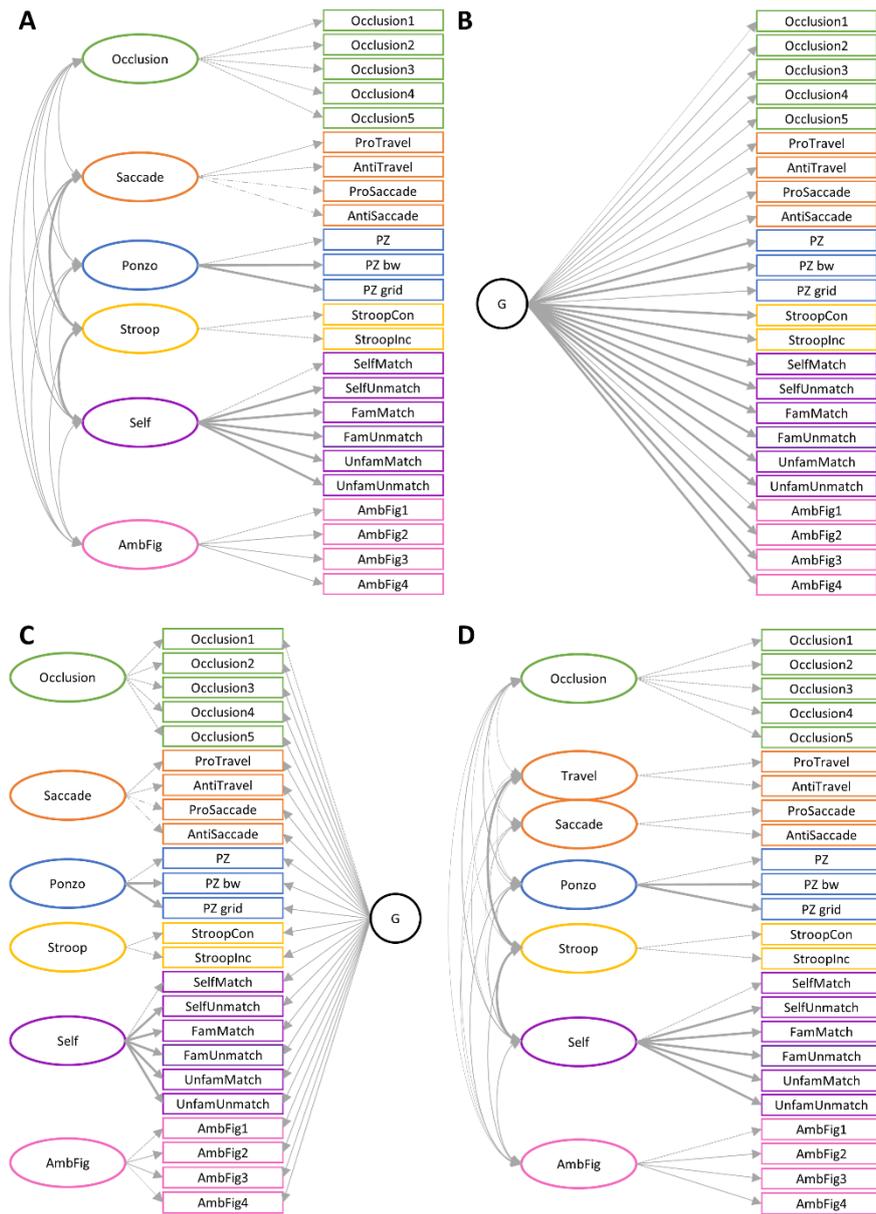


Figure 19. Study 6. Structural equation modeling: (A) Model A is a multiple, correlated factor model. Note that Model A.2 (not illustrated here) is a multiple, uncorrelated factor model (i.e., same as Model A, but without the covariances between factors). (B) Model B is a single-factor model. (C) Model C is a bifactor model with both specific factors from Model A and a general, common factor as in Model B. (D) Model D is a multiple, correlated factor model, which is similar to Model A, except for the fact that the factor for the saccade paradigm was split into two factors. For all models, dashed lines indicate variables that were used to set the scale of the factors to 1 or loadings that were equalized because of Heywood cases (i.e., variables showing negative variance). Double-headed arrows represent the covariances between factors. Bold arrows indicate significant estimates ($p < 0.05$). Loadings are not shown for the sake of readability.

4.3.3. Study 7: Action video game players¹⁰

It has been claimed that video gamers possess increased perceptual and cognitive skills compared to non-video gamers. Here, I measured visual abilities in CS:GO (Counter-Strike: Global Offensive, i.e., an action video game) players. First, I investigated whether there is a strong common factor for visual perception in action video game players. Second, I wondered whether the players' rank can be predicted from their visual performances.

Ninety-four players ranging from beginners to professionals were tested with a battery of 12 visual paradigms: crowding (CrwSize, CrwPeri), contrast sensitivity (Ctr), the Honeycomb and Extinction illusions (HcBlack, HcWhite, ExBlack, ExWhite), a battery of other illusions (CS, EB, ML, PD, PZ, TT, VH, WH, ZN), N-back (NB), orientation discrimination (Ori), random dot kinematograms (RDKh, RDKr), simple reaction times (RT), pro- and anti-saccades (proTravel, proSac, antiTravel, antiSac), Freiburg visual acuity (VA), visual backward masking (VBM), and visual search (VS4, VS16). In total, 29 variables (spelled out in brackets) were extracted from the visual paradigms. In addition, performance in specific gaming skills (flicking, holding, peeking, shooting, spraying, and tracking) was assessed and personality traits were administered. Note that the sign of visual variables was flipped when necessary, so that higher values always indicate better performance.

Pairwise correlations were computed for all variables extracted, including the players' actual rank (Actual CS:GO rank), players' best rank ever (Best CS:GO rank) and the weekly (NbHoursPerWeek) and total (NbTotalHours) number of hours they played CS:GO (Table 10). In general, correlations were weak between visual variables, except between pairs of variables that were extracted from the same paradigm. In line with previous studies (e.g., Cretenoud et al., 2019; Grzeczowski et al., 2018; Grzeczowski, Clarke, et al., 2017), between-illusion correlations were rather weak ($M_r = 0.108$; $SD_r = 0.102$), except between the Honeycomb and Extinction illusions ($M_r = 0.648$; $SD_r = 0.133$; see Bertamini et al., 2019). Gaming variables showed stronger intercorrelations ($M_r = 0.291$; $SD_r = 0.121$). In addition, correlations between questionnaire variables, which are related to CS:GO (NbHoursPerWeek, NbTotalHours, Actual CS:GO rank, Best CS:GO rank), and gaming variables were high too ($M_r = 0.314$; $SD_r = 0.182$). The correlations between CS:GO related questionnaire variables were even higher ($M_r = 0.586$; $SD_r = 0.256$).

¹⁰ This section is based on Cretenoud, A. F., Barakat, A., Milliet, A., Choung, O.-H., Bertamini, M., Constantin, C., & Herzog, M. H. (submitted). How does action video game performance relate to visual skills? (Appendix F)

Variables that showed an unacceptable measure of sampling adequacy (i.e., MSA < 0.5) were removed sequentially until all variables showed an acceptable MSA. The global MSA index after variable removal was 0.695. Following an EFA, a three-factor model was suggested by scree plot inspection and explained 30% of the variance (RF1: 12%; RF2: 10%; RF3: 8%). Loadings are reported in Table 11. The first factor was mainly composed of the CS:GO related and gaming variables. The second factor was mainly related to the Honeycomb and Extinction variables. The third factor mainly loaded on the visual variables related to reaction times, such as RT, VS4, and VS16, and to the pro- and anti-saccades paradigm. Inter-factor correlations were weak.

Table 11. Study 7. Factor loadings from an EFA after promax rotation for visual (green), gaming (orange), and questionnaire (purple) variables. A color scale from blue (negative loadings) to red (positive loadings) is shown. Absolute factor loadings larger than 0.55 are highlighted (bold) because they are considered as significant with a sample size of 100 (Hair et al., 2018). Reprinted from Cretenoud, Barakat, et al. (submitted).

	RF1	RF2	RF3
Ctr	-0.110	-0.226	0.275
HcBlack	0.011	0.798	0.176
HcWhite	0.070	0.836	0.159
ExBlack	-0.082	0.751	0.251
ExWhite	0.092	0.580	0.182
PZ	-0.301	0.115	0.094
ZN	-0.243	-0.024	0.276
NB	-0.161	-0.093	0.365
RT	0.175	0.049	0.538
proTravel	0.111	-0.038	0.352
proSac	0.174	0.060	0.543
antiSac	0.111	-0.008	0.569
VA	-0.089	-0.200	0.237
VBM	-0.055	-0.267	0.525
VS4	0.060	0.168	0.537
VS16	-0.022	0.272	0.489
Shoot	0.701	0.110	0.108
Spray	0.500	0.043	-0.011
Track	0.596	-0.076	0.133
Hold	0.212	0.020	0.289
Flick	0.407	-0.109	0.022
Peek	0.220	0.069	0.146
AQ	-0.003	-0.263	-0.253
CI	-0.183	0.454	-0.060
HEXACO EM	-0.073	-0.405	0.056
HEXACO EX	-0.095	0.129	0.216
HEXACO AG	0.166	-0.143	0.057
HEXACO CO	-0.087	0.277	-0.128
HEXACO OP	-0.270	0.018	-0.023
Handedness	0.142	0.271	-0.148
PRFd	-0.333	0.186	0.095
NbHoursPerWeek	0.405	0.151	-0.072
NbTotalHours	0.806	0.140	0.054
Actual CS:GO rank	0.848	0.131	0.122
Best CS:GO rank	0.876	0.105	0.083

To predict the actual CS:GO players' rank, an elastic net model (Zou & Hastie, 2005), i.e., a regressor, which both uses L1 and L2 regularizations to reduce the dimensionality of the model, was fitted on the visual, gaming, and questionnaire variables. The dataset was split into a training (80%) and test (20%)

set. Using a search grid with a 5-fold cross-validation, the model's generalization performance was optimized on the training set by tuning two hyperparameters, namely alpha and the L1/L2 ratio. The lower alpha, the more complex the model (i.e., less strict regularization).

Performance on the training set was optimized for $\alpha = 0.15$ and with an L1/L2 ratio of 0.45. With these values for the hyperparameters, the training and test set accuracies were $r^2 = 0.640$ and $r^2 = 0.263$, respectively. The *MSEs* for the training and test sets were 0.21 and 0.25, respectively. In contrast, a dummy regressor resulted in an *MSE* of 0.60 and 0.37 in the training and test sets, respectively.

As expected, most gaming variables showed non-zero coefficients (Shoot: 0.225; Spray: 0.059; Track: 0.169; Flick: 0.002). The results also suggest that the Honeycomb illusion (HcWhite: 0.085) and one crowding variable (CrwSize: 0.031) are predictors of the players' rank, i.e., participants who show better performances in those two tasks tend to have a higher rank. Note that both paradigms are related to visual perception in the periphery. In contrast, participants with higher ranks tend to have worse visual acuity in the fovea (VA: -0.038), which suggests that central and peripheral vision are independently engaged while playing video games. In addition, the results suggest that faster reaction times (proSac: 0.087; VS4: 0.007) and stronger susceptibility to the Zöllner illusion (ZN: -0.110) are associated with higher ranks. Lastly, participants with weaker conscientiousness (HEXACO CO: -0.011), weaker openness to experience (HEXACO OP: -0.099), and with higher score on the Barratt Impulsiveness Scale (BIS: 0.045), tend to have higher ranks.

Hence, rather than a strong common factor for visual perception in CS:GO players, specific visual paradigms seem to be strongly predictive of the players' rank. Even though causative relationships cannot be derived from these results, this investigation gives clues about visual paradigms, such as those measuring visual perception in the periphery, which may be part of future training programs for esports.

5. Discussion

5.1. Summary

Previous studies reported only weak evidence for a strong common factor for vision, i.e., a factor that would explain a substantial proportion of the individual variance across basic visual skills. In addition, only weak correlations were previously observed between the susceptibility to different visual illusions, suggesting that there is no strong common factor for illusions (e.g., Axelrod et al., 2017; Grzeczowski, Clarke, et al., 2017; Taylor, 1976).

Here, I first examined the multidimensionality of the visual space underlying illusions. In Studies 1 and 2, weak between-illusion correlations were observed, as in previous studies. However, I observed strong within-illusion correlations, e.g., between two variants of the same illusion varying in size, suggesting that the factors underlying visual illusions are not specific to the spatial, luminance, orientation, and contextual features. Hence, it seems that more or less each illusion makes up its own factor. Importantly, it is unlikely that the weak between-illusion correlations are due to methodological flaws or lack of power, since strong within-illusion correlations were observed. In Study 3, I investigated whether and how factors for illusions interact. To this aim, pairs of illusions were merged. The susceptibility to a merged illusion was strongly predicted from the susceptibility to the two illusions of which it was made, suggesting that the factors for illusions combine independently. Similar to Lavoisier's famous chemistry principle, which claims that nothing is lost, nothing is created but everything is transformed (Lavoisier, 1789), it seems that in the merged illusions, no factor is lost, and no factor is created. Factors are not even transformed; they are just combined.

Second, I wondered whether the large individual differences in illusion susceptibility are robust over manipulations of the eye(s) to which the stimulus is presented, and whether they are stable within a month period as well as independent of the measurement method (adjustment procedure versus method of constant stimuli). The results from Study 4 suggest that this is the case. Hence, the mixed results previously reported regarding the existence of a unique, strong common factor or several more specific factors for illusions are unlikely to be related to unstable individual differences across eyes, time, and measurement methods.

Third and last, I investigated whether a strong common factor for vision emerges in special populations that tend to show altered performances in visual tasks compared to the average healthy young adult. It may be expected that there are strong between-task correlations in special populations (or at least stronger than in control groups), either because of a general perceptual decline (healthy elderly in Study 5; patients with schizophrenia in Study 6) or contrarily because of a general enhancement of perceptual skills (action video game players in Study 7). For example, if there is a general, reliable

performance decline in the visual abilities of patients with schizophrenia, a strong common factor is expected to explain a large proportion of the individual variance. However, the pattern of between-task correlations was similar in healthy aging, patients with schizophrenia, and action video gamers compared to control groups, i.e., correlations were in general weak and non-significant between different tasks, while there were strong correlations between different measures extracted from the same task (which may somehow be considered as a measure of test-retest reliability). These results suggest that no strong common factor for vision emerges in special populations.

5.2. A multifactorial structure for visual illusions does not preclude from a more general – albeit weak – common factor

While most within-task correlations were strong, weaker between-task correlations were observed, suggesting that there is no strong common factor for vision (and more specifically for visual illusions). The heterogeneity of the visual structure seems to be better represented by a multitude of factors for visual abilities, such as a factor specific for the Ebbinghaus illusion (Study 1: Feature-unspecific). However, in addition to many specific factors, it is not completely unlikely that a more general common factor accounts for a certain – likely small – proportion of the individual variability.

To investigate whether a general – albeit weak – common factor for visual illusions coexist with more specific ones, I examined the effect size of the between-illusion correlations in Studies 1, 2, and 3. In Study 1 (luminance-unspecific), when 10 illusions were tested with four luminance conditions, between-illusion correlations were significantly different from zero (one-sample two-tailed t -test: $t[719] = 15.093, p < 0.001, d = 0.562$). In Study 2, the average between-illusion correlation coefficient was $r = 0.171$ ($SD = 0.092$) with 35 out of 36 correlations being positive. A one-sample two-tailed t -test revealed that the effect size was significantly different from zero ($t[35] = 11.146, p < 0.001, d = 1.858$; similar in Study 3), suggesting that a small proportion of the variance underlying the Müller-Lyer and Ponzo illusions is accounted for by a common factor. Note, however, that between-illusion correlations were not significantly different from zero (one-sample two-tailed t -test: $t[159] = -0.780, p = 0.437, d = -0.062$) when five illusions were tested with four different orientations in Study 1 (orientation-unspecific). Importantly, the specific factors extracted from dimensionality reduction techniques showed a large range of intercorrelations, arguing in favor of a weak but general common factor for illusions.

Therefore, even though the structure underlying visual illusions seems to be highly dimensional, different illusions are likely not completely uncorrelated. For example, Coren and colleagues (1976) found that 45 measures of spatial illusions loaded on five specific factors. In addition, a second-order factor solution, which was computed from the intercorrelations between the five first-order factors,

revealed only two more general factors. Hence, it seems that a common mechanism, such as a constancy phenomenon (Gregory, 1963, 1964), may account for a small proportion of the variance in the susceptibility to illusions that were loading on different factors in the first-order solution. However, because within-illusion correlations were stronger and significantly larger than between-illusion correlations, a large part of the variance seems to specifically tap into unique aspects of each illusion.

Similarly, when testing batteries of paradigms (Studies 5 to 7) in special populations, multiple-factor models better fitted the data compared to single-factor models. For example, patients with schizophrenia and healthy controls were tested with several paradigms in Study 6. A seven-factor model showed a better fit to the data from both groups compared to a single-factor model. Rather than suggesting exactly seven unique dimensions to underlie this dataset, our results suggest that a substantially larger proportion of the variance is explained by multiple specific factors compared to a unique common factor. Importantly, it is not excluded that both a general – albeit weak – and more specific factors coexist, as in visual illusions.

While studies usually concluded for (e.g., Halpern et al., 1999) or against (e.g., Cappe et al., 2014) a strong common factor for vision, Bosten and colleagues (2017) reported evidence for both a general factor for visual and auditory perception and for independent perceptual abilities. Indeed, an unrotated factor solution (i.e., identifying a first factor that accounts for a maximum amount of the variability) resulted in a first factor accounting for about 20% of the correlated variance. In contrast, an orthogonal factor solution produced a set of eight factors explaining 57% of the underlying variance (note that only 12% of the variance was accounted for by the first rotated factor). The authors suggested that both unrotated and rotated (orthogonal) factor solutions should be examined when investigating the perceptual space.

5.3. Implications for vision research

The study of individual differences in vision aims at identifying a limited number of factors which underlie all visual mechanisms. Hence, psychological, cognitive, and neurophysiological mechanisms that mediate visual processing should gather into specific categories. Ultimately, unrolling the structure of vision may help better understanding the fundamental mechanisms of perception (Tulver, 2019).

However, only a few factors can hardly explain the individual differences in vision and more specifically in the perception of visual illusions. As mentioned above, there were only weak correlations between the susceptibility to different illusions. For example, out of 15 between-illusion correlations, only one was significant when testing a large sample of participants (Grzeczowski, Clarke, et al., 2017). Similarly, in Study 1 (luminance-unspecific), only 65 out of 720 between-illusion correlations were

significant. These results are surprising since in the latter study, all illusions were spatial illusions, which are often implicitly or explicitly assumed to share a similar mechanism (e.g., Coren et al., 1976; Ninio, 2014; Piaget, 1961). However, a shared mechanism should have led to substantial correlations. The weak correlations cannot be explained by large intraobserver variability because test-retest reliabilities were mostly significant. In addition, individual differences in monocular and binocular viewing conditions are robust, suggesting that the mechanisms underlying illusions occur after binocular fusion (Study 4: Stability across eyes; but see Song et al., 2011; Yildiz et al., 2021). Similarly, individual differences in the perception of visual illusions are robust within a month and independent of the method used (Study 4: Stability over time and across measurement methods), which back up other results. Hence, the relevance of taxonomies of illusions is questioned. In other terms, what can we learn from detailed analyses of mechanisms underlying visual illusions when they are rather idiosyncratic?

The findings reported in this thesis fit nicely into a larger picture about common factors for vision in general, which suggests that there are similarly many factors underlying several visual paradigms. For example, individual differences in contrast sensitivity (Peterzell, 2016; Peterzell, Scheffrin, et al., 2000), hue scaling (e.g., Emery et al., 2017a, 2017b), color matching (Webster & MacLeod, 1988), stereopsis (Peterzell et al., 2017), and in the effects of priors (Tulver et al., 2019) were suggested to rely on several specific factors. All these studies clearly argue against the widespread intuition that there are only a few mechanisms behind vision, which just need to be unearthed. It rather seems that there is a plethora of idiosyncratic mechanisms, arguing against the monolithic theory of visual processing.

5.4. Analogy with the classic model of vision

In the classic model of vision, stimuli are processed in a hierarchical and feedforward manner. For example, low-level features, such as lines and orientations, are processed by neurons in the primary visual cortex (V1), while neurons in higher visual areas (e.g., V4 and IT) are sensitive to more complex features, such as shapes or faces (e.g., Hubel & Wiesel, 1959, 1962). When a complex shape is viewed, high-level neurons fire according to the responses of lower-level neurons. For example, when a square is viewed, the response of high-level neurons depends on the firing of lower-level neurons, which are sensitive to the orientation of the lines making up the squares (i.e., horizontal and vertical). While the definition of a common factor is vague, it may be that visual factors simply refer to the way visual stimuli are processed in the brain.

During my doctoral studies, I observed that the susceptibility to several variants of the Ebbinghaus illusion, which varied in shape, size, color, or texture, strongly correlated, suggesting the existence of a factor specific to the Ebbinghaus illusion (Study 1: Feature-unspecific). Similarly, I observed strong correlations between the susceptibility to different variants of the same illusion, which were presented

in different orientations (Study 1: Orientation-unspecific). Hence, it seems that factors underlying visual illusions are not hyperspecific, unlike in perceptual learning, where learning is usually feature-specific. For example, training with a vertical vernier stimulus does not improve performance for the same stimulus rotated by 90 degrees (e.g., Spang et al., 2010), contrary to the illusions where individual differences are stable when an illusion is rotated. Importantly, however, within-illusion correlations were not perfect (i.e., $r < 1.0$), indicating that individual differences may slightly vary from one illusion variant to the other. While the truly uncorrelated variability is hardly disentangled from noise (i.e., intraobserver variation, instrumental variation, or variation in the experimenter), it is not completely unlikely that there are even more than one factor underlying a visual illusion. In other words, we cannot exclude that factors are as specific as in perceptual learning. For example, a slight change in the stimulus, such as switching from one variant of the Ebbinghaus illusion to another, may result in the same set of factors being activated (i.e., the same visual pathway), except for one (or very few) factor that differs between the two processing routes (i.e., mostly at low level, where neurons are specifically tuned to a feature). Hence, visual factors may simply describe the intrinsic nature of visual processing, i.e., the more similar the neuronal firing, the stronger the correlation between two viewed stimuli.

Factors may therefore combine independently to underlie the representation of complex stimuli, as in Study 3, where the susceptibility to a merged illusion was strongly predicted from the susceptibility to the two illusions of which it was made. In contrast, two different illusions, which conceptually look similar, may still tap into very distinct factors or into factors that combine differently (e.g., to represent higher-order features such as depth), which shows up as a weak correlation between the susceptibility to the two illusions. For example, only a weak association was observed between two contrast illusions (Grzeczowski, Clarke, et al., 2017), even though they look alike from a conceptual point of view.

The correlational strength (i.e., the effect size) may give strong clues about the similarity between underlying factors (i.e., mechanisms) in the absence of measurement noise (note that this is hardly achievable, if not impossible). For instance, a strong correlation was observed between the magnitudes of a corridor variant of the Ponzo illusion (i.e., Ponzo hallway) and the Ebbinghaus illusion (Study 4: Stability across eyes; Grzeczowski, Clarke, et al., 2017). However, the correlations between the corridor variant (PZc) and other variants of the Ponzo illusion (PZ, PZw, and PZg) were weaker than between other pairs of Ponzo variants (Study 4: Stability across eyes), which suggests that the corridor variant of the Ponzo illusion is – factorially speaking – more similar to the Ebbinghaus illusion than to other variants of the Ponzo illusion. Similitudes are noticed between the two illusions, e.g., the circular shape of the target and reference, their spatial position, and the size constancy mechanism that may act behind the scene. A full and partial interocular transfer was observed when the Ponzo and Ebbinghaus illusions were tested in a dichoptic viewing condition, respectively (Song et al., 2011). In

addition, a partial interocular transfer was observed in the effect of linear cues in the dichoptic perception of a corridor illusion (Yildiz et al., 2021), suggesting that the corridor and Ebbinghaus illusions may rely on more similar mechanisms than the corridor and other variants of the Ponzo illusion.

5.5. Islands of significant correlations within an ocean of non-correlations

Hence, there seem to be some associations between specific pairs of visual illusions or paradigms, which may not be obvious at first sight, as between the Ebbinghaus and corridor variant of the Ponzo illusion. In addition, Grzeczowski, Clarke, et al. (2017) found that the Ponzo illusion magnitude correlates with cognitive disorganization. Likewise, contrast detection and visual backward masking were shown to strongly correlate in action video gamers, as previously reported in healthy young adults (da Cruz, Shaqiri, et al., 2020).

Given the large number of statistical comparisons and the absence of clear replications (for example, the association between the Ebbinghaus and corridor illusions was not significant in Schwarzkopf and colleagues, 2011), it may be that these few significant correlations are false positives. Alternatively, it may be that there are some strong singular correlations within an ocean of weak correlations. While these strong correlations may be surprising from a vision science perspective, they may be less surprising from a genetic point of view. As a metaphor, a hypothetical gene may code for a protein, which plays an important role in the visual cortex and, say, in the liver. Variability in that gene may cause significant correlations between visual and liver functions, which may appear bizarre as long as one does not expect visual and liver functions to be related.

Importantly, the aim of the studies reported here was not to make conclusions about specific comparisons, such as the relationship between the Ebbinghaus and corridor variant of the Ponzo illusion. To examine all individual comparisons, the studies were mainly underpowered. Instead, the general pattern of correlations was of interest. I specifically wondered whether there is a strong positive manifold, i.e., whether there are strong correlations in general. Hence, correlation tables were reported and dimensionality reduction techniques, such as factor analysis, were used as a way to unravel the underlying structure.

Until recently, null results, i.e., datasets that do not show clear and significant effects, were considered as not worth being published, resulting in publication bias. However, it is important to publish null results to give an accurate global picture of the field of research. For example, in clinical research, publications usually only report deficits in patients compared to healthy controls. However, patients do not always perform worse compared to healthy controls (i.e., diseases are heterogeneous), which can only be emphasized when null results are published. Here, I observed that the structure underlying

several visual paradigms was similar in healthy controls and patients with schizophrenia (Study 6). Similarly, only publishing studies, which suggest a strong common factor for vision, may result in a truncated, monolithic vision of the true visual structure.

5.6. The notion of a common factor is ill-posed

Scientists often rely on the percentage of variance explained by the first component or factor of a factor analysis to draw a conclusion about the existence – or absence – of a common factor underlying the dataset. For example, Halpern et al. (1999) reported 30% of explained variance in the tasks they tested, which was sufficient for the authors to conclude in favor of an underlying common factor. At first sight, however, it seems that results from different studies failed to agree. Hence, with a larger proportion of variance explained by the first factor compared to Halpern et al. (1999), i.e., 34%, Cappe et al. (2014) concluded that there is no strong common factor for visual perception. Note that both studies examined individual differences in the performance in basic visual tasks. Should scientists simply agree about a benchmark, i.e., a threshold in terms of proportion of explained variance, above which one can consider that there is an underlying common factor? For example, 50% of the variance was suggested to be the lower threshold for a common factor of cognitive abilities such as the *g* factor (Jensen, 1998; Lubinski, 2000).

However, such a threshold is meaningless as long as the intrinsic properties of the dataset, such as the sample size and number of measures, are not considered. To demonstrate this, I randomly simulated datasets of a certain size, i.e., with a predetermined sample size and number of measures, such that the data points for each measure were normally distributed ($M = 0, SD = 1$). The proportion of variance explained by the first eigenvalue was computed and averaged across 10,000 simulations. Results clearly show that the proportion of explained variance is a function of the sample size and number of measures (Figure 20). The smaller the sample size and the less measures, the more likely it is to obtain a large percentage of explained variance by chance. Indeed, the less data points, the more likely it is that there is a correlational pattern underlying these data points, even though they were randomly picked from a normal distribution. This is illustrated in Figure 21, where I simulated two vectors of 1,000,000 data points each from a normal distribution ($M = 0, SD = 1$) with a predetermined intercorrelation (effect size of 0.1, 0.2, 0.3, or 0.5). Samples of different sizes were randomly picked from these two vectors and correlated. The correlation coefficient r , which was averaged across 10,000 simulations, was centered around the true effect size. However, the smaller the sample size, the larger the 95% confidence interval. Hence, with a small sample size (i.e., in underpowered experiments), it is not unlikely to obtain an effect size that is very dissimilar to the true effect size. In contrast, with a larger sample size, the likelihood of a correlation which largely differs from the true effect size is weaker.

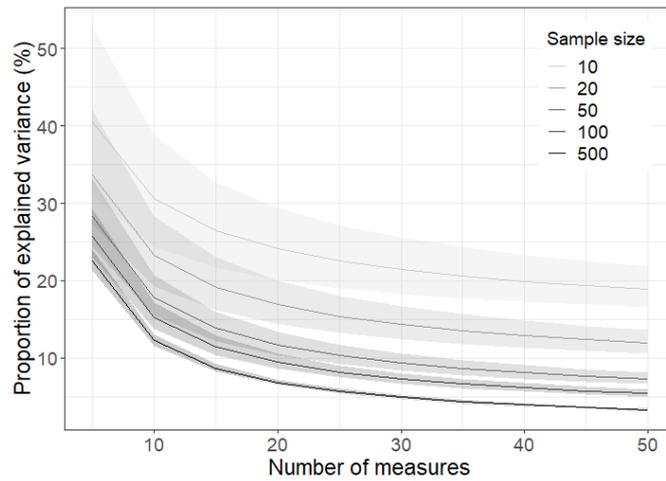


Figure 20. Proportion of variance explained by the first eigenvalue in randomly simulated datasets as a function of the number of measures and sample size. The smaller the sample size and the less measures, the more likely it is to obtain a large percentage of variance explained by the first eigenvalue just by chance. Results were averaged across 10,000 simulations (solid lines) and shaded areas represent 95% confidence intervals.

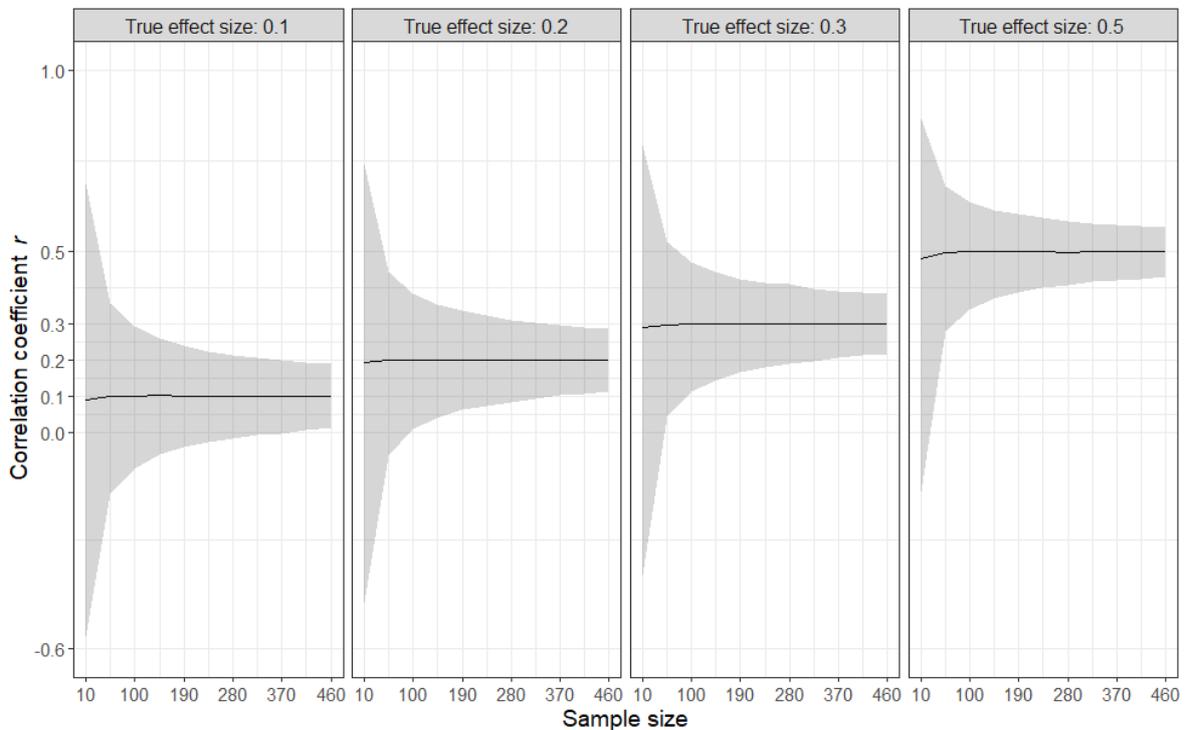


Figure 21. Underpowered experiments are likely to result in an effect size that is dissimilar to the true effect size. Here, two vectors of 1,000,000 data points each were simulated so that they correlated with a predetermined effect size (referred to as the true effect size; $r = 0.1$, $r = 0.2$, $r = 0.3$, and $r = 0.5$ are shown in different subpanels). Samples of different sizes were randomly picked from these two vectors and correlated. The resulting correlation coefficient r was averaged across 10,000 simulations (solid lines) and 95% confidence intervals are reported (shaded areas).

Hence, the notion of a common factor is not as straightforward as it may be thought. In fact, a common factor is an ill-posed problem, in the sense that the same solution, i.e., the proportion of variance explained by the first eigenvalue, may arise from different situations. For example, in Figure 20, both (1) a dataset with 14 measures and 20 subjects and (2) a dataset with 40 measures and 10 subjects resulted in a first eigenvalue explaining on average 20% of the variance in randomly but normally

distributed datasets. To summarize, there is no reason to compare the proportion of variance explained by the first component (or factor) across studies as long as they do not share the same intrinsic properties.

The lack of coherence between the results from different studies may also arise from the lack of agreement regarding the statistical methods. While a factor analysis is computed in most cases, scientists face multiple methodological choices (e.g., Costello & Osborne, 2005). First, there is a fundamental difference between a principal component analysis, in which the full variance (i.e., common and unique variances) is brought into the factor matrix, and a common factor analysis, in which only the common variance is employed in the estimation of the factors. Technically speaking, the former implies inserting unities (i.e., 1.0 values) in the diagonal of the correlation matrix, while the communalities are inserted in the latter case. Note, however, that both are techniques of dimensionality reduction, which often result in similar factor structures. Second, scientists must decide how to determine the number of factors to retain. A commonly used rule is the Kaiser rule (also known as Latent Root Criterion), which suggests to retain all eigenvalues larger than 1.0 (Kaiser, 1970). Another way to identify the optimum number of factors to extract is to look for the inflection point in the scree plot, which is often referred to as the elbow. Scientists then usually consider all factors with eigenvalues preceding the inflection point for inclusion. Unlike the Kaiser rule and scree plot inspection, a parallel analysis is based on the intrinsic properties (i.e., number of measures and sample size) of the dataset being analyzed. In a parallel analysis, simulated datasets of the same size as the original dataset are randomly generated. Eigenvalues are computed and averaged across a large number of simulations (as in Figure 20). These values are then compared to the eigenvalues extracted for the original dataset. All factors with eigenvalues larger than those from the simulated datasets are retained (Horn, 1965). Third, an orthogonal or oblique rotational procedure can be used. These and more discrepancies in the data analysis make the comparison across studies and interpretation complex.

5.7. Batteries of tasks are powerful

Scientists commonly focus on a single paradigm, which leads to impaired performances in a special population compared to a control group. For example, patients with schizophrenia and their unaffected siblings were shown to have impaired visual backward masking compared to healthy controls, which showed up both as an increased interstimulus interval (ISI) needed to reach 75% of correct responses and strongly reduced magnitudes in the electroencephalogram (da Cruz, Shaqiri, et al., 2020; Favrod et al., 2018). The authors suggested that unaffected siblings overactivate specific brain networks to partially compensate for the deficit. Hence, visual backward masking was proposed as a candidate endophenotype of schizophrenia, i.e., a trait marker that is abnormally present in

patients and, to a lesser extent, in their unaffected relatives (Gottesman & Gould, 2003). Patients with schizophrenia not only show deficits in specific visual tasks but also in cognition (e.g., Gur et al., 2007; Meyer-Lindenberg et al., 2001), electrophysiology (Price et al., 2006) and immunology (e.g., Cullen et al., 2019). In addition, more than a hundred genetic loci were associated with schizophrenia in genome-wide association studies (Pardiñas et al., 2018; Ripke et al., 2014).

In contrast, only a few studies investigated the relationship between several endophenotypes in patients with schizophrenia. For example, the Consortium on the Genetics of Schizophrenia measured 15 neurocognitive and neurophysiological tasks in a large sample of patients with schizophrenia, siblings of patients with schizophrenia, and healthy controls (Seidman et al., 2015; see also Dickinson et al., 2011). A factor analysis resulted in a very similar factor structure across groups, with 10 tasks strongly loading on five factors.

Given the high heterogeneity of schizophrenia (i.e., patients do not show abnormal patterns of performance in all tasks), a battery of tasks may be more appropriate than a single task to diagnose the disease. If a set of tasks strongly correlate with each other, it is likely that one of them is sufficiently representative of the whole domain tested. In contrast, a common factor accounting for only a small proportion of the variance in the tasks, as observed in Studies 5 to 7, indicates that a combination of tasks may produce a much better outcome prediction compared to a single task. For example, while only weak correlations were observed between electrophysiological features, combining features considerably improved the predictability of the patients' clinical outcomes compared to individual features (da Cruz, Gordillo, et al., 2020). That is the magic of combinatorics.

5.8. Limitations

5.8.1. Power

Power may be an issue when only small sample sizes were tested (e.g., Study 1: Luminance- and orientation-unspecific). Such low powered experiments do not allow to make conclusions about specific pairwise correlations, which was nevertheless not the main goal of the studies. Note, however, that similar patterns of correlations, i.e., strong within-paradigm and weaker between-paradigm correlations, were observed in all studies independently of the sample size.

5.8.2. Experimental flaws

In Studies 1 to 4, which explored the individual differences in the perception of visual illusions, a few weak correlations are likely explained by experimental flaws. For example, there were no significant correlations between the different luminance conditions of the Ebbinghaus and Delboeuf illusions in Study 1 (luminance-unspecific), while both illusions were often proposed to rely on the same

mechanism (Coren et al., 1976; Girgus et al., 1972; Roberts et al., 2005) and observed to strongly correlate (e.g., Sherman & Chouinard, 2016). Importantly, the two targets in the Ebbinghaus and Delboeuf illusions were vertically aligned in Study 1 (luminance-unspecific). Hence, participants may have performed the task by drawing imaginary vertical lines between the targets without considering the context (also note that power was low). Likewise, there were only weak associations between the Judd and Müller-Lyer illusions, which however look very similar. An explanation may be that the two segments compared in the Judd illusion were collinear while vertically aligned in the Müller-Lyer illusion. Therefore, results should always be considered with caution, since false positives and false negatives can result from such experimental flaws (making between-study comparisons even more complicated).

5.8.3. Non-perceptual factors

Importantly, components other than purely perceptual may also play a role in visual perception. For example, physiological (e.g., Pollack & Silvar, 1967; Silvar & Pollack, 1967) or cultural (e.g., Brislin & Keating, 1976; Caparos et al., 2012; de Fockert et al., 2007; Leibowitz & Pick, 1972) factors were shown to contribute to the individual differences in the perception of visual illusions. Children growing up in an urban environment were reported to experience more linear perspectives than children growing up in a rural environment, thereby enhancing their susceptibility to some illusions (Brislin, 1974; Segall et al., 1963; Stewart, 1973). Likewise, genetic variance may act as a hidden covariate, resulting in biased results. Indeed, a modest sample size often results in an unbalanced distribution of genotypes (with some being more represented than others), which may obscure the emergence of a clear factorial structure (e.g., Maksimov et al., 2013). Overall, such non-perceptual factors intermingle in a complex manner with perceptual ones, resulting in a highly dimensional space.

5.8.4. Strengths and weaknesses of the adjustment procedure

More than being playful (e.g., when testing children), the adjustment procedure as used here and previously (Grzeczowski et al., 2018; Grzeczowski, Clarke, et al., 2017) showed overall similar individual differences in the perception of visual illusions compared to a more traditional method of constant stimuli (most illusion studies used forced choice responses; see King et al., 2017). Indeed, in Study 4 (stability across measurement methods), individual differences in illusion magnitudes as measured with an adjustment procedure and a method of constant stimuli were shown to strongly correlate in three out of four conditions, suggesting that individual differences are largely stable across the two measurement methods. Similarly, Schwarzkopf and colleagues (2011) observed strong correlations in the estimation of the susceptibility to the Ebbinghaus and corridor illusions between an adjustment procedure (with 8 trials per condition) and a method of constant stimuli. Hence, it seems that an adjustment procedure is amenable to replace a method of constant stimuli when testing

healthy subjects. In contrast, the adjustment procedure was suggested to be susceptible to atypical decisional bias, e.g., in clinical populations such as patients with autism (Manning et al., 2017). Note, however, that the susceptibility to visual illusions when measured with an adjustment procedure did not significantly differ between patients with schizophrenia and healthy controls, except in the estimation of a contrast illusion (Grzeczowski et al., 2018).

Nevertheless, the adjustment procedure as used here is not free from weaknesses. For example, it may be argued that two trials of each condition are too few. Indeed, a misclick or poor attention paid to one of the trials biases the average, therefore leading to a poor estimate of the illusion magnitude. However, test-retest reliabilities were in general significant (even though not always strong).

Another weakness of the adjustment procedure is its non-speeded nature, i.e., the stimuli were shown until a response was given. It was recently suggested that perceptual bias is reduced with longer response times in orientation discrimination tasks (Dekel & Sagi, 2020a, 2020b). Similarly, the susceptibility to the Ebbinghaus illusion was suggested to decrease with slower responses (for exposure times of several seconds; Bressan & Kramer, 2013, 2021; but the opposite relationship was observed for shorter time scales; see Schmidt et al., 2016). In contrast, priming effects induced by a Ponzo illusion increased with slower responses (for SOAs ≤ 130 ms), suggesting that the Ponzo illusion is not only based on rapid feedforward processing but also relies on recurrent processing (Schmidt & Haberkamp, 2016). Hence, the variability in response times may strongly bias the results. Note, however, that the aforementioned studies used a binary response mode (e.g., two-alternative forced choice) rather than an adjustment procedure. Importantly, in Studies 1 to 3, correlations between illusion magnitudes and response times were weak and non-significant. Hence, future work may further examine how susceptibility to visual illusions varies with response time when an adjustment procedure is used.

5.8.5. Non-zero control bias in illusion magnitudes

Non-zero effect size in illusion magnitude may arise even in the absence of context, i.e., there may be a control bias. For example, in Study 1 (feature-unspecific), the control condition (i.e., a condition with no flankers) was expected to show a null effect (i.e., an illusion magnitude that does not significantly differ from zero), which however was not the case. In all other conditions, participants overadjusted the size of the target compared to the reference, because the former was always surrounded by larger flankers compared to the reference. Therefore, it may have induced a small but significant bias in the control condition as well. Partial correlations were computed to control for this bias. However, not all experiments included control conditions. Correlation coefficients must therefore be considered with caution. Note, however, that there should be no (or at least weaker) control bias when the target and reference were randomized across trials. For instance, in Study 4 (stability across measurement

methods), the control conditions led to non-significant effects, likely because there were two reference-dependent conditions for each illusion, i.e., one element was in turn the target or the reference.

5.8.6. Cross-sectional design

The cross-sectional design of the studies exploring the structure of vision in special populations (Studies 5 to 7) does not allow to draw conclusions about how perception changes with age (Study 5), when developing psychopathologies (Study 6), or when becoming a video game expert (Study 7). While longitudinal studies (or an intervention study in the case of Study 7) would allow a closer examination of the changes and the direction of the effects, the results of the present investigations only suggest that there is no strong common factor for vision in these special populations. For example, a patient with schizophrenia can show a deficit in a specific visual function but not in others. Similarly, it seems that becoming an expert in action video games is not related with a consistent enhancement of all visual skills. In contrast, optical or neurological pathologies can lead to reliably poor performances in several visual paradigms, as in the case of a blind person.

5.8.7. Tasks other than visual

Tasks tested in Studies 5 to 7 were not all purely visual. While the main goal of these studies was to span a wide range of visual functions (i.e., from basic to more complex), some of these tasks also tap into more cognitive aspects, such as inhibition and attention. In Study 5, three cognitive tasks (verbal fluency, digit span, and Wisconsin Card Sorting Test) were also included in the battery to assess a potential relationship between cognitive and perceptual functions. Unlike previously reported (e.g., Lindenberger & Ghisletta, 2009), only weak correlations were observed between visual and cognitive paradigms, suggesting that differences in the three cognitive functions tested are not strongly related to perception, and inversely. However, more than three cognitive tasks are necessary to gather evidence in favor or against a common factor of cognition related to perception.

5.9. Future work

There are still several open questions that can be answered in the near future regarding the unknowns of visual perception. For example, the mechanism(s) underlying each visual illusion may be unraveled. Likewise, the significant associations observed in Studies 1 to 4 may be further investigated to determine whether they are false positives, arise from experimental weaknesses, or are true effects. In addition, the genetic component in the perception of visual illusions may be further examined. For example, the pattern of individual differences may be compared between twins and siblings. Likewise, cultural differences may be systematically investigated by testing visual illusions around the world (controlling for other confounding variables, such as age, gender, and education).

Importantly, rethinking the statistical methods may help unravelling the structure underlying vision. More specifically, I suggest that the methodology must be homogenized so that comparisons across studies are possible. For such a purpose, the strength of the factors must be quantified regarding to the intrinsic properties of the datasets. If individual differences were perfectly reliable across visual tasks, i.e., in case of an absolute common factor for vision, an individual ranking first in one visual task should rank first in all other visual tasks as well. Numerical ranks could therefore be compared to randomly simulated ranks to determine whether a pattern of “good” or “bad” individuals emerges, i.e., is an individual on average better or worse than when attributed with random ranks? Likewise, a meta-analytical approach may be used to summarize previous results and to examine whether conclusions generalize across different research groups.

5.10. Conclusions

While there are stable individual differences in the perception of variants of the same illusion, different illusions are rather uncorrelated, which suggests that there is no strong common factor underlying visual illusions. In addition, the results presented in this thesis provide evidence in favor of a complex, multifactorial structure underlying visual perception in control groups as well as healthy elderly, patients with schizophrenia, and action video game players. Contrary to the wide held belief that similar tasks rely on similar neural mechanisms, performance did not generalize across similar visual tasks. Hence, an individual’s performance in one visual task is not as informative for other visual tasks as it may be thought. Rather than a unique and strong common factor, vision is likely composed of a multitude of specific and independent abilities that come together to make each one of us have a unique view of the world.

6. References

- Andermane, N., Bosten, J. M., Seth, A. K., & Ward, J. (2019). Individual differences in change blindness are predicted by the strength and stability of visual representations. *Neuroscience of Consciousness*, 5(1), 1–12. <https://doi.org/10.1093/nc/niy010>
- Andreasen, N. C. (1984a). *Scale for the Assessment of Negative Symptoms (SANS)*. University of Iowa, Iowa City.
- Andreasen, N. C. (1984b). *Scale for the Assessment of Positive Symptoms (SAPS)*. University of Iowa, Iowa City.
- Axelrod, V., Schwarzkopf, D. S., Gilaie-Dotan, S., & Rees, G. (2017). Perceptual similarity and the neural correlates of geometrical illusions in human brain structure. *Scientific Reports*, 7(1), 1–16. <https://doi.org/10.1038/srep39968>
- Ball, K., & Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science*, 218(4573), 697–698. <https://doi.org/10.1126/science.7134968>
- Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research*, 27(6), 953–965. [https://doi.org/10.1016/0042-6989\(87\)90011-3](https://doi.org/10.1016/0042-6989(87)90011-3)
- Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging*, 12(1), 12–21. <https://doi.org/10.1037/0882-7974.12.1.12>
- Bargary, G., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). Individual differences in human eye movements: An oculomotor signature? *Vision Research*, 141, 157–169. <https://doi.org/10.1016/j.visres.2017.03.001>
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*, 31(1), 5–17. <https://doi.org/10.1023/A:1005653411471>
- Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological Reports*, 19(1), 3–11. <https://doi.org/10.2466/pr0.1966.19.1.3>
- Berardi, N., & Fiorentini, A. (1987). Interhemispheric transfer of visual information in humans: spatial characteristics. *The Journal of Physiology*, 384(1), 633–647.

<https://doi.org/10.1113/jphysiol.1987.sp016474>

- Bertamini, M., Cretienoud, A. F., & Herzog, M. H. (2019). Exploring the Extent in the Visual Field of the Honeycomb and Extinction Illusions. *I-Perception*, *10*(4), 1–19. <https://doi.org/10.1177/2041669519854784>
- Bertamini, M., Herzog, M. H., & Bruno, N. (2016). The Honeycomb illusion: Uniform textures not perceived as such. *I-Perception*, *7*(4), 1–15. <https://doi.org/10.1177/2041669516660727>
- Binet, A. (1895). La mesure des illusions visuelles chez les enfants. *Revue Philosophique de La France et de l'Étranger*, *40*, 11–25.
- Bosten, J. M., Goodbourn, P. T., Bargary, G., Verhallen, R., Lawrance-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). An exploratory factor analysis of visual performance in a large population. *Vision Research*, *141*, 303–316. <https://doi.org/10.1016/j.visres.2017.02.005>
- Bosten, J. M., & Mollon, J. D. (2010). Is there a general trait of susceptibility to simultaneous contrast? *Vision Research*, *50*(17), 1656–1664. <https://doi.org/10.1016/j.visres.2010.05.012>
- Bosten, J. M., Mollon, J. D., Peterzell, D. H., & Webster, M. A. (2017). Individual differences as a window into the structure and function of the visual system. *Vision Research*, *141*, 1–3. <https://doi.org/10.1016/j.visres.2017.11.003>
- Braddick, O. (1972). The psychology of visual illusion and some illusions of visual psychology. *Perception*, *1*, 239–241.
- Braddick, O. (2018). Illusion Research: An Infantile Disorder? *Perception*, *47*(8), 805–806. <https://doi.org/10.1177/0301006618774658>
- Braff, D. L., Freedman, R., Schork, N. J., & Gottesman, I. I. (2007). Deconstructing schizophrenia: An overview of the use of endophenotypes in order to understand a complex disorder. *Schizophrenia Bulletin*, *33*(1), 21–32. <https://doi.org/10.1093/schbul/sbl049>
- Brascamp, J. W., Becker, M. W., & Hambrick, D. Z. (2018). Revisiting individual differences in the time course of binocular rivalry. *Journal of Vision*, *18*(7), 1–20. <https://doi.org/10.1167/18.7.3>
- Brascamp, J. W., Qian, C. S., Hambrick, D. Z., & Becker, M. W. (2019). Individual differences point to two separate processes involved in the resolution of binocular rivalry. *Journal of Vision*, *19*(12), 1–17. <https://doi.org/10.1167/19.12.15>
- Bressan, P., & Kramer, P. (2013). The relation between cognitive-perceptual schizotypal traits and the

- Ebbinghaus size-illusion is mediated by judgment time. *Frontiers in Psychology*, 4:343, 1–8. <https://doi.org/10.3389/fpsyg.2013.00343>
- Bressan, P., & Kramer, P. (2021). Most findings obtained with untimed visual illusions are confounded. *PsyArXiv*, 1–12. <https://doi.org/10.31234/OSF.IO/PZHRD>
- Brislin, R. W. (1974). The Ponzo illusion: Additional Cues, Age, Orientation, and Culture. *Journal of Cross-Cultural Psychology*, 5(2), 139–161. <https://doi.org/10.1177/002202217400500201>
- Brislin, R. W., & Keating, C. F. (1976). Cultural Differences in the Perception of a Three-Dimensional Ponzo Illusion. *Journal of Cross-Cultural Psychology*, 7(4), 397–412. <https://doi.org/10.1177/002202217674002>
- Burghoorn, F., Dingemanse, M., van Lier, R., & van Leeuwen, T. M. (2020). The Relation Between Autistic Traits, the Degree of Synaesthesia, and Local/Global Visual Perception. *Journal of Autism and Developmental Disorders*, 50(1), 12–29. <https://doi.org/10.1007/s10803-019-04222-7>
- Cao, T., Wang, L., Sun, Z., Engel, S. A., & He, S. (2018). The independent and shared mechanisms of intrinsic brain dynamics: insights from bistable perception. *Frontiers in Psychology*, 9:589, 1–11. <https://doi.org/10.3389/fpsyg.2018.00589>
- Caparos, S., Ahmed, L., Bremner, A. J., de Fockert, J. W., Linnell, K. J., & Davidoff, J. (2012). Exposure to an urban environment alters the local bias of a remote culture. *Cognition*, 122(1), 80–85. <https://doi.org/10.1016/j.cognition.2011.08.013>
- Cappe, C., Clarke, A., Mohr, C., & Herzog, M. H. (2014). Is there a common factor for vision? *Journal of Vision*, 14(8), 1–11. <https://doi.org/10.1167/14.8.4>
- Ćepulić, D. B., Wilhelm, O., Sommer, W., & Hildebrandt, A. (2018). All categories are equal, but some categories are more equal than others: The psychometric structure of object and face cognition. *Journal of Experimental Psychology: Learning Memory and Cognition*, 44(8), 1254–1268. <https://doi.org/10.1037/xlm0000511>
- Chamberlain, R., Van der Hallen, R., Huygelier, H., Van de Cruys, S., & Wagemans, J. (2017). Local-global processing bias is not a unitary individual difference in visual processing. *Vision Research*, 141, 247–257. <https://doi.org/10.1016/j.visres.2017.01.008>
- Chen, L., Wu, B., Qiao, C., & Liu, D. Q. (2020). Resting EEG in alpha band predicts individual differences in visual size perception. *Brain and Cognition*, 145, 1–7. <https://doi.org/10.1016/j.bandc.2020.105625>

- Chkonia, E., Roinishvili, M., Makhatadze, N., Tsverava, L., Stroux, A., Neumann, K., Herzog, M. H., & Brand, A. (2010). The Shine-Through Masking Paradigm Is a Potential Endophenotype of Schizophrenia. *PLoS ONE*, *5*(12), 1–7. <https://doi.org/10.1371/journal.pone.0014268>
- Chouinard, P. A., Unwin, K. L., Landry, O., & Sperandio, I. (2016). Susceptibility to Optical Illusions Varies as a Function of the Autism-Spectrum Quotient but not in Ways Predicted by Local–Global Biases. *Journal of Autism and Developmental Disorders*, *46*(6), 2224–2239. <https://doi.org/10.1007/s10803-016-2753-1>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.
- Coren, S., & Enns, J. T. (1993). Size contrast as a function of conceptual similarity between test and inducers. *Perception & Psychophysics*, *54*(5), 579–588. <https://doi.org/10.3758/BF03211782>
- Coren, S., & Girgus, J. S. (1972). Illusion decrement in intersecting line figures. *Psychonomic Science*, *26*(2), 108–110. <https://doi.org/10.3758/BF03335451>
- Coren, S., Girgus, J. S., Erlichman, H., & Hakstian, A. R. (1976). An empirical taxonomy of visual illusions. *Perception & Psychophysics*, *20*(2), 129–137. <https://doi.org/10.3758/BF03199444>
- Coren, S., & Miller, J. (1974). Size contrast as a function of figural similarity. *Perception & Psychophysics*, *16*(2), 355–357. <https://doi.org/10.3758/BF03203955>
- Costello, A. B., & Osborne, J. W. (2005). Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research & Education*, *10*(7), 86–99. <https://doi.org/10.1.1.110.9154>
- Cretenoud, A. F., Barakat, A., Milliet, A., Choung, O.-H., Bertamini, M., Constantin, C., & Herzog, M. H. (submitted). How do visual skills relate to action video game performance?
- Cretenoud, A. F., Francis, G., & Herzog, M. H. (2020). When illusions merge. *Journal of Vision*, *20*(8), 1–15. <https://doi.org/10.1167/JOV.20.8.12>
- Cretenoud, A. F., Grzeczowski, L., Bertamini, M., & Herzog, M. H. (2020). Individual differences in the Müller-Lyer and Ponzo illusions are stable across different contexts. *Journal of Vision*, *20*(6), 1–14. <https://doi.org/10.1167/JOV.20.6.4>
- Cretenoud, A. F., Grzeczowski, L., Kunchulia, M., & Herzog, M. H. (submitted). Individual differences in the perception of visual illusions are stable across eyes, time, and measurement methods.
- Cretenoud, A. F., Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019).

- Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, 19(14), 1–21. <https://doi.org/10.1167/19.14.12>
- Cullen, A. E., Holmes, S., Pollak, T. A., Blackman, G., Joyce, D. W., Kempton, M. J., Murray, R. M., McGuire, P., & Mondelli, V. (2019). Associations Between Non-neurological Autoimmune Disorders and Psychosis: A Meta-analysis. *Biological Psychiatry*, 85(1), 35–48. <https://doi.org/10.1016/j.biopsych.2018.06.016>
- da Cruz, J. R., Favrod, O., Roinishvili, M., Chkonia, E., Brand, A., Mohr, C., Figueiredo, P., & Herzog, M. H. (2020). EEG microstates are a candidate endophenotype for schizophrenia. *Nature Communications*, 11(1), 1–11. <https://doi.org/10.1038/s41467-020-16914-1>
- da Cruz, J. R., Gordillo, D., Chkonia, E., Lin, W. H., Favrod, O., Brand, A., Figueiredo, P., Roinishvili, M., & Herzog, M. H. (2020). The EEG multiverse of schizophrenia. *MedRxiv*, 1–16. <https://doi.org/10.1101/2020.12.21.20248665>
- da Cruz, J. R., Shaqiri, A., Roinishvili, M., Favrod, O., Chkonia, E., Brand, A., Figueiredo, P., & Herzog, M. H. (2020). Neural Compensation Mechanisms of Siblings of Schizophrenia Patients as Revealed by High-Density EEG. *Schizophrenia Bulletin*, 46(4), 1009–1018. <https://doi.org/10.1093/schbul/sbz133>
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage*, 9(2), 179–194. <https://doi.org/10.1006/nimg.1998.0395>
- de Fockert, J., Davidoff, J., Fagot, J., Parron, C., & Goldstein, J. (2007). More Accurate Size Contrast Judgments in the Ebbinghaus Illusion by a Remote Culture. *Journal of Experimental Psychology: Human Perception and Performance*, 33(3), 738–742. <https://doi.org/10.1037/0096-1523.33.3.738>
- Dekel, R., & Sagi, D. (2020a). Interaction of contexts in context-dependent orientation estimation. *Vision Research*, 169, 58–72. <https://doi.org/10.1016/j.visres.2020.02.006>
- Dekel, R., & Sagi, D. (2020b). A decision-time account of individual variability in context-dependent orientation estimation. *Vision Research*, 177, 20–31. <https://doi.org/10.1016/j.visres.2020.08.002>
- Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *The Journal of Physiology*, 357(1), 241–265. <https://doi.org/10.1113/jphysiol.1984.sp015499>

- Dickinson, D., Goldberg, T. E., Gold, J. M., Elvevg, B., & Weinberger, D. R. (2011). Cognitive factor structure and invariance in people with schizophrenia, their unaffected siblings, and controls. *Schizophrenia Bulletin*, *37*(6), 1157–1167. <https://doi.org/10.1093/schbul/sbq018>
- Dobkins, K. R., Gunther, K. L., & Peterzell, D. H. (2000). What covariance mechanisms underlie green/red equiluminance, luminance contrast sensitivity and chromatic (green/red) contrast sensitivity? *Vision Research*, *40*(6), 613–628. [https://doi.org/10.1016/S0042-6989\(99\)00211-4](https://doi.org/10.1016/S0042-6989(99)00211-4)
- Doherty, M. J., Campbell, N. M., Tsuji, H., & Phillips, W. A. (2010). The Ebbinghaus illusion deceives adults but not young children. *Developmental Science*, *13*(5), 714–721. <https://doi.org/10.1111/j.1467-7687.2009.00931.x>
- Dougherty, R. F., Koch, V. M., Brewer, A. A., Fischer, B., Modersitzki, J., & Wandell, B. A. (2003). Visual field representations and locations of visual areas v1/2/3 in human visual cortex. *Journal of Vision*, *3*(10), 586–598. <https://doi.org/10.1167/3.10.1>
- Dresslar, F. B. (1894). Studies in the Psychology of Touch. *The American Journal of Psychology*, *6*(3), 313–368. <https://doi.org/10.2307/1411644>
- Emery, K. J., Volbrecht, V. J., Peterzell, D. H., & Webster, M. A. (2017a). Variations in normal color vision. VI. Factors underlying individual differences in hue scaling and their implications for models of color appearance. *Vision Research*, *141*, 51–65. <https://doi.org/10.1016/j.visres.2016.12.006>
- Emery, K. J., Volbrecht, V. J., Peterzell, D. H., & Webster, M. A. (2017b). Variations in normal color vision. VII. Relationships between color naming and hue scaling. *Vision Research*, *141*, 66–75. <https://doi.org/10.1016/j.visres.2016.12.007>
- Fahle, M., & Morgan, M. (1996). No transfer of perceptual learning between similar stimuli in the same retinal position. *Current Biology*, *6*(3), 292–297. [https://doi.org/10.1016/S0960-9822\(02\)00479-7](https://doi.org/10.1016/S0960-9822(02)00479-7)
- Faivre, N., Filevich, E., Solovey, G., Kühn, S., & Blanke, O. (2018). Behavioural, modeling, and electrophysiological evidence for supramodality in human metacognition. *Journal of Neuroscience*, *38*(2), 263–277. <https://doi.org/10.1523/JNEUROSCI.0322-17.2017>
- Favrod, O., Roinishvili, M., da Cruz, J. R., Brand, A., Okruashvili, M., Gamkrelidze, T., Figueiredo, P., Herzog, M. H., Chkonia, E., & Shaqiri, A. (2018). Electrophysiological correlates of visual backward masking in patients with first episode psychosis. *Psychiatry Research - Neuroimaging*, *282*, 64–72. <https://doi.org/10.1016/j.psychresns.2018.10.008>

- Fisher, R. A. (1992). Statistical Methods for Research Workers. In *Breakthroughs in Statistics*. (pp. 66–70). Springer, New York, NY. https://doi.org/10.1007/978-1-4612-4380-9_6
- Frenzel, H., Bohlender, J., Pinsker, K., Wohlleben, B., Tank, J., Lechner, S. G., Schiska, D., Jaijo, T., Rüschemdorf, F., Saar, K., Jordan, J., Millán, J. M., Gross, M., & Lewin, G. R. (2012). A genetic basis for mechanosensory traits in humans. *PLoS Biology*, *10*(5), 1–15. <https://doi.org/10.1371/journal.pbio.1001318>
- Ghisletta, P., & Lindenberger, U. (2005). Exploring structural dynamics within and between sensory and intellectual functioning in old and very old age: Longitudinal evidence from the Berlin Aging Study. *Intelligence*, *33*(6), 555–587. <https://doi.org/10.1016/J.INTELL.2005.07.002>
- Gignac, G. E., & Kretschmar, A. (2017). Evaluating dimensional distinctness with correlated-factor models: Limitations and suggestions. *Intelligence*, *62*, 138–147. <https://doi.org/10.1016/j.intell.2017.04.001>
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78. <https://doi.org/10.1016/j.paid.2016.06.069>
- Girgus, J. S., Coren, S., & Agdern, M. (1972). The interrelationship between the Ebbinghaus and Delboeuf illusions. *Journal of Experimental Psychology*, *95*(2), 453–455. <https://doi.org/10.1037/h0033606>
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25. [https://doi.org/10.1016/0166-2236\(92\)90344-8](https://doi.org/10.1016/0166-2236(92)90344-8)
- Goodbourn, P. T., Bosten, J. M., Hogg, R. E., Bargary, G., Lawrance-Owen, A. J., & Mollon, J. D. (2012). Do different “magnocellular tasks” probe the same neural substrate? *Proceedings of the Royal Society B: Biological Sciences*, *279*(1745), 4263–4271. <https://doi.org/10.1098/rspb.2012.1430>
- Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *American Journal of Psychiatry*, *160*(4), 636–645. <https://doi.org/10.1176/appi.ajp.160.4.636>
- Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature*, *199*, 678–680.
- Gregory, R. L. (1964). Illusory Perception as a Constancy Phenomenon. *Nature*, *204*, 302–303.
- Grzeczowski, L., Clarke, A. M., Francis, G., Mast, F. W., & Herzog, M. H. (2017). About individual

- differences in vision. *Vision Research*, *141*, 282–292.
<https://doi.org/10.1016/j.visres.2016.10.006>
- Grzeczowski, L., Cretenoud, A. F., Herzog, M. H., & Mast, F. W. (2017). Perceptual learning is specific beyond vision and decision making. *Journal of Vision*, *17*(6), 1–11. <https://doi.org/10.1167/17.6.6>
- Grzeczowski, L., Cretenoud, A. F., Mast, F. W., & Herzog, M. H. (2019). Motor response specificity in perceptual learning and its release by double training. *Journal of Vision*, *19*(6), 1–14. <https://doi.org/10.1167/19.6.4>
- Grzeczowski, L., Roinishvili, M., Chkonia, E., Brand, A., Mast, F. W., Herzog, M. H., & Shaqiri, A. (2018). Is the perception of illusions abnormal in schizophrenia? *Psychiatry Research*, *270*, 929–939. <https://doi.org/10.1016/j.psychres.2018.10.063>
- Gur, R. E., Calkins, M. E., Gur, R. C., Horan, W. P., Nuechterlein, K. H., Seidman, L. J., & Stone, W. S. (2007). The Consortium on the Genetics of Schizophrenia: Neurocognitive Endophenotypes. *Schizophrenia Bulletin*, *33*(1), 49–68. <https://doi.org/10.1093/schbul/sbl055>
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2018). *Multivariate data analysis* (8th ed.). Cengage Learning EMEA.
- Halpern, S. D., Andrews, T. J., & Purves, D. (1999). Interindividual variation in human visual performance. *Journal of Cognitive Neuroscience*, *11*(5), 521–534. <https://doi.org/10.1162/089892999563580>
- Hamburger, K., Hansen, T., & Gegenfurtner, K. R. (2007). Geometric-optical illusions at isoluminance. *Vision Research*, *47*(26), 3276–3285. <https://doi.org/10.1016/j.visres.2007.09.004>
- Herzog, M. H., Cretenoud, A. F., & Grzeczowski, L. (2017). What is new in perceptual learning? *Journal of Vision*, *17*(1), 1–4. <https://doi.org/10.1167/17.1.23>
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. <https://doi.org/10.1007/BF02289447>
- Hu, L.-T., & Bentler, P. M. (1998). Fit Indices in Covariance Structure Modeling: Sensitivity to Underparameterized Model Misspecification. *Psychological Methods*, *3*(4), 424–453. <https://doi.org/10.1037/1082-989X.3.4.424>
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55.

<https://doi.org/10.1080/10705519909540118>

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3), 574–591. <https://doi.org/10.1113/jphysiol.1959.sp006308>

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1), 106–154. <https://doi.org/10.1113/jphysiol.1962.sp006837>

Iglewicz, B., & Hoaglin, D. (1993). Volume 16: how to detect and handle outliers. In *The ASQC basic references in quality control: statistical techniques*. Asq Press.

Jaeger, T., & Klahs, K. (2015). The Ebbinghaus Illusion: New Contextual Effects and Theoretical Considerations. *Perceptual and Motor Skills*, 120(1), 177–182. <https://doi.org/10.2466/24.27.PMS.120v13x4>

Jastrzębowska, M. A. (2020). *Bayesian modeling of brain function and behavior* [EPFL, Lausanne, Switzerland]. <https://doi.org/10.5075/epfl-thesis-7388>

Jeffreys, H. (1961). *Theory of probability*. Oxford University Press.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Praeger.

Judd, C. H. (1902). Practice and its effects on the perception of illusions. *Psychological Review*, 9(1), 27–39. <https://doi.org/10.1037/h0073071>

Kaiser, H. F. (1970). A Second-Generation Little Jiffy. *Psychometrika*, 35, 401–415.

Kantrowitz, J. T., Butler, P. D., Schecter, I., Silipo, G., & Javitt, D. C. (2009). Seeing the World Dimly : The Impact of Early Visual Deficits on Visual Experience in Schizophrenia. *Schizophrenia*, 35(6), 1085–1094. <https://doi.org/10.1093/schbul/sbp100>

Kiely, K., & Anstey, K. (2017). Common Cause Theory in Aging. In *Encyclopedia of Geropsychology* (pp. 559–569). Springer Science + Business Media.

King, D. J., Hodgekins, J., Chouinard, P. A., Chouinard, V. A., & Sperandio, I. (2017). A review of abnormalities in the perception of visual illusions in schizophrenia. *Psychonomic Bulletin and Review*, 24(3), 734–751. <https://doi.org/10.3758/s13423-016-1168-5>

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155–163.

<https://doi.org/10.1016/j.jcm.2016.02.012>

Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.

Leibowitz, H. W., Brislin, R., Perlmutter, L., & Hennessy, R. (1969). Ponzo perspective illusion as a manifestation of space perception. *Science*, *166*(3909), 1174–1176. <https://doi.org/10.1126/science.166.3909.1174>

Leibowitz, H. W., & Heisel, M. A. (1958). L'évolution de l'illusion de Ponzo en fonction de l'âge. *Archives de Psychologie*, *36*, 328–331.

Leibowitz, H. W., & Pick, H. A. (1972). Cross-cultural and educational aspects of the Ponzo perspective illusion. *Perception & Psychophysics*, *12*(5), 430–432. <https://doi.org/10.3758/BF03205856>

Liljequist, D., Elfving, B., & Skavberg Roaldsen, K. (2019). Intraclass correlation – A discussion and demonstration of basic features. *PLOS ONE*, *14*(7), 1–35. <https://doi.org/10.1371/journal.pone.0219854>

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging*, *9*(3), 339–355. <https://doi.org/10.1037/0882-7974.9.3.339>

Lindenberger, U., & Ghisletta, P. (2009). Cognitive and Sensory Declines in Old Age: Gauging the Evidence for a Common Cause. *Psychology and Aging*, *24*(1), 1–16. <https://doi.org/10.1037/a0014986>

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, *240*(4853), 740–749. <https://doi.org/10.1126/science.3283936>

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, *7*(11), 3416–3468. <https://doi.org/10.1523/JNEUROSCI.07-11-03416.1987>

Lubinski, D. (2000). Scientific and social significance of assessing individual differences : “Sinking Shafts at a Few Critical Points.” *Annual Review of Psychology*, *51*(1), 405–444. <https://doi.org/10.1146/annurev.psych.51.1.405>

Mackintosh, N. J. (2011). *IQ and Human Intelligence*. Oxford University Press.

Maksimov, M., Vaht, M., Harro, J., & Bachmann, T. (2013). Can Common Functional Gene Variants

- Affect Visual Discrimination in Metacontrast Masking? *PLoS ONE*, 8(1), 1–7. <https://doi.org/10.1371/journal.pone.0055287>
- Manning, C., Morgan, M. J., Allen, C. T. W., & Pellicano, E. (2017). Susceptibility to Ebbinghaus and Müller-Lyer illusions in autistic children: a comparison of three different methods. *Molecular Autism*, 8, 1–18. <https://doi.org/10.1186/s13229-017-0127-y>
- Massaro, D. W., & Anderson, N. H. (1971). Judgmental model of the Ebbinghaus illusion. *Journal of Experimental Psychology*, 89(1), 147–151. <https://doi.org/10.1037/h0031158>
- Meyer-Lindenberg, A., Polin, J. B., Kohn, P. D., Holt, J. L., Egan, M. F., Weinberger, D. R., & Berman, K. F. (2001). Evidence for abnormal cortical functional connectivity during working memory in schizophrenia. *American Journal of Psychiatry*, 158(11), 1809–1817. <https://doi.org/10.1176/appi.ajp.158.11.1809>
- Milne, E., & Szczerbinski, M. (2009). Global and local perceptual style, field-independence, and central coherence: An attempt at concept validation. *Advances in Cognitive Psychology*, 5, 1–26. <https://doi.org/10.2478/v10053-008-0062-8>
- Mollon, J. D., Bosten, J. M., Peterzell, D. H., & Webster, M. A. (2017). Individual differences in visual science: What can be learned and what is good experimental practice? *Vision Research*, 141, 4–15. <https://doi.org/10.1016/j.visres.2017.11.001>
- Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, 36(1), 27–43. <https://doi.org/10.1007/s10803-005-0040-7>
- Moutsiana, C., De Haas, B., Papageorgiou, A., Van Dijk, J. A., Balraj, A., Greenwood, J. A., & Schwarzkopf, D. S. (2016). Cortical idiosyncrasies predict the perception of object size. *Nature Communications*, 7, 1–12. <https://doi.org/10.1038/ncomms12110>
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Ninio, J. (2014). Geometrical illusions are not always where you think they are: a review of some classical and less classical illusions, and ways to describe them. *Frontiers in Human Neuroscience*, 8:856, 1–21. <https://doi.org/10.3389/fnhum.2014.00856>

- Osborne, J. W. (2002). Effect Sizes and the Disattenuation of Correlation and Regression Coefficients: Lessons from Educational Psychology Coefficients. *Practical Assessment, Research, and Evaluation*, 8(11), 1–8. <https://doi.org/10.7275/0K9H-TQ64>
- Pardiñas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S. E., Bishop, S., Cameron, D., Hamshere, M. L., Han, J., Hubbard, L., Lynham, A., Mantripragada, K., Rees, E., MacCabe, J. H., McCarroll, S. A., Baune, B. T., Breen, G., ... Walters, J. T. R. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics*, 50(3), 381–389. <https://doi.org/10.1038/s41588-018-0059-2>
- Peterzell, D. H. (2016). Discovering Sensory Processes Using Individual Differences : A Review and Factor Analytic Manifesto. *Electronic Imaging*, 2016(16), 1–11. <https://doi.org/10.2352/ISSN.2470-1173.2016.16HVEI-112>
- Peterzell, D. H., Chang, S. K., & Teller, D. Y. (2000). Spatial frequency tuned covariance channels for red-green and luminance- modulated gratings: Psychophysical data from human infants. *Vision Research*, 40(4), 431–444. [https://doi.org/10.1016/S0042-6989\(99\)00188-1](https://doi.org/10.1016/S0042-6989(99)00188-1)
- Peterzell, D. H., Scheffrin, B. E., Tregear, S. J., & Werner, J. S. (2000). Spatial frequency tuned covariance channels underlying scotopic contrast sensitivity. In *Vision Science and its Applications: OSA Technical Digest*. Optical Society of America.
- Peterzell, D. H., Serrano-Pedraza, I., Widdall, M., & Read, J. C. A. (2017). Thresholds for sine-wave corrugations defined by binocular disparity in random dot stereograms: Factor analysis of individual differences reveals two stereoscopic mechanisms tuned for spatial frequency. *Vision Research*, 141, 127–135. <https://doi.org/10.1016/j.visres.2017.11.002>
- Peterzell, D. H., & Teller, D. Y. (1996). Individual differences in contrast sensitivity functions: The lowest spatial frequency channels. *Vision Research*, 36(19), 3077–3085. [https://doi.org/10.1016/0042-6989\(96\)00061-2](https://doi.org/10.1016/0042-6989(96)00061-2)
- Phillips, W. A., Chapman, K. L. S., & Daniel Berry, P. (2004). Size perception is less context-sensitive in males. *Perception*, 33(1), 79–86. <https://doi.org/10.1068/p5110>
- Piaget, J. (1961). *Les Mécanismes perceptifs, Modèles probabilistes, Analyse génétique, Relations avec l'intelligence* (2nd ed.). Presses universitaires de France.
- Piaget, J. (1963). Le développement des perceptions en fonction de l'âge. In *Traité de psychologie expérimentale. Vol. VI, La Perception* (pp. 1–162). Presses universitaires de France.

- Pollack, R. H. (1964). Simultaneous and successive presentation of elements of the Mueller-Lyer figure and chronological age. *Perceptual and Motor Skills*, *19*(1), 303–310. <https://doi.org/10.2466/pms.1964.19.1.303>
- Pollack, R. H., & Silvar, S. D. (1967). Magnitude of the Mueller-Lyer illusion in children as a function of pigmentation of the Fundus oculi. *Psychonomic Science*, *8*(2), 83–84. <https://doi.org/10.3758/BF03330678>
- Pooresmaeili, A., Arrighi, R., Biagi, L., & Morrone, M. C. (2013). Blood oxygen level-dependent activation of the primary visual cortex predicts size adaptation illusion. *Journal of Neuroscience*, *33*(40), 15999–16008. <https://doi.org/10.1523/JNEUROSCI.1770-13.2013>
- Preacher, K. J., Zhang, G., Kim, C., & Mels, G. (2013). Choosing the Optimal Number of Factors in Exploratory Factor Analysis: A Model Selection Perspective. *Multivariate Behavioral Research*, *48*(1), 28–56. <https://doi.org/10.1080/00273171.2012.710386>
- Predebon, J. (2006). Decrement of the Müller-Lyer and Poggendorff illusions: The effects of inspection and practice. *Psychological Research*, *70*(5), 384–394. <https://doi.org/10.1007/s00426-005-0229-6>
- Price, G. W., Michie, P. T., Johnston, J., Innes-Brown, H., Kent, A., Clissa, P., & Jablensky, A. V. (2006). A Multivariate Electrophysiological Endophenotype, from a Unitary Cohort, Shows Greater Research Utility than Any Single Feature in the Western Australian Family Study of Schizophrenia. *Biological Psychiatry*, *60*(1), 1–10. <https://doi.org/10.1016/j.biopsych.2005.09.010>
- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/>
- Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning Memory and Cognition*, *44*(4), 501–526. <https://doi.org/10.1037/xlm0000450>
- Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K. H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., Collier, D. A., Huang, H., Pers, T. H., Agartz, I., Agerbo, E., Albus, M., Alexander, M., Amin, F., Bacanu, S. A., Begemann, M., Belliveau, R. A., ... O'Donovan, M. C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*(7510), 421–427. <https://doi.org/10.1038/nature13595>
- Roberts, B., Harris, M. G., & Yates, T. A. (2005). The roles of inducer size and distance in the Ebbinghaus illusion (Titchener circles). *Perception*, *34*(7), 847–856. <https://doi.org/10.1068/p5273>

- Robinson, J. O. (1968). Retinal inhibition in visual distortion. *British Journal of Psychology*, *59*(1), 29–36. <https://doi.org/10.1111/j.2044-8295.1968.tb01113.x>
- Roff, M. (1953). A Factorial Study of Tests in the Perceptual Area. *Psychometric Monographs*, No. 8, 1–41.
- Rose, D. (2018). Book Review of The Oxford Compendium of Visual Illusions by Shapiro, A. G., & Todorović, D. *Perception*, *47*(8), 892–898. <https://doi.org/10.1177/0301006618778214>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. <http://www.jstatsoft.org/v48/i02/>
- Schmidt, F., & Haberkamp, A. (2016). Temporal processing characteristics of the Ponzo illusion. *Psychological Research*, *80*(2), 273–285. <https://doi.org/10.1007/s00426-015-0659-8>
- Schmidt, F., Weber, A., & Haberkamp, A. (2016). Dissociating early and late visual processing via the Ebbinghaus illusion. *Visual Neuroscience*, *33*(e016), 1–9. <https://doi.org/10.1017/S0952523816000134>
- Schoups, A. A., Vogels, R., & Orban, G. A. (1995). Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularly. *The Journal of Physiology*, *483*(3), 797–810. <https://doi.org/10.1113/jphysiol.1995.sp020623>
- Schwarzkopf, D. S., De Haas, B., & Alvarez, I. (2019). *SamSrf 6 - Toolbox for pRF modelling*. <https://doi.org/10.17605/OSF.IO/2RGSM>
- Schwarzkopf, D. S., & Rees, G. (2013). Subjective Size Perception Depends on Central Visual Cortical Magnification in Human V1. *PLoS ONE*, *8*(3), 1–12. <https://doi.org/10.1371/journal.pone.0060550>
- Schwarzkopf, D. S., Song, C., & Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nature Neuroscience*, *14*(1), 28–30. <https://doi.org/10.1038/nn.2706>
- Segall, M. H., Campbell, D. T., & Herskovits, M. J. (1963). Cultural Differences in the Perception of Geometric Illusions. *Science*, *139*(3556), 769–771. <https://doi.org/10.1126/science.139.3556.769>
- Seidman, L. J., Helleman, G., Nuechterlein, K. H., Greenwood, T. A., Braff, D. L., Cadenhead, K. S., Calkins, M. E., Freedman, R., Gur, R. E., Gur, R. C., Lazzaroni, L. C., Light, G. A., Olincy, A., Radant, A. D., Siever, L. J., Silverman, J. M., Sprock, J., Stone, W. S., Sugar, C., ... Green, M. F. (2015). Factor structure and heritability of endophenotypes in schizophrenia: Findings from the Consortium on

- the Genetics of Schizophrenia (COGS-1). *Schizophrenia Research*, 163(1–3), 73–79. <https://doi.org/10.1016/j.schres.2015.01.027>
- Shapiro, A. (2018). Visual Illusions: Nothing to Lose but Your Chains—A Reply to Oliver Braddick. *Perception*, 47(9), 901–904. <https://doi.org/10.1177/0301006618785964>
- Shapiro, A., & Todorović, D. (2017). *The Oxford Compendium of Visual Illusions*. Oxford University Press.
- Shaqiri, A., Roinishvili, M., Grzeczkowski, L., Chkonia, E., Pilz, K., Mohr, C., Brand, A., Kunchulia, M., & Herzog, M. H. (2018). Sex-related differences in vision are heterogeneous. *Scientific Reports*, 8:7521, 1–10. <https://doi.org/10.1038/s41598-018-25298-8>
- Sherman, J. A., & Chouinard, P. A. (2016). Attractive contours of the Ebbinghaus illusion. *Perceptual and Motor Skills*, 122(1), 88–95. <https://doi.org/10.1177/0031512515626632>
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428. <https://doi.org/10.1037/0033-2909.86.2.420>
- Silvar, S. D., & Pollack, R. H. (1967). Racial differences in pigmentation of the Fundus oculi. *Psychonomic Science*, 7(4), 159–159. <https://doi.org/10.3758/BF03328514>
- Simpson, W. A., & McFadden, S. M. (2005). Spatial frequency channels derived from individual differences. *Vision Research*, 45(21), 2723–2727. <https://doi.org/10.1016/j.visres.2005.01.015>
- Song, C., Schwarzkopf, D. S., & Rees, G. (2011). Interocular induction of illusory size perception. *BMC Neuroscience*, 12(1), 1–9. <https://doi.org/10.1186/1471-2202-12-27>
- Sowden, P. T., Rose, D., & Davies, I. R. L. (2002). Perceptual learning of luminance contrast detection: Specific for spatial frequency and retinal location but not orientation. *Vision Research*, 42(10), 1249–1258. [https://doi.org/10.1016/S0042-6989\(02\)00019-6](https://doi.org/10.1016/S0042-6989(02)00019-6)
- Spang, K., Grimsen, C., Herzog, M. H., & Fahle, M. (2010). Orientation specificity of learning vernier discriminations. *Vision Research*, 50(4), 479–485. <https://doi.org/10.1016/j.visres.2009.12.008>
- Spearman, C. (1904a). “General Intelligence” objectively determined and measured. *The American Journal of Psychology*, 15(2), 201–292.
- Spearman, C. (1904b). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101. <https://doi.org/10.1037/h0065390>
- Stewart, V. M. (1973). Tests of the “carpentered world” hypothesis by race and environment in

- America and Zambia. *International Journal of Psychology*, 8(2), 83–94.
<https://doi.org/10.1080/00207597308247065>
- Taylor, T. R. (1974). A factor analysis of 21 illusions: The implications for theory. *Psychologia Africana*, 15, 137–148.
- Taylor, T. R. (1976). The factor structure of geometric illusions: A second study. *Psychologia Africana*, 16, 177–200.
- Thurstone, L. L. (1944). *A factorial study of perception*. The University of Chicago Press.
- Todorović, D. (2014). On the notion of visual illusions. *Psihologija*, 47(3), 359–367.
<https://doi.org/10.2298/PSI1403359T>
- Todorović, D. (2018). In Defence of Illusions: A Reply to Braddick (2018). *Perception*, 47(9), 905–908.
<https://doi.org/10.1177/0301006618787613>
- Todorović, D. (2020). What Are Visual Illusions? *Perception*, 49(11), 1128–1199.
<https://doi.org/10.1177/0301006620962279>
- Todorović, D., & Jovanović, L. (2018). Is the Ebbinghaus illusion a size contrast illusion? *Acta Psychologica*, 185, 180–187. <https://doi.org/10.1016/j.actpsy.2018.02.011>
- Tulver, K. (2019). The factorial structure of individual differences in visual perception. *Consciousness and Cognition*, 73, 1–8. <https://doi.org/https://doi.org/10.1016/j.concog.2019.102762>
- Tulver, K., Aru, J., Rutiku, R., & Bachmann, T. (2019). Individual differences in the effects of priors on perception: A multi-paradigm approach. *Cognition*, 187(3), 167–177.
<https://doi.org/10.1016/j.cognition.2019.03.008>
- van Dijk, J. A., de Haas, B., Moutsiana, C., & Schwarzkopf, D. S. (2016). Intersession reliability of population receptive field estimates. *NeuroImage*, 143, 293–303.
<https://doi.org/10.1016/j.neuroimage.2016.09.013>
- Verhallen, R. J., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Bargary, G., & Mollon, J. D. (2017). General and specific factors in the processing of faces. *Vision Research*, 141, 217–227.
<https://doi.org/10.1016/j.visres.2016.12.014>
- Vuk, M., & Podlesek, A. (2005). The effect of inner elements of the context figures on the Ebbinghaus illusion. *Horizons of Psychology*, 14(4), 9–22.

- Wagner, D. A. (1977). Ontogeny of the Ponzo illusion: Effects of age, schooling, and environment. *International Journal of Psychology*, 12(3), 161–176. <https://doi.org/10.1080/00207597708247386>
- Wang, L. L. (2010). Disattenuation of Correlations Due to Fallible Measurement. *Newborn and Infant Nursing Reviews*, 10(1), 60–65. <https://doi.org/10.1053/j.nainr.2009.12.013>
- Ward, J., Rothen, N., Chang, A., & Kanai, R. (2017). The structure of inter-individual differences in visual ability : Evidence from the general population and synaesthesia. *Vision Research*, 141, 293–302. <https://doi.org/10.1016/j.visres.2016.06.009>
- Webster, M. A., & MacLeod, D. I. A. (1988). Factors underlying individual differences in the color matches of normal observers. *Journal of the Optical Society of America A*, 5(10), 1722–1735. <https://doi.org/10.1364/JOSAA.5.001722>
- Wechsler, D. (2003). *Wechsler intelligence scale for children—Fourth Edition (WISC-IV)*. The Psychological Corporation.
- Weintraub, D. J., & Schneck, M. K. (1986). Fragments of Delboeuf and Ebbinghaus illusions: Contour/context explorations of misjudged circle size. *Perception & Psychophysics*, 40(3), 147–158. <https://doi.org/10.3758/BF03203010>
- Wexler, M., Duyck, M., & Mamassian, P. (2015). Persistent states in vision break universality and time invariance. *Proceedings of the National Academy of Sciences of the United States of America*, 112(48), 14990–14995. <https://doi.org/10.1073/pnas.1508847112>
- Wilmer, J. B. (2008). How to use individual differences to isolate functional organization, biology, and utility of visual functions; With illustrative proposals for stereopsis. *Spatial Vision*, 21(6), 561–579. <https://doi.org/10.1163/156856808786451408>
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17(1), 37–49. [https://doi.org/10.1016/0022-3956\(82\)90033-4](https://doi.org/10.1016/0022-3956(82)90033-4)
- Yildiz, G. Y., Sperandio, I., Kettle, C., & Chouinard, P. A. (2021). Interocular transfer effects of linear perspective cues and texture gradients in the perceptual rescaling of size. *Vision Research*, 179, 19–33. <https://doi.org/10.1016/j.visres.2020.11.005>
- Yu, C., Klein, S. A., & Levi, D. M. (2004). Perceptual learning in contrast discrimination and the (minimal) role of context. *Journal of Vision*, 4(3), 169–182. <https://doi.org/10.1167/4.3.4>

Yung, Y. F., Thissen, D., & McLeod, L. D. (1999). On the relationship between the higher-order factor model and the hierarchical factor model. *Psychometrika*, *64*(2), 113–128.
<https://doi.org/10.1007/BF02294531>

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301–320.
<https://doi.org/10.1111/J.1467-9868.2005.00503.X>

Appendix A

Cretenoud, A. F., Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, 19(14):12, 1–21, <https://doi.org/10.1167/19.14.12>.

Copyright holder: © ARVO

Published under CC-BY-NC-ND 4.0

Factors underlying visual illusions are illusion-specific but not feature-specific

Aline F. Cretenoud

Laboratory of Psychophysics, Brain Mind Institute,
École Polytechnique Fédérale de Lausanne (EPFL),
Lausanne, Switzerland



Harun Karimpur

Experimental Psychology, Justus Liebig University,
Giessen, Germany
Center for Mind, Brain and Behavior (CMBB),
University of Marburg and Justus Liebig University
Giessen, Germany

Lukasz Grzeczowski

Laboratory of Psychophysics, Brain Mind Institute,
École Polytechnique Fédérale de Lausanne (EPFL),
Lausanne, Switzerland
General and Experimental Psychology,
Psychology Department,
Ludwig-Maximilian-University of Munich,
Munich, Germany

Gregory Francis

Psychological Sciences, Purdue University,
West Lafayette, IN, USA

Kai Hamburger

Experimental Psychology and Cognitive Science,
Justus Liebig University, Giessen, Germany

Michael H. Herzog

Laboratory of Psychophysics, Brain Mind Institute,
École Polytechnique Fédérale de Lausanne (EPFL),
Lausanne, Switzerland

Common factors are ubiquitous. For example, there is a common factor, *g*, for intelligence. In vision, there is much weaker evidence for such common factors. For example, visual illusion magnitudes correlate only weakly with each other. Here, we investigated whether illusions are hyper-specific as in perceptual learning. First, we tested 19 variants of the Ebbinghaus illusion that differed in color, shape, or texture. Correlations between the illusion magnitudes of the different variants were mostly significant. Second, we reanalyzed a dataset from a previous experiment where 10 illusions were tested under four conditions of luminance and found significant correlations between the different luminance conditions of each illusion. However, there were only very weak correlations between the 10 different illusions. Third, five visual illusions were tested with four

orientations. Again, there were significant correlations between the four orientations of each illusion, but not across different illusions. The weak inter-illusion correlations suggest that there is no unique common mechanism for the tested illusions. We suggest that most illusions make up their own factor.

Introduction

Common factors are ubiquitous. For example, it is widely held that there is a common factor, *g*, for intelligence (Spearman, 1904a). This factor is not measurable per se but is inferred from several indicator

Citation: Cretenoud, A. F., Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, 19(14):12, 1–21, <https://doi.org/10.1167/19.14.12>.



variables, such as the Wechsler scale (Wechsler, 2003). In metacognition and somato-sensation, there are common factors between different modalities, for example, between touch and audition (Frenzel et al., 2012). Faivre, Filevich, Solovey, Kühn, and Blanke (2018) showed that participants with high performance in one metacognitive modality are likely to show high performance in other metacognitive modalities.

However, there seems to be no unique common factor explaining individual differences in vision (for a review, see Tulver, 2019). For example, no unique common factor was found for oculomotor tasks (Bargary et al., 2017), bistable paradigms such as binocular rivalry paradigms (Brascamp, Becker, & Hambrick, 2018; Cao, Wang, Sun, Engel, & He, 2018; Wexler, 2005), and face recognition (Verhallen et al., 2017). Rather than a unique common factor, several specific factors underlying individual differences were often suggested. For example, two factors were suggested to underlie the activity of magnocellular and parvocellular systems (Peterzell & Teller, 1996; Simpson & McFadden, 2005; Ward, Rothen, Chang, & Kanai, 2017; but see Goodbourn et al., 2012). Similarly, some very specific factors underlying individual differences have been found in hue scaling (e.g., Emery, Volbrecht, Peterzell, & Webster, 2017a, 2017b) and stereopsis (Peterzell, Serrano-Pedraza, Widdall, & Read, 2017).

Hence, it seems that the structure of individual differences in perception is better represented by a multifactorial space than by a unique common factor. Bosten and colleagues (2017) found that a model with a unique common factor underlying 25 visual and auditory measures only explained about 20% of the total variance. However, eight more specific factors were identified, e.g., a factor related to stereoacuity and one related to oculomotor control, altogether explaining about 57% of the total variance. Additionally, only weak or nonsignificant correlations were found between performance in six very basic visual tasks such as visual backward masking and bisection discrimination (Cappe, Clarke, Mohr, & Herzog, 2014). Aging was expected to strengthen the correlations between visual paradigms because aging effects occur more quickly or strongly for some people. However, only weak correlations were also observed between visual paradigms in older people (Shaqiri et al., 2019).

Interestingly, we also found very weak correlations—except for one—between the magnitudes of visual illusions (Grzeczowski, Clarke, Francis, Mast, & Herzog, 2017; see also Axelrod, Schwarzkopf, Gilaie-Dotan, & Rees, 2017). Patients with schizophrenia similarly showed only weak correlations between different illusion magnitudes (Grzeczowski et al., 2018; see also Kaliuzhna et al., 2018).

Improvements in perceptual learning are very specific. For example, when trained with a vertical stimulus, there is usually no transfer of learning to the same stimulus rotated by 90° (e.g., Ball & Sekuler, 1987; Fahle & Morgan, 1996; Grzeczowski, Cretenoud, Herzog, & Mast, 2017; Schoups, Vogels, & Orban, 1995). Learning was shown to transfer only for stimuli rotated up to 10° as compared to the trained one (Spang, Grimsen, Herzog, & Fahle, 2010). Such a high degree of specificity for the trained orientation suggests that, if perceptual learning plays a role in the perception of illusions, we may even observe only weak correlations between different variants of a given illusion.

Here, we investigated whether factors for illusions are hyper-specific. In Experiment 1, we tested 19 variants of the same illusion, namely the Ebbinghaus illusion, which differed in size, color, shape, or texture. In Experiment 2, the effects of luminance contrast on illusion susceptibility were tested for 10 different visual illusions. In Experiment 3, we measured susceptibility to five visual illusions at different orientations to test whether illusion susceptibility is orientation-specific as in perceptual learning.

Experiment 1

Previously, we observed only weak correlations between the susceptibility to different visual illusions, i.e., a high susceptibility to one illusion does not inevitably imply a high susceptibility to another illusion (Grzeczowski et al., 2017). Here, we examined to what extent the susceptibility to a single visual illusion (the Ebbinghaus illusion) differs as a function of its size, color, shape, and texture.

Methods

Participants

Participants were 87 visitors of a public event at the École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. Seven of them were considered as outliers and removed from the dataset (see Data analysis section). The age of the remaining 80 participants ranged from 14 to 75 years (mean age: 48 years, 52 females). Adults signed informed consent and parents signed consent for their children. Participation was not compensated for in any form. Procedures were conducted in accordance with the Declaration of Helsinki except for the preregistration (§ 35) and were approved by the local ethics committee.

Apparatus

A BenQ XL2420T monitor (BenQ, Taipei, Taiwan) driven by a PC computer using MATLAB (R2015b, 64 bits; MathWorks, Natick, MA) and the Psychophysics toolbox (Brainard, 1997; Pelli, 1997; version 3.1, 64 bits) was used to present stimuli at a 1920×1080 pixel resolution with a 60 Hz refresh rate and a 32-bit color depth. The distance to the screen was approximately 60 cm. Participants adjusted stimuli with a Logitech LS1 computer mouse. Prior to the experiments the color look-up tables of the monitor were linearized and calibrated with a Minolta LS-100 luminance meter (Konica Minolta, Tokyo, Japan). An experimental room was especially built for this experiment to ensure controlled light conditions.

Stimuli

Each participant was tested on 19 variants of the Ebbinghaus illusion and a control condition (20 conditions in total), shown in Figure 1. For each condition, participants adjusted the size of the right central disk (adjustable target) to match the size of the left central disk (reference target) by moving the mouse on the horizontal axis. The reference target had a fixed diameter of 2.7° . Participants pressed the left mouse button to validate their adjustments. The centers of the left and right central disks were 8.6° to the left and to the right, respectively, and 2.7° to the top and to the bottom compared to the center of the screen.

In a standard variant of the Ebbinghaus illusion (STD), the reference target was a yellow disk surrounded by eight smaller yellow disks (flankers), 1.35° diameter each. The distance between the centers of the reference target and of the small flankers was 2.3° . The right adjustable target was surrounded by 6 large flankers, 3.5° diameter each, which were 4.2° away from the center of the adjustable target. The luminance was approximately 1 cd/m^2 for the background and 146 cd/m^2 for yellow disks.

Bigger and smaller right flankers compared to the STD were shown in the FBIG (right flankers were bigger) and FSMA (right flankers were smaller) conditions. The diameter of the right flankers was 4° and 3° , respectively, and their distance to the center of the adjustable target was 4.7° and 3.6° , respectively.

In the three conditions with either blue flankers, or blue targets, or both (FBLU, TBLU, and BLU), blue-cyan color was used instead of yellow without any further changes compared to the STD condition. Luminance of the blue-cyan color was the same as for the yellow color.

Three conditions used squares instead of disks, either for the flankers (FSQU) or targets (TSQU) or for both flankers and targets (SQU). The sides of the squares

were computed so that their surface equals the area of the disks used in the STD condition.

Soccer ball images were shown instead of yellow disks in the FSOC (flankers were soccer balls), TSOC (targets were soccer balls), and SOC (flankers and targets were soccer balls) conditions. Similarly, tennis ball images were used in the FTEN (flankers were tennis balls), TTEN (targets were tennis balls), and TEN (flankers and targets were tennis balls) conditions.

In two conditions, left and right flankers rotated clockwise on a circular orbit with a radius of 2.3° and 4.2° , respectively. The motion speed was either “slow” with 0.5 radian per refreshing screen (SRT for slow rotation) or “fast” with 1 radian per refreshing screen (FRT for fast rotation).

In the MRF (missing right flankers) conditions, the right flankers were removed. Likewise, the left flankers were missing in the MLF (missing left flankers) condition. Finally, we used a control condition without flankers (CON).

Procedure

The experimenter first explained the task to the participants who completed one trial of the standard variant of the illusion (STD) to familiarize themselves with the adjustment method. Then, each condition was presented twice. The two trials for each condition were presented sequentially, i.e., one after the other, and without time constraint. The order of the 20 conditions was randomized across participants. The initial size of the adjustable target was pseudorandomly chosen for each trial. The adjustable target appeared with a diameter in the range of 0° to 4.5° in the case of a circular target, except for the FBIG and FSMA conditions where the ranges were 0° to 4.9° and 0° to 3.8° , respectively. When the adjustable target was a square, the initial side of the square was pseudorandomly chosen between 0° and 3.99° , to match the global target area observed in case of a circular target.

The experimenter stayed in the experimental room during the whole experiment to answer any questions. Participants were asked to base their adjustments on their subjective perception only and to ignore potential prior knowledge about the illusion. At the end, participants were shown their results for the nine following conditions: STD, FBIG, TBLU, BLU, TSQU, SQU, SOC, TEN, and SRT.

Data analysis

As a measure of illusion magnitude for each participant and each condition, the adjusted radii (or side lengths, in case of a square target) from both trials were averaged into a mean adjustment, from which the reference disk radius (or reference side length) was

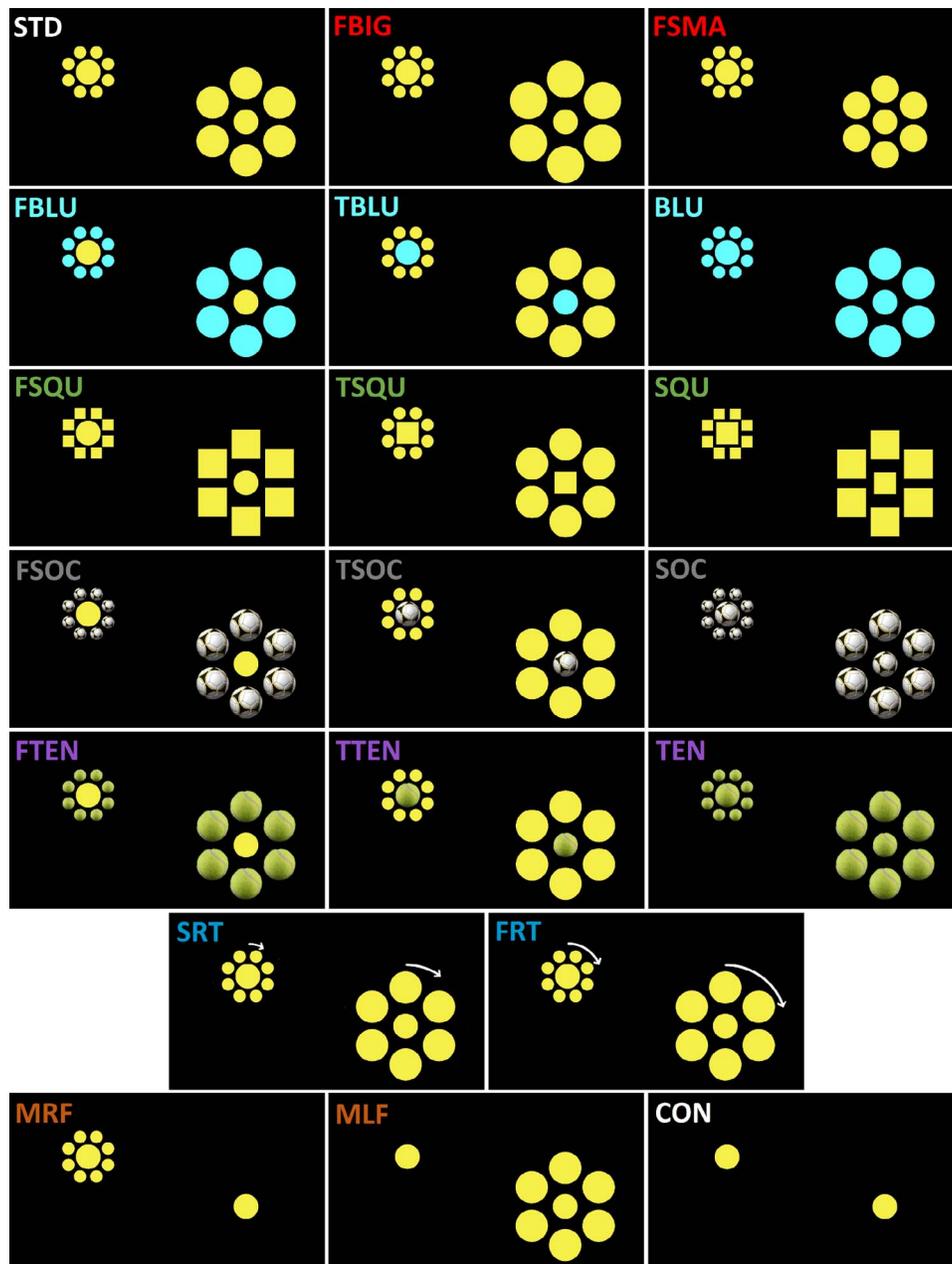


Figure 1. In 19 variants of the Ebbinghaus illusion and a control condition, participants adjusted the size of the right central disk (adjustable target) to match the size of the left central disk (reference target) by moving the mouse on the horizontal axis. The conditions of the illusion varied from the standard one (STD) as follows: bigger right flankers (FBIG), smaller right flankers (FSMA), blue flankers (FBLU), blue targets (TBLU), both blue flankers and targets (BLU), square flankers (FSQU), square targets (TSQU), both square flankers and targets (SQU), soccer ball flankers (FSOC), soccer ball targets (TSOC), both soccer ball flankers and targets (SOC), tennis ball flankers (FTEN), tennis ball targets (TTEN), both tennis ball flankers and targets (TEN), slow clockwise rotation of the flankers (SRT), fast clockwise rotation of the flankers (FRT), missing right flankers (MRF), missing left flankers (MLF), and missing all flankers (CON). For each condition, participants performed two adjustment trials. The different conditions were presented in a random order for each participant. The acronyms of similar conditions are presented in the same color.

subtracted. The result was divided by the reference disk radius (or reference side length) to express the illusion magnitude proportionally to the reference disk radius (or reference side length). A positive illusion magnitude indicates that the adjustable target looked smaller than the reference target and thereby needed to be over-

adjusted in order to appear the same size. A negative illusion magnitude indicates that the adjustable target looked larger than the reference target and thereby needed to be underadjusted in order to appear the same size. Analyses were performed with MATLAB (R2015b, 64 bit) and R (R Development Core Team, 2018).

Shapiro–Wilk tests, which test the null hypothesis that a sample comes from a normally distributed population, indicated that all distributions were approximately normally distributed. Data were therefore not transformed for further analyses.

Since data are roughly normally distributed and to ensure maximal power, we computed parametric tests with outlier removal rather than nonparametric statistics. For outlier detection, we used a modified *z*-score, which is more robust than the commonly used *z*-score (Iglewicz & Hoaglin, 1993). The modified *z*-score takes into account the median (\tilde{x}) and median absolute deviation (MAD) of a given condition instead of the mean and standard deviation, respectively, and is computed as:

$$M_i = \frac{0.6745(x_i - \tilde{x})}{MAD}$$

As suggested by Iglewicz and Hoaglin (1993), modified *z*-scores with an absolute value greater than 3.5 were considered as outliers. Seven participants showed at least one condition with a modified *z*-score larger than ± 3.5 and were removed from the dataset.

Following Shrout and Fleiss (1979) and Koo and Li (2016), intrarater reliability was assessed by computing two-way mixed effects models (intraclass correlations of type (3, 1) or ICC_{3,1}) between the two adjustments of each condition. Bravais–Pearson’s correlations were computed between the mean magnitude of each condition and participants’ age.

The mean illusion magnitude of the control condition was significantly different from zero (see the Magnitudes of the illusions section), which indicates a bias that probably occurred in all conditions. Therefore, we computed Bravais–Pearson’s partial correlations to examine the relationships between the mean magnitudes for each pair of variants, controlling for the control condition variability. A cutoff value of 0.3 for the correlation coefficient (*r*) reflects the lower limit for a medium effect size according to Cohen (1988), and a relatively large effect size according to Hemphill (2003; see also Gignac & Szodorai, 2016). The observed between-variant partial correlations were underestimated because of measurement errors (Spearman, 1904b; Wang, 2010), which are reflected by moderate intrarater reliabilities (see the Intrarater reliability section). To account for these measurement errors, which put an upper limit on the between-variant correlations, we also computed disattenuated partial correlations (Osborne, 2003):

$$r_{xy,z'} = \frac{r_{zz}r_{xy} - r_{xz}r_{yz}}{\sqrt{r_{xx}r_{zz} - r_{xz}^2}\sqrt{r_{yy}r_{zz} - r_{yz}^2}}$$

where $r_{xy,z'}$ is the disattenuated relationship between *x* and *y* controlling for the *z* variable; r_{xx} , r_{yy} , and r_{zz} are

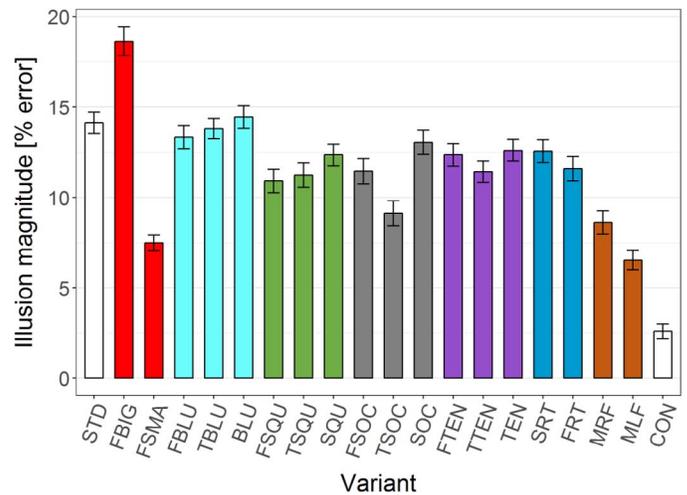


Figure 2. Illusion magnitudes [% error] ± SEM for the 19 variants of the Ebbinghaus illusion and a control condition. Similar conditions are presented in the same color, as in Figure 1.

the intrarater reliabilities of the *x*, *y*, and *z* variables and r_{xy} , r_{xz} , and r_{yz} are the attenuated (i.e., non-disattenuated) correlation coefficients between the pairs of variables.

An exploratory factor analysis (EFA) was computed to explore the factors underlying the global structure of the Ebbinghaus illusion using the guidelines outlined in Preacher, Zhang, Kim, and Mels (2013).

Results

Magnitudes of the illusions

Participants significantly overadjusted the right target compared to the left reference (Figure 2) in all 20 conditions (independent one-sample *t* test for each condition: $p < 0.001$), including the control condition (CON, illusion magnitude: 2.595% ± 0.399%, $t[79] = 6.508$, $p < 0.001$, Cohen’s $d = 0.728$).

Intrarater reliability

Moderate but significant intraclass correlations were observed for all 20 conditions even after correcting for inflated family-wise errors (Table 1, diagonal).

Correlations between illusion magnitudes and age

Contrary to previous findings (Grzeczowski et al., 2017), none of the conditions correlated significantly with age after Bonferroni correction was applied for multiple comparisons.

Partial correlations between illusion magnitudes

To control for the CON condition variability, which showed up as a significant effect in the CON condition,

	STD	FBIG	FSMA	FBLU	TBLU	BLU	FSQU	TSQU	SQU	FSOC	TSOC	SOC	FTEN	TTEN	TEN	SRT	FRT	MRF	MLF
STD	.59	.59	.52	.39	.56	.40	.41	.38	.39	.29	.34	.50	.43	.48	.29	.48	.44	.30	.23
FBIG	.93	.72	.50	.60	.63	.49	.50	.33	.59	.45	.37	.57	.55	.35	.43	.63	.56	.33	.36
FSMA	.96	.84	.50	.46	.54	.30	.34	.38	.57	.36	.42	.62	.52	.30	.49	.52	.49	.27	.38
FBLU	.65	.86	.81	.67	.63	.60	.36	.28	.45	.52	.39	.45	.58	.35	.30	.45	.48	.30	.34
TBLU	.98	.97	1	1	.60	.52	.53	.44	.53	.50	.52	.61	.59	.47	.43	.55	.55	.29	.35
BLU	.67	.73	.53	.95	.86	.63	.42	.30	.40	.49	.38	.41	.46	.43	.36	.47	.47	.24	.30
FSQU	1	1	.84	.71	1	.91	.44	.46	.39	.43	.42	.43	.51	.61	.41	.49	.47	.21	.14
TSQU	.72	.48	.74	.39	.76	.48	1	.59	.46	.36	.32	.46	.39	.45	.31	.33	.27	.25	.29
SQU	.74	1	1	.79	1	.72	.95	.89	.50	.45	.46	.65	.55	.41	.58	.51	.39	.40	.41
FSOC	.46	.63	.61	.77	.79	.75	.90	.56	.77	.68	.47	.40	.54	.29	.40	.58	.49	.18	.30
TSOC	.57	.51	.74	.56	.84	.57	.87	.48	.81	.71	.65	.58	.58	.47	.45	.48	.46	.37	.22
SOC	.89	.89	1	.70	1	.66	.95	.79	1	.63	.95	.60	.57	.50	.54	.54	.50	.33	.25
FTEN	.79	.86	1	.94	1	.76	1	.63	1	.87	.97	.99	.60	.51	.53	.55	.53	.27	.31
TTEN	.86	.52	.55	.54	.81	.71	1	.78	.76	.43	.77	.87	.90	.58	.49	.40	.49	.31	.16
TEN	.61	.69	1	.44	.78	.63	.85	.44	1	.68	.76	1	.97	.93	.60	.53	.42	.23	.40
SRT	.82	.96	.97	.69	.93	.76	1	.53	.94	.91	.76	.92	.96	.66	1	.61	.76	.10	.37
FRT	.76	.83	.88	.72	.90	.74	1	.37	.68	.75	.69	.81	.87	.82	.70	1	.65	.15	.24
MRF	.56	.49	.51	.44	.46	.37	.32	.35	.76	.24	.59	.54	.40	.51	.26	.10	.15	.57	.16
MLF	.38	.52	.69	.51	.55	.47	.23	.46	.73	.44	.31	.39	.49	.22	.71	.59	.35	.22	.63

Table 1. Diagonal (in gray): Intrarater reliability expressed as intraclass correlation coefficients (ICC_{3,1}) for each variant. All of them were significant. Upper triangle: Attenuated partial correlation coefficients between each pair of variants (Bravais–Pearson’s *r*), controlling for the control condition variability. Lower triangle: Disattenuated partial correlation coefficients between each pair of variants (Bravais–Pearson’s *r*), controlling for the control condition variability. Italics and bold font indicate significant results without and with Bonferroni correction, respectively. The color scale from white to red reflects effect sizes from *r*=0 to *r*=1. The acronyms of similar variants are presented in the same color, as in Figure 1.

partial correlations were computed for each pair of variants based on the mean of both adjustments for each participant. We analyzed attenuated (Table 1, upper triangle) and disattenuated (Table 1, lower triangle) partial correlation coefficients without and with Bonferroni correction and observed strong between-variant effects. Indeed, out of 171 attenuated correlations only 26 showed *r* < 0.3. Similarly, only seven disattenuated correlations showed *r* < 0.3 (Figure 3).

Exploratory factor analysis (EFA)

A one-factor model (explaining ~44% of the variance) was suggested by a parallel analysis and scree plot inspection. All conditions except the CON condition highly loaded onto a unique factor (factor loadings—STD: 0.600; FBIG: 0.764; FSMA: 0.671; FBLU: 0.692; TBLU: 0.799; BLU: 0.642; FSQU: 0.676; TSQU: 0.580;

SQU: 0.729; FSOC: 0.647; TSOC: 0.670; SOC: 0.778; FTEN: 0.787; TTEN: 0.647; TEN: 0.684; SRT: 0.757; FRT: 0.722; MRF: 0.436; MLF: 0.451; CON: 0.293).

Data simulation

As an estimate of the experimental power, we simulated data to estimate the likelihood to observe at least 145 out of 171 correlations with *r* > 0.3 if there truly are no correlations between the variants. Individual mean adjustments for all 19 variants were simulated from a normal distribution (*M* = 0, *SD* = 1) and correlation coefficients were computed from these simulated values. The probability of observing at least 145 correlations with *r* > 0.3 averaged across 10,000 simulations was smaller than 0.001. Similarly, this probability was also smaller than 0.001 when we simulated weak correlation coefficients from a normal distribution with mean and standard deviation computed from the inter-illusion

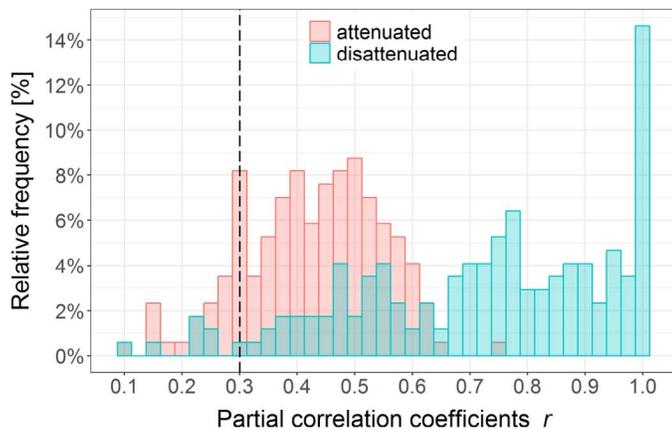


Figure 3. Relative frequency (in percentage) of attenuated (red) and disattenuated (turquoise) partial correlation coefficients r in Experiment 1. According to Cohen’s convention (Cohen, 1988), $r = 0.3$ (vertical dashed line) is the lower limit corresponding to a medium effect size. Only 26 attenuated and seven disattenuated correlation coefficients out of 171 showed $r < 0.3$.

correlation coefficients of Experiment 3 ($M = 0.16$, $SD = 0.12$; see Experiment 3 Relations between illusion magnitudes section), suggesting that our results are unlikely to be false alarms. However, when simulating r coefficients from a normal distribution centered on the average attenuated correlation coefficient, the likelihood of observing at least 145 out of 171 correlations with $r > 0.3$ was 0.74, which suggests that a new study with our sample size has a pretty good chance of showing a similar pattern of results.

Experiment 2

Hamburger, Hansen, and Gegenfurtner (2007) previously tested 10 visual illusions under different luminance conditions and observed high correlations between the different luminance conditions for each illusion. However, the authors did not analyze the relationships between the different illusions, which we did here. We added four subjects who joined a pilot experiment in 2007. There are no changes between the pilot and the reported experiment.

Methods

Participants

Twenty-one students (nine females, mean age: 26 years, age range: 20–48 years) of the Justus Liebig University Giessen (JLU), Germany, were considered for the analysis (three outliers; see Data analysis section). Participants signed informed consent and received course credits for participation. Participants

had to correctly answer all Ishihara pseudo-isochromatic plates to take part in this experiment. Procedures were conducted in accordance with the Declaration of Helsinki except for the preregistration (§ 35) and were approved by the local ethics committee.

Apparatus

The experiment was conducted in a dark room. The distance between the participant and the monitor was 60 cm and a chinrest was used for head stabilization, but participants were free to move their eyes. Observation was binocular and the stimuli were presented in the center of the monitor at the line of sight. The stimuli were presented on a 21-in. Iiyama Vision Master Pro 513 CRT monitor at a refresh rate of 85 Hz noninterlaced with a resolution of 1154×768 pixels, driven by an NVIDIA graphics card with a color resolution of 8 bits for each of the three monitor primaries. We linearized the relationship between luminance and voltage output by a color look-up table for each primary.

Color space

Four different luminance/color conditions were included (Figure 4): a high luminance contrast condition (50%; Lum), a low luminance contrast condition (10%; LumLow), a “red-green” [L – M] isoluminant condition (Iso), and a “blue-yellow” [S – (L + M)] isoluminant condition (IsoS). Both isoluminant conditions were derived from the DKL color space (Derrington, Krauskopf, & Lennie, 1984; Krauskopf, Williams, & Heeley, 1982).

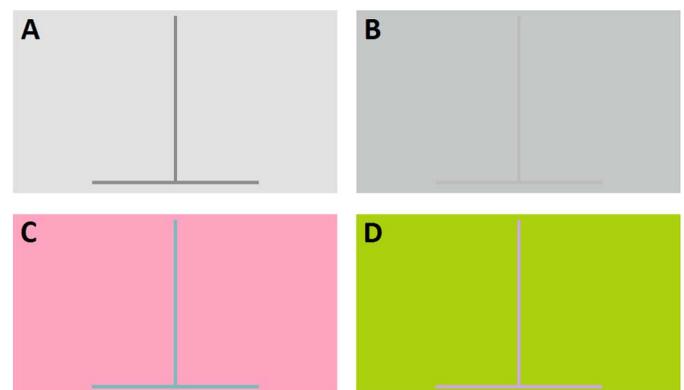


Figure 4. Exemplary horizontal-vertical stimulus with two luminance conditions: (A) 50% (Lum) and (B) 10% (LumLow) contrast and two isoluminant color conditions: (C) L – M (Iso) and (D) S – (L + M) (IsoS) according to DKL color space (Derrington et al., 1984). Please note that the colors here may vary in luminance due to reproduction.

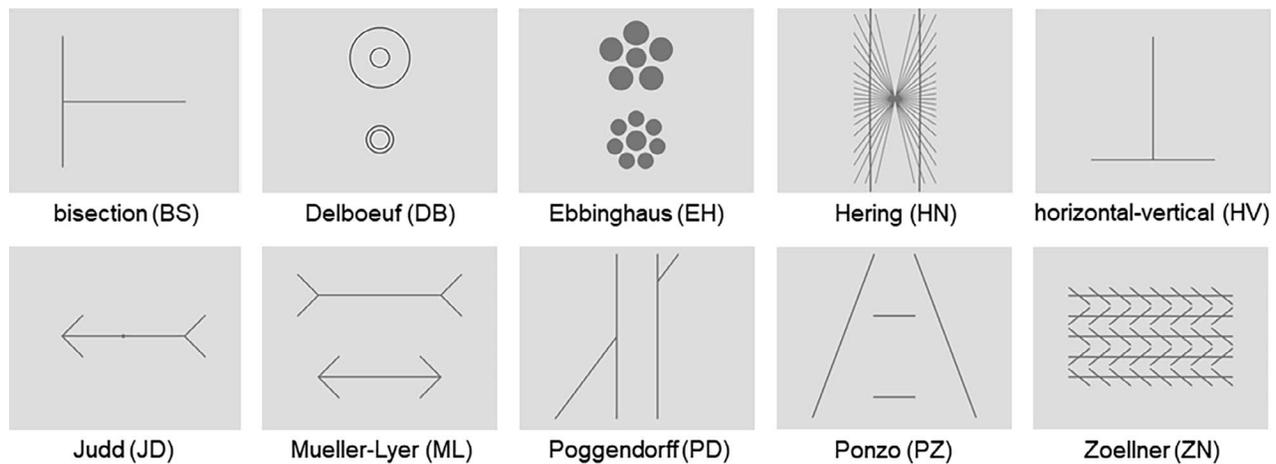


Figure 5. The 10 visual illusions used in Experiment 2. Observers were asked to reach subjective equality in length (BS, HV, JD, ML, PZ), size (DB, EH), curvature (HN), position (PD), and orientation (ZN) with an adjustment procedure. For a detailed description of the different adjustments, please see Hamburger et al. (2007).

Stimuli

The stimuli consisted of 10 classic visual illusions (Figure 5). These illusions were: bisection (BS), Delboeuf (DB), Ebbinghaus (EH), Hering (HN), horizontal-vertical (HV), Judd (JD), Müller-Lyer (ML), Poggendorff (PD), Ponzo (PZ), and Zöllner (ZN). Stimuli were created with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) in MATLAB (version 7, 32 bit). Stimulus size (including the background) was 17° of the visual field for all illusions with a constant line width of 4 pixels (0.14°). Only in the HN illusion the radial lines have a smaller width of 1 pixel (0.035°) in order to retain clearly visible edges within the illusion.

Procedure

The instruction for the participants was to reach subjective equality in length (BS, HV, JD, ML, PZ), size (DB, EH), curvature (HN), position (PD), and orientation (ZN) by pressing the right and left arrow keys of a standard keyboard. The stimulus conditions were randomly intermixed and each illusion was presented under four different luminance conditions, four trials each (160 trials in total). For a detailed description of the methods please see Hamburger et al. (2007).

Data analysis

We first aggregated the 10 illusion magnitudes for each of the four luminance conditions and all participants over all trials. We then standardized these results for each illusion by computing modified z -scores as in Experiment 1 to allow comparison across the different illusions. Three participants showed modified z -scores larger than ± 3.5 in at least one condition and were removed from the dataset. A repeated-measures

analysis of variance (ANOVA) with two main factors (illusion and luminance condition) was computed. As in Experiment 1, we computed intrarater reliability between the four adjustments of each condition. We computed Bravais–Pearson’s r correlation coefficients between illusion magnitudes over all illusions and luminance conditions and computed disattenuated correlation coefficients as:

$$r_{xy'} = \frac{r_{xy}}{\sqrt{r_{xx}r_{yy}}}$$

where $r_{xy'}$ is the disattenuated relationship between x and y , r_{xx} and r_{yy} are the intrarater reliabilities of the x and y variables, and r_{xy} is the attenuated correlation coefficient between x and y . Similarly to Experiment 1, we conducted an EFA. Oblique rotations allow factors to correlate, while uncorrelated factors result from orthogonal rotations. Since we have no reason to preclude correlated factors from our datasets, we used an oblique rotation (promax) for the maximum likelihood estimation rather than an orthogonal rotation. If the factors are uncorrelated, both oblique and orthogonal rotations produce very similar results (e.g., Costello & Osborne, 2005).

Results

Magnitudes of the illusions

Figure 6 shows the standardized illusion magnitudes as a function of illusion and luminance conditions. The z -scores show to what extent the illusion magnitudes deviate as a function of luminance for each illusion type (see Hamburger et al., 2007, for the nontransformed illusion magnitudes). A repeated-measures ANOVA with illusion and luminance conditions as within-subject factors yielded no significant main effect of

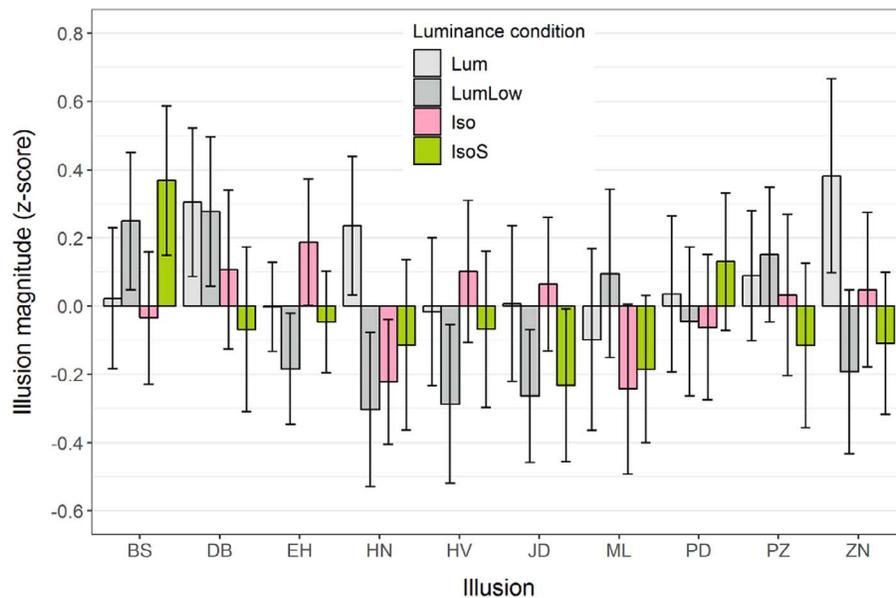


Figure 6. Standardized illusion magnitudes \pm SEM as a function of illusion and luminance conditions. The z-scores show to what extent the illusion magnitudes deviate as a function of luminance for each illusion type.

illusion ($p = 0.980$) but a significant main effect of luminance condition ($F[3, 60] = 4.292$, $p = 0.008$, $\eta_p^2 = 0.177$) and a significant interaction between illusion and luminance condition ($F[27, 540] = 3.679$, $p < 0.001$, $\eta_p^2 = 0.155$).

Intrarater reliability

For each combination of illusion and luminance condition, we computed intraclass correlations ($ICC_{3,1}$) between the adjustments of all participants (Table 2, diagonal). For all but three conditions we found significant intraclass correlations after correcting the alpha level for inflated family-wise errors. In those three cases, the intraclass correlations yielded significance without correction (EH-IsoS: ICC coef. = 0.225, 95% CI [0.027, 0.488], $F[20, 60] = 2.164$, $p = 0.011$; ZN-Iso: ICC coef. = 0.218, 95% CI [0.022, 0.481], $F[20, 60] = 2.117$, $p = 0.013$; ZN-IsoS: ICC coef. = 0.228, 95% CI [0.030, 0.491], $F[20, 60] = 2.182$, $p = 0.011$).

Correlations between illusion magnitudes

We show both attenuated (Table 2, upper triangle) and disattenuated (Table 2, lower triangle) correlation coefficients without and with Bonferroni correction. Among 780 correlations, 253 attenuated and 455 disattenuated correlations resulted in $r > 0.3$. For each illusion, the different luminance conditions were highly correlated ($r > 0.3$ for all intra-illusion correlations), while 27% (193 out of 720) and 55% (395 out of 720) of the attenuated and disattenuated inter-illusion correlations showed $r > 0.3$, respectively

(Figure 7). More specifically, it seems that most illusions were not strongly related to each other, except the BS illusion, which was strongly linked to the HV illusion. In fact, the BS illusion is an HV illusion rotated by 90° .

Exploratory factor analysis (EFA)

We inspected the scree plot of the EFA and identified nine factors. These factors accounted for $\sim 90\%$ of the variance (RF1: 18%, RF2: 11%, RF5: 10%, RF3: 10%, RF4: 9%; RF6: 8%; RF8: 8%; RF7:

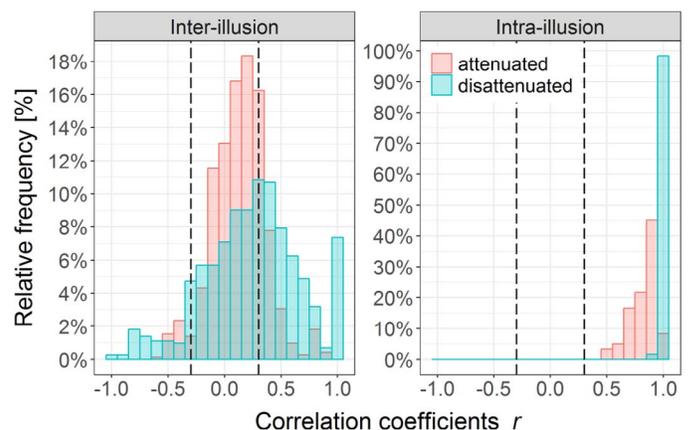


Figure 7. Relative frequency (in percentage) of attenuated (red) and disattenuated (turquoise) correlation coefficients r for inter-illusion (left panel) and intra-illusion (right panel) correlations in Experiment 2. According to Cohen's convention (Cohen, 1988), $r = \pm 0.3$ (vertical dashed lines) is the lower limit corresponding to a medium effect size. Please note the different scales of the y axes.

	BS Lum	BS LumLow	BS Iso	BS IsoS	DB Lum	DB LumLow	DB Iso	DB IsoS	EH Lum	EH LumLow	EH Iso	EH IsoS	HN Lum	HN LumLow	HN Iso	HN IsoS	HV Lum	HV LumLow	HV Iso	HV IsoS	JD Lum	JD LumLow	JD Iso	JD IsoS	ML Lum	ML LumLow	ML Iso	ML IsoS	PD Lum	PD LumLow	PD Iso	PD IsoS	PZ Lum	PZ LumLow	PZ Iso	PZ IsoS	ZN Lum	ZN LumLow	ZN Iso	ZN IsoS
BS Lum	.67	.90	.92	.87	.00	.02	-.08	-.06	.10	.03	.09	-.02	.01	.01	-.24	.16	.80	.78	.79	.77	.39	.13	.27	.30	.05	.08	.11	.03	.44	.32	.32	.32	-.19	-.41	-.47	-.39	.11	.06	-.07	-.13
BS LumLow	1	.42	.91	.96	.25	.29	.19	.16	.06	.00	.04	.07	.20	.11	-.07	.32	.81	.82	.85	.84	.27	-.02	.33	.34	.22	.24	.21	.15	.39	.31	.30	.31	-.11	-.50	-.56	-.42	.26	.20	.12	.09
BS Iso	1	1	.70	.92	.06	.14	.06	.05	.14	.04	.16	.13	.08	.10	-.08	.28	.88	.87	.88	.84	.41	.20	.44	.41	.29	.30	.31	.28	.46	.32	.33	.35	-.14	-.39	-.46	-.32	.26	.31	.06	.03
BS IsoS	1	1	1	.55	.09	.13	.03	.00	.04	-.01	.06	.03	.13	.04	-.13	.25	.81	.80	.82	.82	.36	.07	.36	.37	.23	.22	.24	.17	.31	.27	.29	.27	-.10	-.43	-.50	-.34	.21	.19	.10	.03
DB Lum	.00	.48	.09	.15	.67	.96	.93	.89	-.21	-.33	-.08	.20	.15	.16	.16	.34	.07	.04	.19	.13	-.48	-.43	.04	-.10	.21	.28	.08	.10	.03	-.09	-.07	-.01	.11	-.14	-.06	-.08	.10	.20	.29	.46
DB LumLow	.02	.53	.19	.21	1	.71	.98	.93	-.10	-.19	.01	.28	.23	.20	.21	.40	.13	.11	.22	.15	-.43	-.40	.17	-.10	.35	.40	.22	.24	.05	-.08	-.09	-.03	.10	-.15	-.08	-.04	.27	.38	.43	.57
DB Iso	-.10	.34	.08	.04	1	1	.79	.97	-.05	-.15	.02	.32	.25	.24	.27	.41	.07	.05	.16	.10	-.49	-.42	.11	-.14	.36	.39	.22	.25	.05	-.08	-.10	.01	.06	-.16	-.06	-.07	.25	.40	.41	.58
DB IsoS	-.08	.27	.07	.00	1	1	1	.79	.04	-.11	.09	.33	.18	.20	.13	.35	.09	.02	.11	.06	-.43	-.37	.12	-.15	.33	.31	.16	.19	.02	-.10	-.15	-.06	-.02	-.14	-.03	-.07	.18	.39	.38	.56
EH Lum	.20	.14	.27	.10	-.43	-.19	-.10	.07	.37	.89	.74	.74	.24	.14	-.02	-.10	.26	.25	.10	.24	.48	.46	.52	.31	-.12	-.28	-.24	-.11	.18	.24	.02	-.02	-.26	-.01	-.08	.03	.57	.56	.45	.37
EH LumLow	-.05	.00	.07	-.01	-.52	-.29	-.22	-.16	1	.61	.60	.64	.33	.21	.11	-.05	.14	.22	.04	.19	.36	.33	.35	.19	-.07	-.17	-.13	.02	.11	.18	-.01	-.05	-.13	.05	.01	.08	.62	.53	.47	.23
EH Iso	-.13	.07	.24	.10	-.12	.02	.03	.12	1	.94	.68	.78	.35	.30	.16	-.05	.30	.19	.19	.35	.35	.53	.21	.03	-.02	-.06	-.01	.18	.25	.19	.06	-.02	.09	.15	.27	.46	.53	.45	.44	
EH IsoS	-.05	.24	.33	.09	.51	.70	.75	.79	1	1	1	.23	.44	.37	.37	.13	.28	.33	.27	.32	.09	.17	.48	.12	.02	.00	-.13	.05	.03	.09	-.02	-.04	-.09	-.08	-.10	.04	.63	.69	.70	.54
HN Lum	.01	.46	.14	.26	.26	.39	.41	.29	.57	.60	.61	1	.48	.83	.75	.72	.13	.24	.24	.26	-.01	-.11	.25	.13	.16	.17	.16	.24	.21	.37	.40	.33	.10	-.19	-.02	-.01	.50	.40	.50	.45
HN LumLow	.02	.25	.18	.08	.28	.35	.40	.33	.33	.38	.54	1	1	.47	.77	.82	.24	.29	.31	.28	.10	.02	.24	.24	.14	.21	.25	.39	.14	.13	.27	.22	.10	-.12	.18	-.02	.38	.33	.33	.29
HN Iso	-.50	-.18	-.15	-.30	.32	.42	.50	.25	-.05	.23	.32	1	1	1	.36	.65	-.04	.14	.19	.10	-.20	-.11	.11	.07	.08	.19	.16	.31	.02	.07	.20	.18	.07	-.06	.14	.10	.33	.28	.35	.27
HN IsoS	.24	.63	.42	.42	.52	.60	.58	.50	-.21	-.09	-.08	.34	1	1	1	.64	.28	.29	.36	.35	-.01	-.13	.22	.15	.32	.36	.44	.51	.16	.12	.23	.25	-.05	-.33	-.10	-.20	.29	.30	.29	.29
HV Lum	1	1	1	1	.10	.18	.09	.12	.51	.21	.43	.70	.23	.41	-.08	.42	.72	.92	.92	.90	.45	.19	.42	.40	.20	.25	.24	.22	.33	.16	.22	.24	-.14	-.47	-.47	-.43	.36	.37	.18	.10
HV LumLow	1	1	1	1	.06	.16	.07	.03	.49	.34	.44	.83	.41	.51	.29	.43	1	.69	.95	.91	.32	.09	.29	.32	.18	.26	.19	.21	.32	.17	.25	.26	.01	-.33	-.38	-.30	.41	.38	.21	.07
HV Iso	1	1	1	1	.32	.35	.24	.17	.22	.07	.31	.78	.46	.62	.43	.61	1	1	.55	.93	.25	.04	.32	.37	.11	.23	.18	.19	.32	.16	.27	.30	-.05	-.40	-.40	-.36	.33	.31	.18	.07
HV IsoS	1	1	1	1	.19	.20	.13	.08	.45	.28	.26	.79	.43	.48	.19	.51	1	1	1	.75	.30	.09	.29	.38	.14	.20	.16	.19	.41	.29	.33	.41	-.12	-.46	-.51	-.46	.39	.39	.22	.07
JD Lum	.68	.61	.71	.70	-.85	-.73	-.79	-.70	1	.67	.61	.28	-.01	.20	-.49	-.01	.76	.55	.48	.50	.48	.81	.70	.78	-.06	-.18	.04	.07	.26	.23	.17	.03	-.24	-.12	-.11	-.04	.25	.11	-.06	-.02
JD LumLow	.26	-.05	.38	.14	-.84	-.75	-.75	-.67	1	.67	.68	.58	-.26	.04	-.30	-.25	.35	.16	.09	.16	1	.39	.68	.75	-.01	-.17	.04	.19	.29	.30	.22	.14	-.36	-.01	.00	.02	.21	.18	-.05	-.03
JD Iso	.53	.80	.83	.76	.08	.31	.20	.22	1	.70	1	1	.58	.54	.28	.45	.78	.56	.68	.52	1	1	.40	.72	.18	.07	.19	.28	.27	.26	.17	.04	-.28	-.20	-.13	.06	.54	.52	.41	.46
JD IsoS	.56	.79	.75	.76	-.19	-.17	-.24	-.26	.79	.37	.39	.39	.30	.54	.17	.28	.71	.59	.75	.66	1	1	1	.44	.01	-.09	.06	.18	.46	.39	.40	.27	-.15	-.05	.02	.01	.24	.21	-.05	.08
ML Lum	.06	.40	.41	.36	.29	.49	.47	.43	-.22	-.11	.04	.06	.27	.23	.15	.47	.27	.25	.18	.18	-.11	-.02	.32	.02	.74	.92	.90	.86	.27	.18	.18	.27	.18	-.20	-.18	-.13	.22	.41	.05	.13
ML LumLow	.12	.47	.44	.37	.43	.59	.55	.43	-.57	-.27	-.04	.00	.31	.38	.39	.55	.36	.38	.39	.29	-.32	-.34	.13	-.17	1	.65	.90	.84	.24	.06	.15	.23	.30	-.23	-.17	-.16	.24	.34	.05	.07
ML Iso	.15	.37	.43	.36	.11	.30	.28	.21	-.45	-.19	-.08	-.31	.27	.42	.30	.63	.33	.26	.27	.21	.06	.07	.35	.11	1	1	.77	.93	.24	.08	.16	.26	.12	-.25	-.12	-.18	.15	.26	-.08	-.05
ML IsoS	.04	.29	.42	.28	.15	.35	.35	.27	-.23	.02	-.02	.13	.43	.70	.63	.80	.32	.32	.32	.27	.13	.38	.55	.34	1	1	1	.65	.18	.04	.10	.21	.10	-.20	-.07	-.16	.30	.39	.07	.02
PD Lum	.75	.84	.75	.57	.06	.07	.08	.03	.41	.20	.30	.09	.42	.29	.04	.27	.53	.53	.60	.65	.51	.65	.58	.96	.44	.41	.38	.32	.52	.89	.86	.86	-.04	-.12	-.14	-.08	.20	.32	-.15	.09
PD LumLow	.58	.69	.55	.53	-.16	-.14	-.12	-.17	.58	.33	.44	.27	.77	.28	.16	.22	.27	.30	.31	.48	.48	.70	.61	.86	.30	.11	.13	.07	1	.47	.93	.87	-.10	-.09	-.13	.01	.21	.32	.00	.20
PD Iso	.51	.61	.52	.51	-.11	-.13	-.14	-.22	.04	-.01	.30	-.06	.75	.52	.43	.38	.33	.39	.47	.50	.32	.45	.35	.78	.27	.24	.24	1	1	.59	.91	.05	.00	.03	.10	.11	.25	-.08	.11	
PD IsoS	.57	.70	.61	.53	-.02	-.06	.01	-.10	-.04	-.09	.10	-.12	.69	.46	.43	.46	.42	.45	.59	.68	.05	.33	.10	.59	.45	.42	.43	1	1	1	.48	-.04	-.16	-.15	-.16	.12	.27	-.12	.06	
PZ Lum	-.34	-.26	-.26	-.21	.21	.18	.09	-.03	-.63	-.25	-.04	-.27	.22	.21	.16	-.10	-.25	.01	-.09	-.21	-.52	-.84	-.65	-.34	.32	.56	.21	.17	-.08	-.22	.09	-.08	.46	.63	.49	.53	.12	.06	.00	.08
PZ LumLow	-.75	-.1	-.70	-.85	-.25	-.27	-.27	-.23	-.03	.10	.17	-.24	-.42	-.26	-.15	-.62	-.82	-.59	-.80	-.80	-.26	-.02	-.47	-.12	-.34	-.43	-.43	-.37	-.25	-.19	.00	-.35	1	.45	.87	.90	-.09	.04	-.06	.06
PZ Iso	-.81	-.1	-.78	-.94	-.11	-.13	-.10	-.04	-.20	.02	.25	-.30	-.03	.36	.32	-.18	-.77	-.65	-.76	-.82	-.22	.00	-.29	.04	-.29															

		RF1	RF2	RF5	RF3	RF4	RF6	RF8	RF7	RF9
BS	Lum	0.90	-0.02	-0.12	-0.08	0.12	-0.15	0.05	-0.05	-0.07
BS	LumLow	0.89	0.09	-0.08	-0.18	0.07	-0.06	0.04	-0.09	0.14
BS	Iso	0.90	0.02	0.11	-0.03	0.05	-0.12	0.15	0.02	-0.01
BS	IsoS	0.91	-0.08	-0.01	-0.21	0.00	-0.11	0.10	-0.01	0.19
DB	Lum	0.09	0.95	-0.10	-0.15	0.00	0.01	-0.03	0.01	0.02
DB	LumLow	0.07	0.90	0.06	-0.06	-0.05	-0.02	-0.01	0.02	0.15
DB	Iso	-0.05	0.92	0.08	0.05	0.00	0.03	-0.08	-0.05	0.06
DB	IsoS	-0.06	0.97	0.06	0.15	-0.03	-0.05	-0.04	-0.06	-0.04
EH	Lum	-0.03	0.00	-0.12	0.91	0.04	-0.13	0.13	-0.14	0.02
EH	LumLow	-0.04	-0.27	0.01	0.82	-0.02	0.01	-0.04	-0.06	0.11
EH	Iso	0.08	0.07	-0.02	0.75	0.06	0.01	0.10	0.17	-0.02
EH	IsoS	0.05	0.23	-0.07	0.86	-0.08	0.12	-0.08	-0.07	0.06
HN	Lum	-0.05	-0.12	-0.08	0.11	0.21	0.78	-0.10	-0.07	0.26
HN	LumLow	0.06	0.03	0.02	0.10	-0.03	0.97	0.11	0.07	-0.13
HN	Iso	-0.13	-0.04	-0.01	0.04	0.01	0.86	-0.07	0.02	0.04
HN	IsoS	0.09	0.11	0.12	-0.27	-0.03	0.79	0.14	-0.13	0.10
HV	Lum	0.90	0.03	0.06	0.18	-0.10	-0.01	0.09	-0.05	-0.09
HV	LumLow	1.01	-0.08	0.03	0.24	-0.09	0.09	-0.11	0.11	-0.09
HV	Iso	1.00	0.08	-0.08	0.06	-0.06	0.18	-0.02	0.06	-0.10
HV	IsoS	0.88	-0.01	-0.05	0.17	0.07	0.11	-0.09	-0.10	-0.05
JD	Lum	0.20	-0.31	0.00	0.03	-0.12	-0.02	0.80	-0.01	0.03
JD	LumLow	-0.19	-0.16	0.12	0.16	0.06	-0.09	0.84	-0.07	-0.09
JD	Iso	0.03	0.16	0.08	0.10	-0.09	0.00	0.84	-0.05	0.29
JD	IsoS	0.17	0.07	-0.07	-0.11	0.12	0.14	0.86	0.11	-0.08
ML	Lum	-0.09	0.09	0.97	0.04	0.11	-0.16	-0.02	-0.02	0.02
ML	LumLow	0.07	0.06	0.94	-0.01	0.03	-0.04	-0.18	0.04	-0.02
ML	Iso	-0.03	-0.05	0.99	-0.11	-0.02	0.06	0.10	-0.03	-0.10
ML	IsoS	-0.09	-0.03	0.96	0.00	-0.10	0.19	0.20	-0.02	-0.10
PD	Lum	0.09	0.14	0.10	0.10	0.88	-0.10	0.10	0.01	-0.17
PD	LumLow	-0.11	-0.07	-0.07	0.06	1.00	-0.04	0.04	-0.08	0.12
PD	Iso	0.06	-0.09	-0.06	-0.10	0.93	0.16	0.03	0.14	-0.01
PD	IsoS	0.02	-0.04	0.07	-0.02	0.93	0.10	-0.13	-0.08	-0.09
PZ	Lum	0.21	-0.09	0.19	-0.08	-0.03	0.00	-0.29	0.74	0.07
PZ	LumLow	-0.07	-0.02	-0.09	0.07	0.01	-0.15	0.01	0.93	-0.04
PZ	Iso	-0.21	0.11	-0.06	0.00	-0.04	0.18	0.17	0.83	-0.22
PZ	IsoS	-0.14	-0.03	-0.11	-0.02	0.03	-0.07	0.15	0.88	0.25
ZN	Lum	0.10	-0.14	0.14	0.40	-0.03	0.08	0.03	0.02	0.62
ZN	LumLow	0.05	0.11	0.27	0.52	0.15	-0.07	0.00	0.12	0.44
ZN	Iso	0.01	0.01	-0.12	0.32	-0.19	0.11	-0.10	-0.03	0.80
ZN	IsoS	-0.14	0.33	-0.16	0.13	0.09	0.01	0.12	0.11	0.73

Table 3. Rotated factor loadings from an exploratory factor analysis (EFA) after promax rotation for all illusions and orientations. A color scale from blue (negative loadings) to red (positive loadings) is shown.

$SD = 1$) and inter-illusion correlation coefficients were computed from these simulated values. The probability of observing less than 100% of intra-illusion and more than 27% of inter-illusion correlation coefficients larger than 0.3 across 10,000 simulations was null, since it never happened that more than 27% of simulated inter-illusion correla-

tions showed $r > 0.3$. On average, 15% and 86% of simulated inter- and intra-illusion correlation coefficients were larger than 0.3, respectively, suggesting that we may underestimate the true inter- and intra-illusion effect sizes.

Importantly, a strongly significant difference was observed when computing a Welch t test between the

observed inter- and intra-illusion correlation coefficients (two-tailed t test, $t[110.46] = 40.81$, $p < 0.001$), suggesting a true difference between inter- and intra-illusion correlations despite the small sample size and the very large number of comparisons.

Experiment 3

In the first experiment, we observed that individual differences in visual illusions are not specific to features such as color, shape, or texture. The second experiment suggested that individual differences in visual illusions are not specific to luminance changes. In addition, there were only very weak associations between different visual illusions, except between the BS illusion and the HV illusion (see also Hamburger & Hansen, 2010), which is in fact a BS illusion rotated by 90° . We here tested whether individual differences for visual illusions are stable across changes in orientation.

Methods

Participants

Twenty students of the EPFL participated in this experiment (seven females, mean age: 23 years, age range: 18–28 years). Participants signed informed consent prior to the experiment and were paid 20 Swiss Francs per hour. Procedures were conducted in accordance with the Declaration of Helsinki except for the pre-registration (§ 35) and were approved by the local ethics committee.

Apparatus

The same experimental setup as in Experiment 1 was used except that the experiment was conducted in the Laboratory of Psychophysics at EPFL, Switzerland.

Stimuli

Five illusions were tested: horizontal-vertical (HV), Müller-Lyer (ML), Poggendorff (PD), Ponzo (PZ) and Zöllner (ZN). Stimuli were presented in white (≈ 176 cd/m²) on a black background (≈ 1 cd/m²). In the HV, ML, and PZ illusions, participants had to adjust the length of a target line to match the length of a reference line by moving the computer mouse on the horizontal axis. In the PD, the right part of the interrupted diagonal had to be vertically adjusted along the right parallel line by moving the computer mouse from left to right so that it appeared to be in a continuum with the left part of the interrupted diagonal. In the ZN illusion, moving the mouse on the

horizontal axis changed the alignment of the main streams (two neighbor streams always bent in opposite directions). Participants validated the trial when they perceived these main streams to be perfectly parallel. All lines were drawn with a 4-pixel width.

Each illusion was presented under four different orientations: -60° , -15° , 30° , and 75° (Figure 8). Illusions are described in detail in Supplementary File S1.

Procedure

The experimenter first explained the task to the participants who adjusted each illusion in the 0° orientation for one trial to familiarize with the task. Then, each of the 20 conditions (5 illusions \times 4 orientations) was presented twice without time restriction. The two trials for each condition were presented sequentially, i.e., one after the other, but the order of the 20 conditions was randomized across participants.

As in Experiment 1, participants were asked to perform the task relying on their percepts only. Participants could ask questions at any time since the experimenter stayed in the experimental room during the whole experiment. Contrary to Experiment 1, participants were not shown their own results at the end of the experiment.

Data analysis

For each participant and each condition, the adjustments from both trials were averaged. Then, the reference value of each condition was subtracted from the averaged values. In order to make the scores comparable across illusions, illusion magnitudes were turned into modified z -scores for illusion type. No outliers were detected based on the same outlier detection method as in Experiments 1 and 2. A repeated-measures ANOVA with two main factors (illusion and orientation) was computed. Similarly to Experiment 2, we computed intrarater reliabilities as well as a correlation table with both attenuated and disattenuated correlation coefficients and an EFA with an oblique rotation method (promax).

Results

Magnitudes of the illusions

Standardized illusion magnitudes are plotted for each illusion and each orientation in Figure 9. A repeated-measures ANOVA was computed with the main factors of illusion (HV, ML, PD, PZ, ZN) and orientation (-60° , -15° , 30° , and 75°). There was a significant interaction ($F[12, 228] = 3.996$, $p < 0.001$, $\eta_p^2 = 0.174$) and a significant main effect of orientation

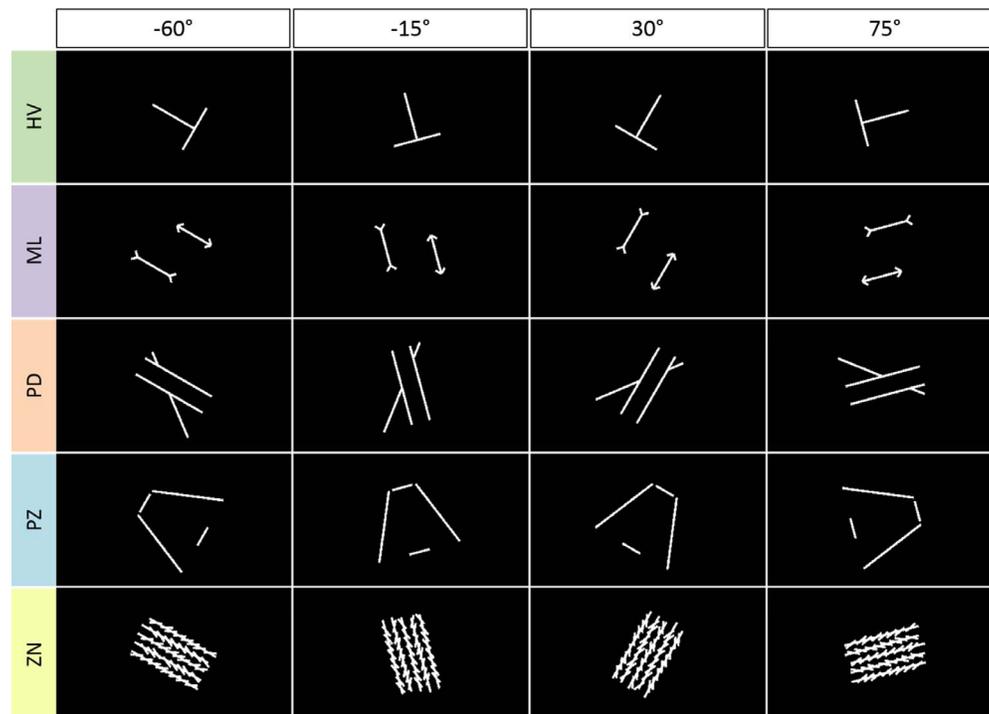


Figure 8. The five illusions used in Experiment 3, tested with four different orientations. By row: horizontal-vertical (HV), Müller-Lyer (ML), Poggendorff (PD), Ponzo (PZ), and Zöllner (ZN) illusions. By column: -60° , -15° , 30° , and 75° . In the HV, ML, and PZ illusions, participants adjusted the length of a line to match the length of a reference line. In the PD illusion, participants adjusted the position of the right part of the interrupted diagonal so that it appeared to be in a continuum with the left part. In the ZN illusion, participants aligned the five main streams in a parallel way. All adjustments were made by moving the computer mouse on the horizontal axis. Each condition was presented twice and the order of presentation was randomized across participants.

($F[3, 57] = 13.827$, $p < 0.001$, $\eta_p^2 = 0.421$). There was no main effect of illusion ($p = 0.880$).

Intrarater reliability

Intraclass correlations ($ICC_{3,1}$) were computed between the first and second adjustments of all participants for each condition (Table 4, diagonal). Five out of 20 intraclass correlations were not significant after Bonferroni correction but yielded significance without correction, suggesting an overall moderate intrarater reliability (PD -15° : ICC coef. = 0.546, 95% CI [0.148, 0.792], $F[19, 19] = 3.405$, $p = 0.005$; PZ 30° : ICC coef. = 0.438, 95% CI [0.006, 0.732], $F[19, 19] = 2.559$, $p = 0.023$; PZ 75° : ICC coef. = 0.431, 95% CI [0.000, 0.728], $F[19, 19] = 2.517$, $p = 0.025$; ZN -60° : ICC coef. = 0.396, 95% CI [0.000, 0.708], $F[19, 19] = 2.311$, $p = 0.038$; ZN -15° : ICC coef. = 0.494, 95% CI [0.077, 0.763], $F[19, 19] = 2.950$, $p = 0.011$).

Correlations between illusion magnitudes

Attenuated (Table 4, upper triangle) and disattenuated (Table 4, lower triangle) correlation coefficients were reported both without and with Bonferroni

correction. Among 190 correlations, 55 attenuated and 81 disattenuated correlations resulted in $r > 0.3$. Interestingly, 97% (29 out of 30) and 100% of the attenuated and disattenuated intra-illusion correlations showed $r > 0.3$, respectively. In contrast, only 16% (26 out of 160) and 32% (51 out of 160) of the attenuated and disattenuated inter-illusion correlations showed $r > 0.3$, respectively (Figure 10). Hence, it seems that intra-illusion correlations were stronger than inter-illusion correlations, even though intrarater reliabilities were not always high.

Exploratory factor analysis (EFA)

We conducted an exploratory factor analysis with an oblique rotation. We identified four factors by scree plot inspection, which accounted for $\sim 61\%$ of the variability of the data (RF1: 17%, RF2: 17%, RF3: 15%, RF4: 12%). In the unrotated factor solution, the first factor explained about 22% of the total variance. Rotated factor loadings are presented in Table 5. The first factor was mainly composed of the PD and ZN conditions while the second factor mainly loaded on the HV and PD conditions. The third and fourth factors were respectively dominated by loadings from the ML and PZ conditions. Each illusion mainly loaded on one

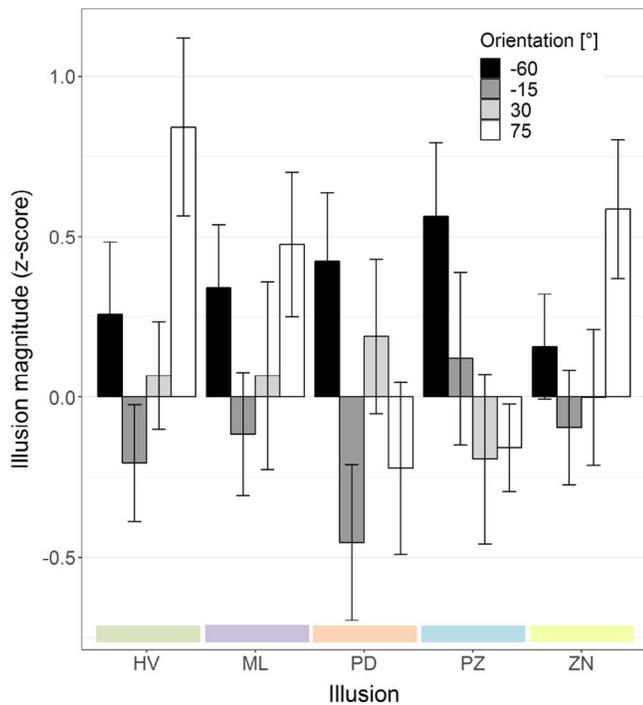


Figure 9. Standardized illusion magnitudes \pm SEM as a function of the orientation [°] for each illusion. The z-scores show to what extent the illusion magnitudes deviate as a function of orientation for each illusion type.

factor, except for the PD conditions which highly loaded on two (or three) factors.

Data simulation

As in Experiment 2, we simulated intra-illusion correlation coefficients from the effect size observed in

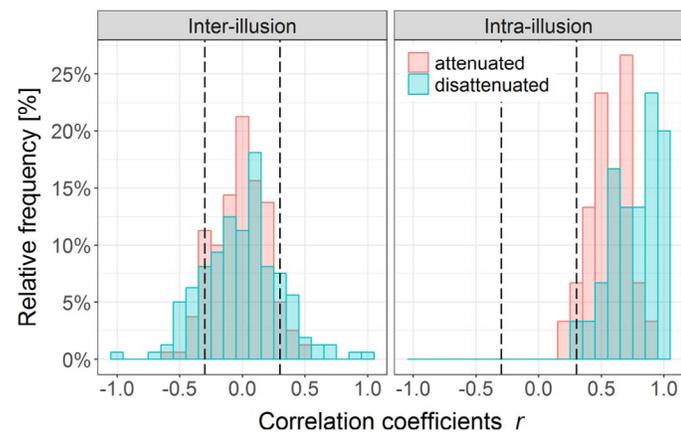


Figure 10. Relative frequency (in percentage) of attenuated (red) and disattenuated (turquoise) correlation coefficients r for inter- (left panel) and intra-illusion (right panel) correlations in Experiment 3. According to Cohen’s convention (Cohen, 1988), $r = \pm 0.3$ (vertical dashed lines) is the lower limit corresponding to a medium effect size.

Experiment 1, while inter-illusion correlation coefficients were computed from simulated data with $M = 0$ and $SD = 1$. The probability of observing less than 97% of intra-illusion and more than 16% of inter-illusion correlation coefficients across 10,000 simulations was 89%. Indeed, there were on average 20% of inter- and 86% of intra-illusion correlation coefficients larger than 0.3, suggesting that we may underestimate the true intra-illusion effect size.

A Welch t test between the observed inter- and intra-illusion correlation coefficients resulted in a strongly significant difference (two-tailed t test, $t[48.37] = 17.37, p < 0.001$), highlighting a difference between inter- and intra-illusion correlations despite the small sample size.

Discussion

Previous research did not find evidence for a common factor for visual illusions (Grzeczowski et al., 2017; Grzeczowski et al., 2018). Here, we systematically examined hyper-specificity of factors for visual illusions. We especially wondered whether these factors are as specific as in perceptual learning. To this end, we presented several variants of different visual illusions. We observed strong correlations between the different variants of each illusion (intra-illusion correlations) but only weak correlations between different illusions (inter-illusion correlations), which suggests that factors are illusion-specific.

In Experiment 1, we tested different variants of the Ebbinghaus illusion and found high correlations between the illusion magnitudes of the different variants. For example, illusion magnitudes significantly correlated between the STD (standard) and the FRT (fast rotating flankers) conditions (Figure 1). These results suggest one mechanism behind the Ebbinghaus illusion, corroborated by a factor analysis showing evidence for only one main factor. A similar result was previously found for the Müller-Lyer illusion (Coren, Gergus, Erlichman, & Hakstian, 1976). A factor analysis showed that 45 measures of different illusions were best represented by a two-factor model, with one factor mainly loading on several variants of the Müller-Lyer illusion.

In Experiment 2, we reanalyzed a dataset, in which 10 different illusions were tested with four luminance conditions. We found high intra-illusion correlations for different luminance conditions, but weak inter-illusion correlations. However, the BS illusion magnitudes strongly correlated with the HV illusion magnitudes. In fact, the HV illusion is a BS illusion rotated by 90°. An EFA identified nine main components which explained $\sim 90\%$ of the variance. Strikingly, each

	HV -60	HV -15	HV 30	HV 75	ML -60	ML -15	ML 30	ML 75	PD -60	PD -15	PD 30	PD 75	PZ -60	PZ -15	PZ 30	PZ 75	ZN -60	ZN -15	ZN 30	ZN 75
HV -60	.74	.43	.50	.71	-.13	.12	.05	-.35	.08	-.08	-.18	-.17	.07	.04	.15	.06	-.08	.07	-.12	-.04
HV -15	.55	.86	.61	.73	-.06	.05	-.25	-.31	-.35	-.31	-.09	-.28	-.19	-.13	-.15	.01	-.06	.15	.04	.22
HV 30	.72	.80	.66	.71	.10	.00	-.13	-.33	-.45	-.31	-.40	-.26	.02	.04	.21	.51	.03	.16	.05	-.05
HV 75	.85	.81	.90	.94	.13	.23	.06	-.29	-.22	-.30	-.28	-.38	.01	.16	.11	0.11	-.03	.10	-.06	.05
ML -60	-.17	-.07	.14	.14	.83	.72	.60	.52	-.18	-.17	-.30	-.29	.12	.27	.11	.01	.01	-.01	.07	-.08
ML -15	.15	.06	.00	.26	.85	.85	.79	.39	-.09	-.11	-.17	-.24	.25	.38	.24	-.02	.23	.18	.18	-.02
ML 30	.06	-.28	-.16	.06	.69	.89	.91	.55	-.01	-.06	-.10	-.25	.11	.15	.17	-.11	.33	.16	.08	-.08
ML 75	-.46	-.37	-.45	-.34	.63	.47	.64	.81	.02	.15	.13	.08	-.11	-.06	.03	-.37	.33	.23	.29	.41
PD -60	.10	-.40	-.59	-.25	-.21	-.11	-.01	.03	.87	.45	.55	.58	.21	.11	-.01	-.26	.23	.08	.18	-.04
PD -15	-.12	-.45	-.51	-.42	-.25	-.16	-.08	.23	.65	.55	.71	.70	-.13	-.33	-.03	.00	.47	.16	.19	.01
PD 30	-.22	-.10	-.52	-.31	-.36	-.20	-.12	.16	.64	1	.85	.78	-.26	-.35	-.43	-.30	.40	.31	.32	.25
PD 75	-.21	-.33	-.34	-.41	-.34	-.28	-.28	.09	.66	1	.90	.88	.04	-.21	-.22	.05	.41	.20	.34	.06
PZ -60	.09	-.25	.04	.01	.16	.32	.14	-.14	.26	-.21	-.33	.05	.73	.54	.52	.47	.04	-.03	.04	-.22
PZ -15	.06	-.16	.06	.20	.35	.48	.18	-.08	.14	-.53	-.45	-.27	.74	.72	.50	.19	-.16	-.24	-.15	.04
PZ 30	.27	-.24	.38	.18	.18	.40	.27	.05	-.02	-.06	-.70	-.36	.93	.89	.44	.57	.03	-.04	.07	-.25
PZ 75	.10	.02	.95	.18	.02	-.03	-.17	-.63	-.42	.00	-.49	.09	.83	.35	1	.43	-.06	-.18	-.07	-.56
ZN -60	-.14	-.10	.06	-.04	.02	.40	.56	.58	.39	1	.69	.69	.07	-.30	.06	-.13	.40	.74	.69	.31
ZN -15	.12	.23	.27	.14	-.02	.27	.24	.37	.13	.31	.47	.31	-.04	-.41	-.09	-.40	1	.49	.92	.44
ZN 30	-.15	.05	.06	-.07	.09	.22	.10	.36	.21	.29	.39	.41	.05	-.20	.12	-.12	1	1	.79	.31
ZN 75	-.05	.28	-.07	.07	-.11	-.02	-.11	.55	-.05	.01	.33	.08	-.31	.06	-.46	-.1	.59	.76	.43	.68

Table 4. Diagonal (in gray): Intrarater reliability expressed as intraclass correlation coefficients (ICC_{3,1}) for each condition. Upper triangle: Attenuated correlation coefficients between each pair of conditions (Bravais–Pearson’s *r*). Lower triangle: Disattenuated correlation coefficients between each pair of conditions (Bravais–Pearson’s *r*). Italics and bold font indicate significant results without and with Bonferroni correction, respectively. The color scale from blue to red reflects effect sizes from *r* = −1 to *r* = 1 (white corresponds to *r* = 0).

component mainly loaded on one illusion except for the BS and HV illusions, which highly loaded on the same component.

The strong correlations between the four luminance conditions indicate that, contrary to a widely held belief (e.g., Livingstone & Hubel, 1987), visual illusions do not break down under conditions of isoluminance and are therefore not primarily processed by the magnocellular system (also after controlling for subjective isoluminance, cf. Hamburger et al., 2007).

In Experiment 3, five visual illusions were tested with four orientations. Again, intra-illusion correlations were mostly high, while inter-illusion correlations were mostly weak. Hence, we found almost no significant inter-illusion correlations in both Experiments 2 and 3, as observed in Grzeczowski et al. (2017). Furthermore, in all three experiments, a dimensionality reduction technique (EFA) showed that more or less

each illusion makes up its own factor. The PD illusion loaded on only one factor in Experiment 2 but was related to the HV, ML, and ZN illusions in Experiment 3, which may be due to chance.

In Experiment 2, we found no link between the EH and DB illusions, which are often proposed to rely on the same mechanism (Coren et al., 1976; Girgus, Coren, & Agdern, 1972; Roberts, Harris, & Yates, 2005). Likewise, there was no link between the JD and ML illusions. An explanation may be that the two segments compared in the JD illusion are collinear while on top of each other in the ML illusion. The absence of significant correlations between the EH and DB illusions is, however, more puzzling. However, it is important to note that we are not making conclusions on specific comparisons in this study such as the comparison between the EH and DB illusion. For such a purpose, we do not have sufficient power. The

	RF1	RF2	RF3	RF4
HV -60	0.12	0.59	-0.14	0.12
HV -15	0.09	0.86	-0.19	-0.27
HV 30	0.17	0.90	-0.15	0.21
HV 75	0.12	0.87	0.01	0.03
ML -60	-0.03	-0.07	0.76	0.06
ML -15	0.21	0.05	0.82	0.18
ML 30	0.16	-0.16	0.82	0.08
ML 75	0.21	-0.40	0.70	-0.26
PD -60	0.33	-0.37	-0.13	0.16
PD -15	0.49	-0.32	-0.20	0.08
PD 30	0.53	-0.28	-0.27	-0.24
PD 75	0.57	-0.33	-0.37	0.14
PZ -60	0.12	-0.05	0.15	0.71
PZ -15	-0.14	-0.04	0.29	0.44
PZ 30	0.11	0.08	0.21	0.72
PZ 75	0.05	0.22	-0.21	0.77
ZN -60	0.90	0.11	0.25	0.07
ZN -15	0.90	0.38	0.21	-0.12
ZN 30	0.85	0.20	0.21	0.02
ZN 75	0.29	0.14	0.15	-0.47

Table 5. Rotated factor loadings from an exploratory factor analysis (EFA) after promax rotation for all illusions and orientations. A color scale from blue (negative loadings) to red (positive loadings) is shown.

purpose of this study was to show that *in general* variants of one illusion type show strong correlations but inter-illusion correlations are weak(er). It is the data as a whole which is important and not single, specific findings.

Perceptual learning is usually very specific to the trained stimuli. For example, when a Vernier offset discrimination task is trained, performance improves. However, learning transfers only when the Vernier is rotated up to 10° (Spang et al., 2010) but not when rotated by 90° (Ball & Sekuler, 1987; Fahle & Morgan, 1996; Schoups et al., 1995). However, we found high correlations when illusions were rotated by -60°, -15°, 30°, and 75° (Experiment 3). In Experiment 2, the BS illusion strongly correlated with the HV illusion, which is just the BS illusion rotated by 90°. Hence, factors for illusions are not as specific as one may have expected from perceptual learning, where learning is usually retinotopic and orientation-specific. Therefore, perceptual learning does not seem to shape the factors underlying visual illusions, as long as everyday perceptual learning is as specific as under laboratory conditions (just to mention, transfer of perceptual

learning in the laboratory can occur; e.g., Aberg, Tartaglia, & Herzog, 2009).

In Experiment 1, we expected the control condition (CON) to show a null effect, which however was not the case. We suggest that our procedure induced a bias. The adjustable target was always surrounded by larger flankers compared to the reference target. Since participants overadjusted the size of the target in each noncontrol condition, we suggest that it induced a small but significant bias in the control condition as well. We hypothesize that this bias would disappear if the adjustable target was randomly surrounded by large or small flankers. Illusion magnitudes varied as a function of the size of the flankers, as highlighted by the FBIG and FSMA conditions, where we observed larger and smaller effects compared to the STD condition, respectively. The size of the flankers can indeed influence illusion magnitudes (e.g., Roberts et al., 2005).

We reported significance both without correction and with the very conservative Bonferroni correction, since we sometimes aimed for null results and sometimes not. Therefore, we considered the most extreme option in each case. In Experiment 1, correlations were significant even when the alpha level was corrected for family-wise errors. In Experiments 2 and 3, inter-illusion correlations were mostly nonsignificant even though we did not correct the alpha level for multiple comparisons and intra-illusion correlations were mostly significant both without and with Bonferroni correction.

Intrarater reliabilities were mostly significant but moderate, reflecting a nonnegligible within-participant variation that can be explained by the low number of repetitions per condition and by measurement errors. This may put an upper limit on the observed pairwise correlations, i.e., it may lead to underestimated correlations (Bosten et al., 2017; Mollon, Bosten, Peterzell, & Webster, 2017). To account for this, we computed disattenuated correlations. Similar patterns of correlations were observed between attenuated and disattenuated correlations (see Tables 1, 2, and 4). Hence, our null results are unlikely to be type II errors. Moreover, a Bayesian approach previously showed that the null hypothesis was more likely than the alternative hypothesis for inter-illusion correlations with significant intrarater reliabilities (Grzeczowski et al., 2017). Intrarater reliabilities were higher in Experiment 2 than in Experiment 1, likely because we had four instead of two trials per condition, respectively. We like to mention that between-condition variations may be higher than within-condition variations in Experiments 1 and 3, since both trials of a condition were always presented sequentially, i.e., one after the other, which may inflate intrarater reliabilities.

Power may be an issue, especially in Experiments 2 and 3, in which we had small sample sizes. Data simulation in Experiment 3 resulted in a high likelihood

of observing a smaller percentage of intra-illusion correlation coefficients and a larger percentage of inter-illusion correlation coefficients larger than 0.3 compared to what we actually observed. On the contrary, this likelihood was null in Experiment 2. Indeed, the proportion of simulated inter-illusion correlations with $r > 0.3$ never reached the 27% observed in the data, which suggests that the true inter-illusion effect size is nonnull. The average proportion of simulated intra-illusion coefficients larger than 0.3 was 86%, while all observed intra-illusion correlations showed $r > 0.3$, indicating that we may also underestimate the true intra-illusion effect size.

Importantly, there are strong differences between the inter- and intra-illusion correlation coefficients in Experiments 2 and 3, supporting our claim. In addition, we like to mention that we observed significant correlation coefficients both without and with the conservative Bonferroni correction.

Often illusions are implicitly or explicitly classified according to certain criteria, such as their geometric-spatial features, and it is assumed that a common mechanism is in operation (e.g., Coren et al., 1976; Ninio, 2014; Piaget, 1961). Our study cannot address the question of which mechanisms are at work but challenges the notion that there are common explanations for classes of illusions (for a critical review see Hamburger, 2016).

The Ebbinghaus illusion magnitude was shown to negatively correlate with V1 cortex size (de Haas, Kanai, Jalkanen, & Rees, 2012; Schwarzkopf & Rees, 2013; Schwarzkopf, Song, & Rees, 2011). However, whether this conclusion extends to other visual illusions may be questioned given the large individual differences we found. Our experiments tested a battery of illusions rather than investigating the mechanisms of one illusion as is common practice (but see Coren et al., 1976; Thurstone, 1944). The factor structure is not only sparse for illusions but also for vision in general. For example, a distinct factor structure was found for contrast: sensitivities correlated between 0.2 and 0.4 $c/^\circ$, between 0.4 and 1.2 $c/^\circ$ and between 1.2 to 3 $c/^\circ$, but correlations were weak between these different ranges (Peterzell, 2016; Peterzell, Scheffrin, Tregear, & Werner, 2000). Similarly, Emery and colleagues (2017a, 2017b) observed several small factors underlying individual differences in hue scaling. In addition, Bosten and Mollon (2010) found “no noteworthy general trait of susceptibility” (p. 1663) for contrast perception. When comparing different spatial tasks, such as bisection discrimination and Vernier offset discrimination, only low correlations were found (Cappe et al., 2014). Likewise, there was little evidence for a common factor for oculomotor tasks (Bargary et al., 2017) and for binocular rivalry and other bistable paradigms (Brascamp et al., 2018; Cao et al., 2018; Wexler, 2005).

Chamberlain, Van der Hallen, Huygelier, Van de Cruys, and Wagemans (2017) also showed poor evidence for a common factor for local and global visual processing. In addition, the effects of priors in perceptual tasks seem not to follow a single mechanism (Tulver, Aru, Rutiku, & Bachmann, 2019). Hence, these studies found very specific factors, similar to the very specific factors we found for visual illusions, and are rather arguing against a general factor for vision as proposed previously (e.g., Halpern, Andrews, & Purves, 1999). It appears that even studies that had a narrowly defined hypothesis by including several tasks that tap into a specific functional ability or theoretical construct of perception have often not succeeded in finding evidence to support the existence of a stable factor in perception (Tulver, 2019).

Surprisingly, whereas most inter-illusion magnitudes are only weakly correlated, there may be links between certain visual illusions. For example, Grzeczowski and colleagues (2017, 2018) found significant correlations between the Ebbinghaus and Ponzo illusions, which we could not, however, reproduce here (Table 2, triangle) and was similarly not found by Schwarzkopf and colleagues (2011). However, the power was much higher in the studies by Grzeczowski and colleagues (2017, 2018) than here. In addition, they also found that the Ponzo illusion magnitude correlated with cognitive disorganization. However, this was the only illusion that correlated with a personality trait. Given the large number of statistical comparisons, it may be that these few significant correlations are false positives. Alternatively, it may be that within an ocean of weak correlations there are some strong singular correlations, such as between the Ponzo illusion and cognitive organization or between two rather unrelated illusions. Whereas these links may be surprising from a vision science perspective, they may be less surprising from a genetic point of view. As a metaphor, a hypothetical gene may code for a protein, which plays an important role in the visual cortex and, say, in the liver. Variability in the gene may cause significant correlations between visual and liver functions, which may appear bizarre as long as one expects that visual functions rather than visual and liver functions go together. In this line, Frenzel et al. (2012) found correlations between auditory and touch perception as both sensory systems rely on mechanoreceptors encoded by the same gene.

In summary, common factors are ubiquitous. However, there is little evidence for a unique common factor for vision (e.g., Bosten & Mollon, 2010; Cappe et al., 2014; Coren, Girgus, & Day, 1973; Grzeczowski et al., 2017; Peterzell, Chang, & Teller, 2000; Peterzell & Teller, 1996; Peterzell, Werner, & Kaplan, 1993, 1995; Webster & MacLeod, 1988). Here, we showed that the factor structure underlying visual illusions is sparse.

However, factors are not as specific as it may be expected from perceptual learning. It seems that most illusions make up their own factor.

Keywords: factors, illusions, individuality, perception

Acknowledgments

We thank Marc Reppow for technical support and helpful comments. This work was supported by the project “Basics of visual processing: From elements to figures” (Project number: 320030_176153 / 1) of the Swiss National Science Foundation (SNSF) and by a National Centre of Competence in Research (NCCR Synapsy) grant from the SNSF (51NF40-185897).

Commercial relationships: none.

Corresponding author: Aline F. Cretienoud.

Email: aline.cretienoud@epfl.ch.

Address: Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

References

- Aberg, K. C., Tartaglia, E. M., & Herzog, M. H. (2009). Perceptual learning with chevrons requires a minimal number of trials, transfers to untrained directions, but does not require sleep. *Vision Research*, *49*(16), 2087–2094, <https://doi.org/10.1016/j.visres.2009.05.020>
- Axelrod, V., Schwarzkopf, D. S., Gilaie-Dotan, S., & Rees, G. (2017). Perceptual similarity and the neural correlates of geometrical illusions in human brain structure. *Scientific Reports*, *7*:39968, <https://doi.org/10.1038/srep39968>.
- Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research*, *27*(6), 953–965.
- Bargary, G., Bosten, J. M., Goodbourn, P. T., Lawrence-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). Individual differences in human eye movements: An oculomotor signature? *Vision Research*, *141*, 157–169, <https://doi.org/10.1016/j.visres.2017.03.001>.
- Bosten, J. M., Goodbourn, P. T., Bargary, G., Verhallen, R. J., Lawrence-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). An exploratory factor analysis of visual performance in a large population. *Vision Research*, *141*, 303–316, <https://doi.org/10.1016/j.visres.2017.02.005>.
- Bosten, J. M., & Mollon, J. D. (2010). Is there a general trait of susceptibility to simultaneous contrast? *Vision Research*, *50*(17), 1656–1664, <https://doi.org/10.1016/j.visres.2010.05.012>.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.
- Brascamp, J. W., Becker, M. W., & Hambrick, D. Z. (2018). Revisiting individual differences in the time course of binocular rivalry. *Journal of Vision*, *18*(7): 3, 1–20, <https://doi.org/10.1167/18.7.3>. [PubMed] [Article]
- Cao, T., Wang, L., Sun, Z., Engel, S. A., & He, S. (2018). The independent and shared mechanisms of intrinsic brain dynamics: Insights from bistable perception. *Frontiers in Psychology*, *9*, 589, <https://doi.org/10.3389/fpsyg.2018.00589>.
- Cappe, C., Clarke, A., Mohr, C., & Herzog, M. H. (2014). Is there a common factor for vision? *Journal of Vision*, *14*(8):4, 1–11, <https://doi.org/10.1167/14.8.4>. [PubMed] [Article]
- Chamberlain, R., Van der Hallen, R., Huygelier, H., Van de Cruys, S., & Wagemans, J. (2017). Local-global processing bias is not a unitary individual difference in visual processing. *Vision Research*, *141*, 247–257, <https://doi.org/10.1016/j.visres.2017.01.008>.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Coren, S., Girgus, J. S., & Day, R. H. (1973, February 2). Visual spatial illusions: Many explanations. *Science*, *179*(4072), 503–504.
- Coren, S., Girgus, J. S., Erlichman, H., & Hakstian, A. R. (1976). An empirical taxonomy of visual illusions. *Perception & Psychophysics*, *20*(2), 129–137, <https://doi.org/10.3758/BF03199444>.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Education*, *10*(7), 86–99, <https://doi.org/10.1.1.110.9154>.
- de Haas, B., Kanai, R., Jalkanen, L., & Rees, G. (2012). Grey matter volume in early human visual cortex predicts proneness to the sound-induced flash illusion. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1749), 4955–4961, <https://doi.org/10.1098/rspb.2012.2132>.
- Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *The Journal of Physiology*, *357*(1), 241–265.
- Emery, K. J., Volbrecht, V. J., Peterzell, D. H., &

- Webster, M. A. (2017a). Variations in normal color vision. VI. Factors underlying individual differences in hue scaling and their implications for models of color appearance. *Vision Research*, *141*, 51–65, <https://doi.org/10.1016/j.visres.2016.12.006>.
- Emery, K. J., Volbrecht, V. J., Peterzell, D. H., & Webster, M. A. (2017b). Variations in normal color vision. VII. Relationships between color naming and hue scaling. *Vision Research*, *141*, 66–75, <https://doi.org/10.1016/j.visres.2016.12.007>.
- Fahle, M., & Morgan, M. (1996). No transfer of perceptual learning between similar stimuli in the same retinal position. *Current Biology*, *6*(3), 292–297, [https://doi.org/10.1016/S0960-9822\(02\)00479-7](https://doi.org/10.1016/S0960-9822(02)00479-7).
- Faivre, N., Filevich, E., Solovey, G., Kühn, S., & Blanke, O. (2018). Behavioural, modeling, and electrophysiological evidence for supramodality in human metacognition. *Journal of Neuroscience*, *38*(2), 263–277, <https://doi.org/10.1523/JNEUROSCI.0322-17.2017>.
- Frenzel, H., Bohlender, J., Pinsker, K., Wohlleben, B., Tank, J., Lechner, S. G., . . . Lewin, G. R. (2012). A genetic basis for mechanosensory traits in humans. *PLoS Biology*, *10*(5): e1001318, <https://doi.org/10.1371/journal.pbio.1001318>.
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78, <https://doi.org/10.1016/j.paid.2016.06.069>.
- Girgus, J. S., Coren, S., & Agdern, M. (1972). The interrelationship between the Ebbinghaus and Delboeuf illusions. *Journal of Experimental Psychology*, *95*(2), 453–455, <https://doi.org/10.1037/h0033606>.
- Goodbourn, P. T., Bosten, J. M., Hogg, R. E., Bargary, G., Lawrance-Owen, A. J., & Mollon, J. D. (2012). Do different “magnocellular tasks” probe the same neural substrate? *Proceedings of the Royal Society B: Biological Sciences*, *279*(1745), 4263–4271, <https://doi.org/10.1098/rspb.2012.1430>.
- Grzeczowski, L., Clarke, A. M., Francis, G., Mast, F. W., & Herzog, M. H. (2017). About individual differences in vision. *Vision Research*, *141*, 282–292, <https://doi.org/10.1016/j.visres.2016.10.006>.
- Grzeczowski, L., Cretenoud, A., Herzog, M. H., & Mast, F. W. (2017). Perceptual learning is specific beyond vision and decision making. *Journal of Vision*, *17*(6):6, 1–11, <https://doi.org/10.1167/17.6.6>. [PubMed] [Article]
- Grzeczowski, L., Roinishvili, M., Chkonia, E., Brand, A., Mast, F. W., Herzog, M. H., & Shaqiri, A. (2018). Is the perception of illusions abnormal in schizophrenia? *Psychiatry Research*, *270*, 929–939, <https://doi.org/10.1016/j.psychres.2018.10.063>.
- Halpern, S. D., Andrews, T. J., & Purves, D. (1999). Interindividual variation in human visual performance. *Journal of Cognitive Neuroscience*, *11*(5), 521–534.
- Hamburger, K. (2016). Visual illusions based on processes: New classification system needed. *Perception*, *45*(5), 588–595, <https://doi.org/10.1177/0301006616629038>.
- Hamburger, K., & Hansen, T. (2010). Analysis of individual variations in the classical horizontal-vertical illusion. *Attention, Perception & Psychophysics*, *72*(4), 1045–1052, <https://doi.org/10.3758/APP.72.4.1045>.
- Hamburger, K., Hansen, T., & Gegenfurtner, K. R. (2007). Geometric-optical illusions at isoluminance. *Vision Research*, *47*(26), 3276–3285, <https://doi.org/10.1016/j.visres.2007.09.004>.
- Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients, *58*(1), 78–79, <https://doi.org/10.1037/0003-066X.58.1.78>.
- Iglewicz, B., & Hoaglin, D. (1993). How to detect and handle outliers. In E. F. Mykytka (Ed.), *The ASQC basic references in quality control: Statistical techniques* (Vol. 16). Milwaukee, WI: Asq Press.
- Kaliuzhna, M., Stein, T., Rusch, T., Sekutowicz, M., Sterzer, P., & Seymour, K. J. (2018). No evidence for abnormal priors in early vision in schizophrenia. *Schizophrenia Research*, *210*, 245–254, <https://doi.org/10.1016/j.schres.2018.12.027>.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163, <https://doi.org/10.1016/j.jcm.2016.02.012>.
- Krauskopf, J., Williams, D. R., & Heeley, D. W. (1982). Cardinal directions of color space. *Vision Research*, *22*(9), 1123–1131.
- Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, *7*(11), 3416–3468.
- Mollon, J. D., Bosten, J. M., Peterzell, D. H., & Webster, M. A. (2017). Individual differences in visual science: What can be learned and what is good experimental practice? *Vision Research*, *141*, 4–15, <https://doi.org/10.1016/j.visres.2017.11.001>.
- Ninio, J. (2014). Geometrical illusions are not always where you think they are: A review of some classical and less classical illusions, and ways to

- describe them. *Frontiers in Human Neuroscience*, 8, 856, <https://doi.org/10.3389/fnhum.2014.00856>.
- Osborne, J. W. (2003). Effect sizes and the disattenuation of correlation and regression coefficients: Lessons from educational psychology. *Practical Assessment, Research and Evaluation*, 8(11), 5–11.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Peterzell, D. H. (2016). Discovering sensory processes using individual differences: A review and factor analytic manifesto. *Electronic Imaging*, 2016(16), 1–11, <https://doi.org/10.2352/ISSN.2470-1173.2016.16HVEI-112>.
- Peterzell, D. H., Chang, S. K., & Teller, D. Y. (2000). Spatial frequency tuned covariance channels for red-green and luminance- modulated gratings: Psychophysical data from human infants. *Vision Research*, 40(4), 431–444, [https://doi.org/10.1016/S0042-6989\(99\)00188-1](https://doi.org/10.1016/S0042-6989(99)00188-1).
- Peterzell, D. H., Scheffrin, B. E., Tregear, S. J., & Werner, J. S. (2000). Spatial frequency tuned covariance channels underlying scotopic contrast sensitivity. In *Vision science and its applications: OSA Technical Digest*. Washington, DC: Optical Society of America.
- Peterzell, D. H., Serrano-Pedraza, I., Widdall, M., & Read, J. C. A. (2017). Thresholds for sine-wave corrugations defined by binocular disparity in random dot stereograms: Factor analysis of individual differences reveals two stereoscopic mechanisms tuned for spatial frequency. *Vision Research*, 141, 127–135, <https://doi.org/10.1016/j.visres.2017.11.002>.
- Peterzell, D. H., & Teller, D. Y. (1996). Individual differences in contrast sensitivity functions: The lowest spatial frequency channels. *Vision Research*, 36(19), 3077–3085, [https://doi.org/10.1016/0042-6989\(96\)00061-2](https://doi.org/10.1016/0042-6989(96)00061-2).
- Peterzell, D. H., Werner, J. S., & Kaplan, P. S. (1993). Individual differences in contrast sensitivity functions: The first four months of life in humans. *Vision Research*, 33(3), 381–396, [https://doi.org/10.1016/0042-6989\(93\)90093-C](https://doi.org/10.1016/0042-6989(93)90093-C).
- Peterzell, D. H., Werner, J. S., & Kaplan, P. S. (1995). Individual differences in contrast sensitivity functions: Longitudinal study of 4-, 6- and 8-month-old human infants. *Vision Research*, 35(7), 961–979, [https://doi.org/10.1016/0042-6989\(94\)00117-5](https://doi.org/10.1016/0042-6989(94)00117-5).
- Piaget, J. (1961). *Les mécanismes perceptifs, modèles probabilistes, analyse génétique, relations avec l'intelligence* [The mechanisms of perception] (2nd ed.). Paris, France: Presses universitaires de France.
- Preacher, K. J., Zhang, G., Kim, C., & Mels, G. (2013). Choosing the optimal number of factors in exploratory factor analysis: A model selection perspective. *Multivariate Behavioral Research*, 48(1), 28–56, <https://doi.org/10.1080/00273171.2012.710386>.
- R Core Team (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>
- Roberts, B., Harris, M. G., & Yates, T. A. (2005). The roles of inducer size and distance in the Ebbinghaus illusion (Titchener circles). *Perception*, 34(7), 847–856, <https://doi.org/10.1068/p5273>.
- Schoups, A. A., Vogels, R., & Orban, G. A. (1995). Human perceptual learning in identifying the oblique orientation: Retinotopy, orientation specificity and monocularly. *The Journal of Physiology*, 483(3), 797–810, <https://doi.org/10.1113/jphysiol.1995.sp020623>.
- Schwarzkopf, D. S., & Rees, G. (2013). Subjective size perception depends on central visual cortical magnification in human V1. *PLoS One*, 8(3): e60550, <https://doi.org/10.1371/journal.pone.0060550>.
- Schwarzkopf, D. S., Song, C., & Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nature Neuroscience*, 14(1), 28–30, <https://doi.org/10.1038/nn.2706>.
- Shaqiri, A., Pilz, K. S., Cretenoud, A. F., Neumann, K., Clarke, A., Kunchulia, M., & Herzog, M. H. (2019). No evidence for a common factor underlying visual abilities in healthy older people. *Developmental Psychology*, 55(8), 1775–1787, <https://doi.org/10.1037/dev0000740>.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428.
- Simpson, W. A., & McFadden, S. M. (2005). Spatial frequency channels derived from individual differences. *Vision Research*, 45(21), 2723–2727, <https://doi.org/10.1016/j.visres.2005.01.015>.
- Spang, K., Grimsen, C., Herzog, M. H., & Fahle, M. (2010). Orientation specificity of learning vernier discriminations. *Vision Research*, 50(4), 479–485, <https://doi.org/10.1016/j.visres.2009.12.008>.
- Spearman, C. (1904a). “General Intelligence” objectively determined and measured. *The American Journal of Psychology*, 15(2), 201–292.
- Spearman, C. (1904b). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101, <https://doi.org/10.1037/h0065390>.

- Thurstone, L. L. (1944). *A factorial study of perception*. Chicago, IL: The University of Chicago Press.
- Tulver, K. (2019). The factorial structure of individual differences in visual perception. *Consciousness and Cognition*, 73, 167–177, <https://doi.org/10.1016/j.concog.2019.102762>.
- Tulver, K., Aru, J., Rutiku, R., & Bachmann, T. (2019). Individual differences in the effects of priors on perception: A multi-paradigm approach. *Cognition*, 187(3), 167–177, <https://doi.org/10.1016/j.cognition.2019.03.008>.
- Verhallen, R. J., Bosten, J. M., Goodbourn, P. T., Lawrance-owen, A. J., Bargary, G., & Mollon, J. D. (2017). General and specific factors in the processing of faces. *Vision Research*, 141, 217–227, <https://doi.org/10.1016/j.visres.2016.12.014>.
- Wang, L. L. (2010). Disattenuation of correlations due to fallible measurement. *Newborn and Infant Nursing Reviews*, 10(1), 60–65, <https://doi.org/10.1053/j.nainr.2009.12.013>.
- Ward, J., Rothen, N., Chang, A., & Kanai, R. (2017). The structure of inter-individual differences in visual ability: Evidence from the general population and synaesthesia. *Vision Research*, 141, 293–302, <https://doi.org/10.1016/j.visres.2016.06.009>.
- Webster, M. A., & MacLeod, D. I. A. (1988). Factors underlying individual differences in the color matches of normal observers. *Journal of the Optical Society of America A*, 5(10), 1722–1735, <https://doi.org/10.1364/JOSAA.5.001722>.
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children—Fourth Edition (WISC-IV)*. San Antonio, TX: The Psychological Corporation.
- Wexler, M. (2005). Anticipating the three-dimensional consequences of eye movements. *Proceedings of the National Academy of Sciences, USA*, 102(4), 1246–1251.

Factors underlying visual illusions are illusion-specific but not feature-specific

Aline F. Cretenoud, Harun Karimpur, Lukasz Grzeczowski, Gregory Francis, Kai Hamburger, & Michael H. Herzog

Supplementary File S1

The illusions used in Experiment 3 are here described in a 0° orientation configuration. In the horizontal-vertical illusion (HV), the length of the horizontal line was 10.4°. This length was taken as a reference to adjust the length of the vertical line whose initial length was randomized between 0° and 17°. The reference line was horizontally centered in the middle of the screen and vertically displaced by 5.2° to the bottom. The adjustable line was always touching the center of the reference line, no matter of its size.

In the Müller-Lyer illusion (ML), the reference line was 8° long, always presented with inward-pointing arrows. The initial length of the adjustable line with outward-pointing arrows varied randomly between 0° and 14.5°. The center of the reference and adjustable lines were placed at a distance of 3.4° to the left and to the right (respectively) of the horizontal screen midline, and 2.1° above and below (respectively) the vertical screen midline. The arrows were 1.5° in length, and inclined by 45° compared to the reference and adjustable lines.

In the Poggendorff illusion (PD), two vertical parallel lines, 16.6° long, were horizontally separated by 4° and centrally displayed. The left part of the interrupted diagonal (which was fixed) was 10.4° long, and tilted by 37° from the center of the left parallel line. The adjustable, right part of the interrupted diagonal was randomly positioned at the beginning of each trial between the top and the center of the right parallel line. In order to avoid that participants use landmarks to facilitate the adjustment, the length of the right, adjustable part of the interrupted diagonal was randomized at the beginning of each trial between 3.1° and 3.7°.

In the Ponzo illusion (PZ), the reference stimulus was the 4.5° long, horizontal, lower line. The adjustable line was the horizontal, upper line whose initial length randomly varied between 0° and 12° across trials. Both the reference and adjustable lines were horizontally centered in the middle of the screen, and at a distance of 7.2° from the vertical screen midline. The ends of the diagonal lines were placed at the same vertical height as the reference and adjustable lines. The two upper and two lower diagonal ends were respectively separated by 6° and 18°.

In the Zöllner illusion (ZN), the five vertical main streams were 16.6° in length and 2° apart from each other. The orientation of these streams was adjustable. They could be bent by 6.9° maximum in both

directions, but two neighbor streams never bent in the same direction. The initial orientation was randomized between -6.9° and 6.9° . Eight tilted lines were intersecting each main stream. They were 2.8° long and tilted by $\pm 25^\circ$ compared to the main streams. Each main stream was vertically and randomly shifted by -0.5° to 0.5° to avoid comparing the extremities of these lines. Similarly, the tilted lines were not perfectly centered on the streams. A random proportion (between 30 and 70 %) of each tilted line's length was depicted on the left of the corresponding main stream, the rest on the right part.

Appendix B

Cretenoud, A. F., Grzeczowski, L., Bertamini, M., & Herzog, M. H. (2020). Individual differences in the Müller-Lyer and Ponzo illusions are stable across different contexts. *Journal of Vision*, 20(6):4, 1–14, <https://doi.org/10.1167/jov.20.6.4>.

Copyright holder: © ARVO

Published under CC-BY-NC-ND 4.0

Individual differences in the Müller-Lyer and Ponzo illusions are stable across different contexts

Aline F. Cretenoud

Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland



Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
Allgemeine und Experimentelle Psychologie,
Department Psychologie,

Lukasz Grzeczowski

Ludwig-Maximilians-Universität München, Germany



Department of Psychological Sciences, University of Liverpool, UK
Department of General Psychology, University of Padova, Italy

Marco Bertamini

Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland



Michael H. Herzog



Vision scientists have attempted to classify visual illusions according to certain aspects, such as brightness or spatial features. For example, Piaget proposed that visual illusion magnitudes either decrease or increase with age. Subsequently, it was suggested that illusions are segregated according to their context: real-world contexts enhance and abstract contexts inhibit illusion magnitudes with age. We tested the effects of context on the Müller-Lyer and Ponzo illusions with a standard condition (no additional context), a line-drawing perspective condition, and a real-world perspective condition. A mixed-effects model analysis, based on data from 76 observers with ages ranging from 6 to 66 years, did not reveal any significant interaction between context and age. Although we found strong intra-illusion correlations for both illusions, we found only weak inter-illusion correlations, suggesting that the structure underlying these two spatial illusions includes several specific factors.

or brightness (Coren, Girgus, Erlichman, & Hakstian, 1976; Ninio, 2014; Piaget, 1961; Thurstone, 1944). Coren and Colleagues (1976) claimed that there are five classes of visual illusions, of which one class includes size-contrast illusions. Other taxonomies are based on whether illusions are innate or acquired. For example, Binet (1895) observed weaker Müller-Lyer illusion magnitudes in older compared with younger children and suggested that the magnitude of innate illusions decreases with age, whereas the magnitude of acquired illusions increases with age. In some cases, for example in a weight illusion, adults were found to be more susceptible than children (Dresslar, 1894).

This taxonomy was further elaborated by Piaget (1961, 1963), who developed a concept called centration. According to this concept, an object in the center of the visual field is overestimated in size compared with surrounding objects. Piaget suggested that some illusions are more systematically explored (i.e., multiple centrations) when children become older because older children make more eye movements as compared with younger children. For that reason, the susceptibility to some illusions (primary or type I illusions) is supposed to decrease with age. Simultaneously, the development of depth and perspective become apparent with more eye movements, giving rise to increasing susceptibility with age for other illusions (secondary or type II illusions). The Müller-Lyer and Ponzo illusions were

Introduction

There is an intuitive appeal in the idea that similar illusions rely on similar neural mechanisms. Consequently, taxonomies often classify illusions according to certain aspects, such as spatial features

Citation: Cretenoud, A. F., Grzeczowski, L., Bertamini, M., & Herzog, M. H. (2020). Individual differences in the Müller-Lyer and Ponzo illusions are stable across different contexts. *Journal of Vision*, 20(6):4, 1–14, <https://doi.org/10.1167/jov.20.6.4>.



first thought to belong to types I (Binet, 1895) and II (Leibowitz & Heisel, 1958; but see Pollack, 1964) respectively. For both illusions, however, there is a possible role of perspective depth features (e.g., Thiéry, 1896).

Piaget (1961) claimed that the difference between both illusion types lies in the complexity of the stimulus rather than in the illusion itself. For example, the Ponzo illusion magnitude was stronger when embedded in a photograph (with perspective cues) than when embedded in an abstract background (Leibowitz, Brislin, Perlmutter, & Hennessy, 1969). Wagner (1977) specifically interpreted the two-factor theory of Piaget as depending on the richness of context in which illusions are embedded and the participants' age. When testing children from 7 to 19 years old, Wagner observed increasing and decreasing illusion magnitudes with age for rich-context, real-world (photograph) and poor-context, abstract (geometrical) Ponzo illusions, respectively. Moreover, strong positive and negative correlations with a depth perception measure were found for both contexts, respectively. These results support the taxonomy of visual illusions proposed by Piaget, as long as the abstract illusion is considered as a primary and the real-world as a secondary type of illusion.

More recent studies about the effect of development on the perception of illusions reported inconsistent results (e.g., Duemmler, Franz, Jovanovic, & Schwarzer, 2008; Káldy & Kovács, 2003; Rival, Olivier, Ceyte, & Bard, 2004). For example, Doherty, Campbell, Tsuji, and Phillips (2010) asked participants to discriminate between a small target surrounded by small inducers and a larger target surrounded by large inducers. Children showed higher accuracy to this task compared to adults. However, Hanisch, Konczak, and Dohle (2001) reported that 5-year-old children and adults are susceptible to the Ebbinghaus illusion to the same extent.

In addition to age, other factors contribute to individual differences in the perception of visual illusions, such as physiological (e.g., Pollack & Silvar, 1967; Silvar & Pollack, 1967) or cultural (e.g., Brislin & Keating, 1976; Leibowitz & Pick, 1972) factors. For example, the “ecological” or “carpentered world” hypothesis posits that children growing up in an urban environment experience more linear perspectives than children growing up in a rural environment, thereby enhancing their susceptibility to some illusions (Brislin, 1974; Segall, Campbell, & Herskovits, 1963; Stewart, 1973). Moreover, there is evidence showing that non-Western observers from rural areas are less susceptible to classic visual illusions as compared with Western observers. For example, Leibowitz and Pick (1972) reported the absence of the Ponzo effect in Uganda villagers. Similarly, Himba participants from Northern Namibia were found to be less susceptible to

the Ebbinghaus illusion than Western participants (de Fockert, Davidoff, Fagot, Parron, & Goldstein, 2007), unless they had a few visits to a town. Interestingly, the Ebbinghaus illusion susceptibility drastically increased in a group of urbanized Himba, i.e., Himba that have moved to a city at an average age of 21 years (Caparos et al., 2012).

We recently found very little evidence for a general common factor in visual illusions, which challenges illusion taxonomies to some extent (Cretenoud et al., 2019; Grzeczowski, Clarke, Francis, Mast, & Herzog, 2017; Grzeczowski et al., 2018). For example, we measured the magnitudes of six different visual illusions and found only weak correlations between the illusion magnitudes (Grzeczowski et al., 2017), suggesting that an individual with a strong susceptibility to one illusion does not necessarily have a strong susceptibility to other illusions. This is even true when illusions fall into the same category (e.g., spatial or contrast illusions). Likewise, common factors for vision in general show rather weak loadings (smaller than 40% of explained variance: Cappe, Clarke, Mohr, & Herzog, 2014; see also Bosten et al., 2017; Shaqiri et al., 2019).

To investigate to which extent factors for visual illusions are specific, we recently tested several variants of the Ebbinghaus illusion, which differed in color, shape, or size, and found strong correlations among these variants. Similarly, we tested several illusions with different luminances and orientations and found strong correlations among different variants of an illusion but only weak correlations across different illusions. These results suggest that the factors underlying the susceptibility to visual illusions are illusion-specific but not feature-specific (Cretenoud et al., 2019).

Here, we tested whether there are factors for illusions based on illusion complexity and age, as proposed by Piaget, rather than on visual apparent features such as spatial or brightness features, as in classic taxonomies. If there is an interaction between age and the stimulus complexity, as suggested by Wagner (1977), we expect a decreasing illusion magnitude with age for primary illusions, i.e., poor-context conditions, and an increasing illusion magnitude with age for secondary illusions, i.e., rich-context conditions. Because Piaget (1961) considered intellectual development to continue until approximately 22 years, we specifically considered participants aged 22 years or younger.

To that aim, we tested the effect of different contextual cues on the susceptibility to the Müller-Lyer and Ponzo illusions. We used a classic version of both illusions (poor-context), the illusions embedded into a line-drawing perspective (moderate-context) and a real-world photograph (rich-context). Last, we wondered whether the weak inter-illusion correlations we previously observed replicated here in the case of the Müller-Lyer and Ponzo illusions.

Methods

Participants

Participants were 86 visitors of an open-door public event at the École Polytechnique Fédérale de Lausanne, Switzerland. After outlier removal (see “Data analysis” section), 76 participants were considered for further analysis (40 females, $M = 26$ years, range: 6–66 years) and 52 of them were younger than 23 years (26 females, $M = 12$ years, $SD = 4.2$ years). Adults signed informed consent and assent of the children was obtained, as was consent of their parents. There was no monetary reward. Procedures were conducted in accordance with the Declaration of Helsinki and were approved by the local ethics committee.

Apparatus

The stimuli were shown on a BenQ XL2420T monitor (1920×1080 pixels, 60 Hz; BenQ, Taipei, Taiwan) driven by a PC computer using Matlab (R2014b, 64 bits) and the Psychophysics toolbox (Brainard, 1997; Pelli, 1997; version 3.1, 64 bits). Before the experiments, the color look-up tables of the monitor were linearized and calibrated with a Minolta LS-100 luminance meter. Participants sat at a distance of approximately 60 cm from the monitor and were asked to minimize their head movements. The experiment was run in a silent room with artificial light conditions.

Stimuli

The “inward” and “outward” Müller-Lyer (ML) illusions were presented in different trials. Similarly, the Ponzo (PZ) illusion was presented either with the upper line only (“up”) or with the lower line only (“down”). Therefore, four configurations were tested. Each configuration was presented with three contexts (poor-context, moderate-context, and rich-context), making up 12 variants in total (Figure 1).

Each variant was presented on the left half of the screen (reference) and an adjustable line on the right half of the screen (vertical for ML variants, horizontal for PZ variants). Participants adjusted the line to match the reference line in length by moving the computer mouse on the orthogonal axis corresponding to the direction of the reference line. Both the reference and adjustable lines were presented in yellow (≈ 149 cd/m²) and additional lines were drawn in white when applicable (≈ 176 cd/m²). The black background luminance was approximately 1 cd/m². Participants pressed the left mouse button to validate each adjustment.

In the rich-context variants, real-world pictures were used. The ML real-world pictures were taken at the École Polytechnique Fédérale de Lausanne (EPFL, Switzerland) and the PZ real-world picture was taken in the countryside of Canton de Vaud (Switzerland) by the first author. Poor-context and moderate-context variants were drawn based on the rich-context backgrounds, so that perspective lines matched the perspective of the real-world pictures. The details about the metrics of the 12 variants are given in the Appendix.

Procedure

The experimenter first explained the task to the participants who adjusted each illusion variant once (warming up trials). The 12 variants were then tested twice in a sequential manner, i.e., one trial after the other, and without time restriction (24 trials in total). The order of presentation of the four configurations was randomized within a context but the three contexts were always presented in the same order, i.e., from poor-context to rich-context. Therefore, we followed the guidelines suggested by Mollon, Bosten, Peterzell and Webster (2017) to avoid carryover effects.

Participants were asked to ignore potential prior knowledge about illusions and to rely on their percepts only. The experimenter stayed in the experimental room during the whole experiment. Participants were debriefed at the end of the experiment and were shown their results.

Data analysis

For each participant and each of the 12 variants, the adjustments from both trials were averaged into a mean adjustment, from which the reference length was subtracted. The result was subsequently divided by the reference length. Hence, illusion magnitudes are a measure of bias as a proportion of the reference line. A positive magnitude (overadjustment) indicates that the adjustable line was longer than the reference, while a negative magnitude (underadjustment) indicates that the adjustable line was shorter than the reference. Analyses were performed with R (R Core Team, 2018).

Outliers were detected using a modified z -score, which is more robust than the commonly used z -score because it makes use of the median and median absolute deviation instead of the mean and standard deviation. Iglewicz and Hoaglin (1993) suggested considering participants with absolute modified z -score bigger than 3.5 as outliers. Data of 10 participants were removed based on that criterion (see “Participants” section).

For each variant, test-retest reliability was assessed by computing a Bravais-Pearson’s correlation between the values of the first and second adjustment trials. We

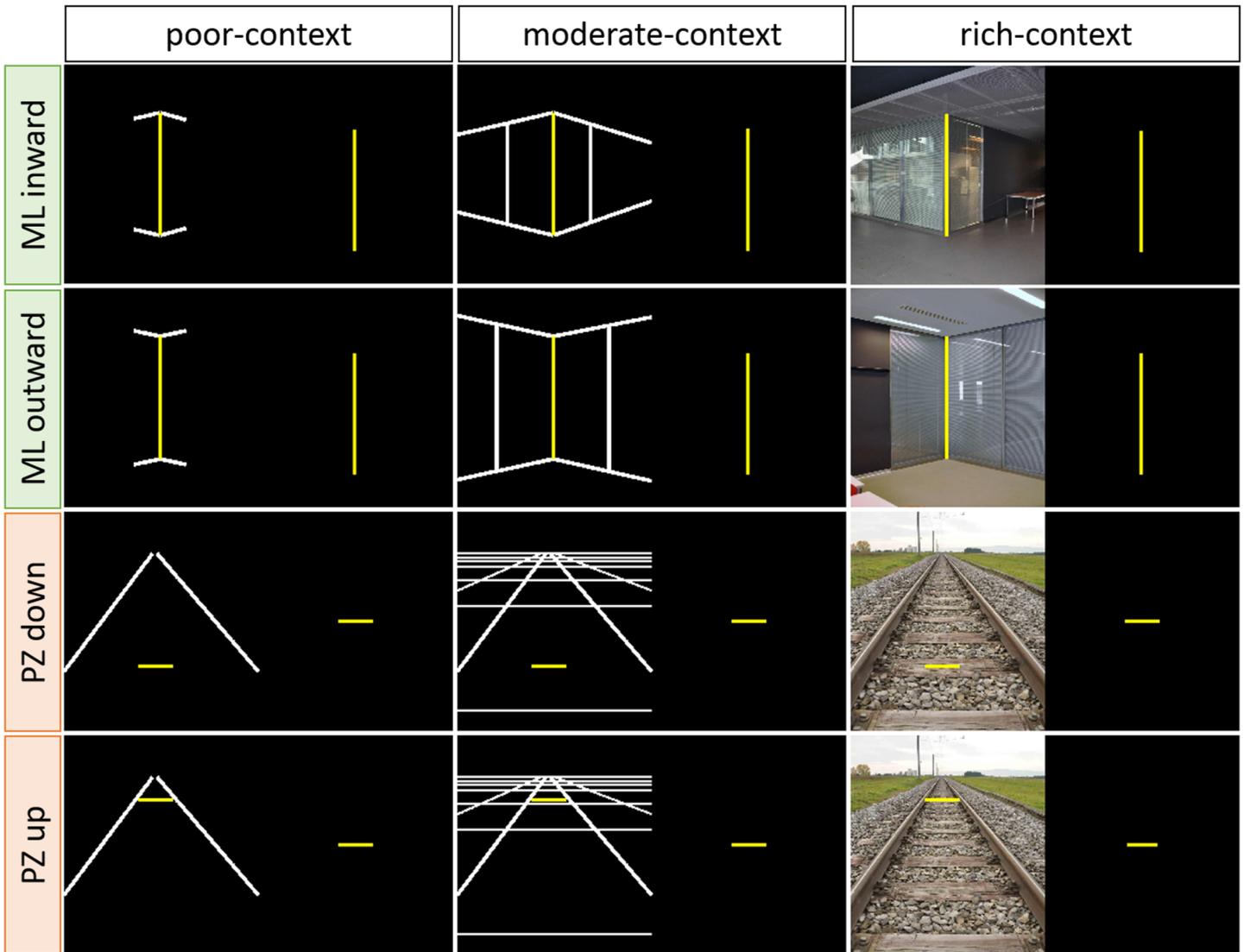


Figure 1. Participants had to adjust the length of the yellow adjustable line (right half screen) to match the length of the yellow reference line (left half screen) by moving the computer mouse. The “inward” and “outward” Müller-Lyer illusions were presented in distinct trials. Similarly, the Ponzo illusion was presented either with the upper line only (“up”) or with the lower line only (“down”). Each configuration was tested with three different contexts, making up 12 variants in total. By row: Müller-Lyer (ML; green) “inward” and “outward” configurations, Ponzo (PZ; orange) “down” and “up” configurations. By column: poor-context, moderate-context and rich-context variants.

also computed Bravais-Pearson’s correlations between the mean magnitudes for each pair of variants (66 correlations). Because of the moderate test-retest correlation coefficients (see “Correlations” section), pairwise correlations may have been underestimated due to measurement errors (e.g., Spearman, 1904b). To account for this, we also computed disattenuated pairwise correlations as:

$$r_{xy'} = \frac{r_{xy}}{\sqrt{r_{xx}r_{yy}}} \quad (1)$$

where $r_{xy'}$ is the disattenuated relationship between x and y , r_{xx} and r_{yy} are the test-retest reliabilities of the x and y variables, and r_{xy} is the attenuated (i.e.,

nondisattenuated) correlation coefficient between x and y (e.g., Osborne, 2003; Wang, 2010).

To account for random variations in baseline among participants and among configurations, mixed-effects models were computed. The fixed effects, also called predictors, were age and context. Intercepts and slopes were taken as random effects to account for differences in individual levels and for differences in the configurations (random intercepts) due to the main factor of context (random slopes). The model significance was obtained through likelihood ratio tests. We computed marginal and conditional effect sizes as measures of explained variance with the random effect structure included (conditional r^2) or excluded (marginal r^2) from the calculation.

Results

Illusion magnitudes

Illusion magnitudes are illustrated in Figure 2. Except from the poor-context outward ML, other ML variants showed the expected illusion susceptibility, i.e., the adjustable line was over- and underadjusted in the ML outward and inward configurations, respectively. The adjustable line was overadjusted in the PZ up configurations, which shows up as a positive illusion magnitude in all three contexts. The PZ down configurations surprisingly showed a very weak-almost null-effect in all three contexts.

Correlations

Significant but moderate test-retest correlations were found between the values of the first and second adjustments for all 12 variants (Table 1, diagonal). Both attenuated (Table 1, upper triangle) and disattenuated (Table 1, lower triangle) pairwise comparisons were computed between the mean magnitudes of each pair of variants. Most intra-illusion correlations were significant, whereas most inter-illusion correlations did not reach significance, suggesting that individual

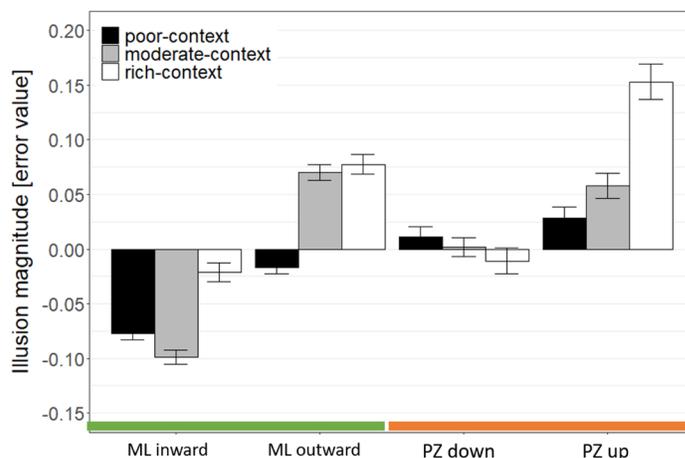


Figure 2. Illusion magnitudes \pm SEM (standard error of the mean) as a function of the configuration and context. Illusion magnitudes represent the proportion of the line over- (positive magnitude) or underadjusted (negative magnitude) compared to the reference line length. For example, 0.10 is a 10% misperception. ML: Müller-Lyer illusion, PZ: Ponzo illusion.

differences are more stable within an illusion (independently of the context) than across illusions. Indeed, a Welch *t*-test between the attenuated inter- and intra-illusion correlation coefficients resulted in

	ML1 inward	ML1 outward	ML2 inward	ML2 outward	ML3 inward	ML3 outward	PZ1 down	PZ1 up	PZ2 down	PZ2 up	PZ3 down	PZ3 up
ML1 inward	0.512	0.478	0.447	0.398	0.572	0.411	0.175	0.316	0.099	0.115	0.211	0.305
ML1 outward	0.966	0.478	0.235	0.573	0.460	0.397	-0.042	0.184	0.055	0.182	0.053	0.132
ML2 inward	0.932	0.508	0.449	0.334	0.556	0.438	0.206	0.238	0.065	0.182	0.189	0.121
ML2 outward	0.655	0.975	0.586	0.723	0.420	0.527	0.130	0.242	0.111	0.260	0.068	0.399
ML3 inward	0.951	0.791	0.986	0.587	0.708	0.674	0.107	0.267	0.160	0.201	0.161	0.290
ML3 outward	0.689	0.688	0.784	0.743	0.960	0.696	0.063	0.122	0.180	0.150	0.117	0.330
PZ1 down	0.370	-0.092	0.465	0.232	0.193	0.114	0.436	0.487	0.382	0.479	0.389	0.337
PZ1 up	0.586	0.354	0.471	0.378	0.421	0.193	0.979	0.567	0.334	0.603	0.470	0.306
PZ2 down	0.196	0.113	0.139	0.186	0.271	0.307	0.823	0.632	0.494	0.376	0.426	0.302
PZ2 up	0.212	0.348	0.357	0.403	0.315	0.237	0.956	1.000	0.705	0.576	0.421	0.396
PZ3 down	0.368	0.096	0.352	0.100	0.239	0.175	0.735	0.779	0.757	0.693	0.642	0.282
PZ3 up	0.542	0.243	0.229	0.597	0.439	0.503	0.649	0.518	0.547	0.664	0.448	0.617

Table 1. Test-retest and pairwise correlation table. Diagonal (in gray): test-retest reliability (Bravais-Pearson's *r*) for each variant. Upper triangle: attenuated correlation coefficients between each pair of variants (Bravais-Pearson's *r*). Lower triangle: disattenuated correlation coefficients between each pair of variants (Bravais-Pearson's *r*). Italics and bold font indicate significant results without and with Bonferroni correction, respectively. The color scale from blue to red reflects effect sizes from $r = -1$ to $r = 1$ (white corresponds to $r = 0$). Intra-illusion correlations are strong while inter-illusion correlations are in general weaker. ML: Müller-Lyer illusion, PZ: Ponzo illusion, 1: poor-context, 2: moderate-context, 3: rich-context.

	Varimax rotation		Promax rotation	
	RF1	RF2	RF1	RF2
ML1 inward	0.719	0.171	0.720	0.053
ML1 outward	0.714	-0.014	0.747	-0.139
ML2 inward	0.624	0.159	0.623	0.058
ML2 outward	0.725	0.146	0.730	0.026
ML3 inward	0.821	0.137	0.833	<0.001
ML3 outward	0.783	0.074	0.804	-0.059
PZ1 down	0.009	0.755	-0.122	0.786
PZ1 up	0.201	0.748	0.080	0.745
PZ2 down	0.042	0.655	-0.069	0.675
PZ2 up	0.139	0.766	0.012	0.775
PZ3 down	0.058	0.713	-0.063	0.733
PZ3 up	0.336	0.516	0.261	0.479

Table 2. Rotated factor (RF) loadings after the principal component analysis (PCA) and varimax or promax rotation. A color scale from red (positive loadings) to blue (negative loadings) is shown. ML: Müller-Lyer illusion, PZ: Ponzo illusion, 1: poor-context, 2: moderate-context, 3: rich-context.

a strongly significant difference (two-tailed t -test, $t[59.402] = 10.834$, $p < 0.001$, $d = 2.690$). Even when accounting for the moderate test-retest correlation coefficients (Table 1, lower triangle), there was a significant difference between disattenuated inter- and intra-illusion correlation coefficients (two-tailed t -test, $t[59.982] = 11.669$, $p < 0.001$, $d = 2.894$).

Thirty-five of 36 attenuated inter-illusion correlations were positive and the mean inter-illusion correlation was $r = 0.171$ ($SD = 0.092$), which is a small effect according to Cohen (1988) and a medium effect according to Gignac and Szodorai (2016) and Hemphill (2003). A one-sample two-tailed t -test on the effect sizes of the 36 attenuated inter-illusion correlations was significantly different from zero ($t[35] = 11.146$, $p < 0.001$, $d = 1.858$), suggesting that a small proportion of the variance underlying the ML and PZ illusions is accounted for by a common factor.

Factor analysis

Components were extracted using a principal component analysis (PCA). By inspecting the scree plot, we identified two components that accounted for $\sim 54\%$ of the total variance in the data ($\sim 35\%$ for the first component). Components were then rotated to achieve a simpler factor structure using a varimax (orthogonal) or a promax (oblique) rotation (Table 2). In both cases, the first and second factors mainly loaded on the ML and PZ variants, respectively. After a promax rotation, the intercorrelation between both factors was $r = 0.324$.

Fixed effects	β Estimate	β Standard error	t Value
(Intercept)	0.004	0.022	0.168
Moderate-context	0.022	0.021	1.024
Rich-context	0.063	0.028	2.291
Age	-0.0007	0.0003	-2.410

Table 3. Estimates from the mixed-effects model with age and context as predictors (no interaction between the two predictors).

Fixed effects	β Estimate	β Standard error	t Value
(Intercept)	0.0001	0.029	-0.005
Moderate-context	0.020	0.025	0.822
Rich-context	0.073	0.032	2.266
Age	-0.0002	0.001	-0.154

Table 4. Estimates from the mixed-effects model with age and context as predictors (no interaction between the two predictors). Data from participants aged younger than 23 years only were considered here.

Mixed linear models

To account for baseline differences across participants and across configurations, we computed mixed-effects models rather than a repeated-measures analysis of variance.

Additive versus interactive models (n = 75)

We tested for an interaction between age and context. A likelihood-ratio test showed a non-significant difference between additive and interactive models ($\chi^2(2) = 1.026$, $p < 0.599$). The estimates for the fixed effects of the additive model are reported in Table 3. Age only showed a tiny negative (Figure 3) but significant effect on illusion magnitudes ($\chi^2(1) = 5.537$, $p = 0.019$). The additive model accounted for 55.8% of the variance in the data, whereas it accounted for only 7.3% of the variance when the random effects and random slopes were not included in the model ($r^2_c = 0.558$; $r^2_m = 0.073$).

Additive versus interactive models in participants aged less than 23 years (n < 23 years = 52)

When considering participants aged 22 years or less only, a likelihood ratio test did not reveal a significant difference between the additive and the interactive model ($\chi^2(2) = 1.472$, $p = 0.479$; Table 4). Age did not show any significant effect on illusion magnitudes ($\chi^2(1)$

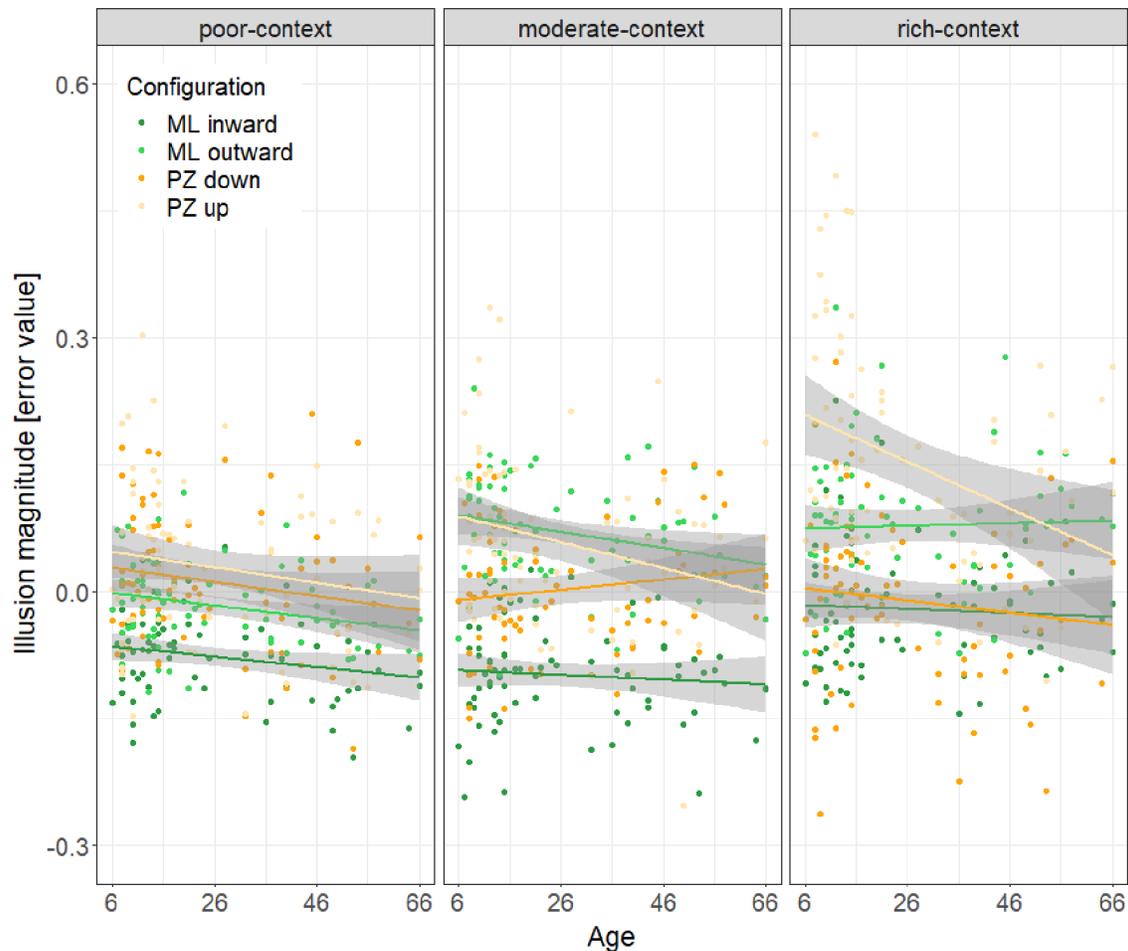


Figure 3. Illusion magnitudes as a function of age for the different contexts (panels from left to right: poor-context, moderate-context, and rich-context). A mixed-effects model indicated that illusion magnitudes slightly decreased with age. Colors represent the four different configurations and linear regression (colored lines) with 95% confidence interval (shadows) are shown. ML: Müller-Lyer illusion, PZ: Ponzo illusion.

$= 0.023$, $p = 0.879$). The conditional and marginal effect sizes are similar to what we observed when considering the full dataset ($r^2_c = 0.573$; $r^2_m = 0.073$).

Discussion

Age effects

We tested the effects of age and context on the Müller-Lyer and Ponzo illusion magnitudes. We observed a weak general decrease in illusion magnitude with age but did not find any significant interaction between the context and participants' age. These results do not support the two-factor theory of Piaget (1961, 1963). Wagner (1977) suggested that the two types of visual illusions suggested by Piaget depend on the context. Primary illusions (poor-context) and secondary illusions (rich-context) were proposed

to show decreasing and increasing effects with age, respectively. This was not the case here, even when we only considered participants aged less than 23 years. Likewise, the slight decrease in illusion magnitude is inconsistent with the empirical theory of Purves (Howe & Purves, 2004, 2005; Purves & Lotto, 2003), according to which we interpret retinal images based on our previous experience, which leads to age-related increases in the magnitude of geometrical illusions. However, Pressey (1974) observed a decrease in the Ponzo illusion magnitude with age, which he suggested to be due to the assimilative nature of the illusion (in opposition with contrast illusions).

Correlations within and between illusions

Although most inter-illusion correlations were weak, we observed strong intra-illusion correlations. Our

results suggest that the structure of the visual space underlying visual illusions is multifactorial (i.e., there seems to be a multitude of factors for illusions). For example, there seems to be at least one factor specific to the ML illusion, with effect sizes (intra-illusion correlation coefficients r) between 0.24 and 0.67, which reflects large effect sizes according to Gignac and Szodorai (2016) and Hemphill (2003) and medium to large effect sizes according to Cohen (1988). The difference between inter- and intra-illusion correlation coefficients was significant even after accounting for the moderate test-retest correlation coefficients (i.e., after disattenuating correlation coefficients).

These results support previous studies, where strong intra-illusion correlations and weak inter-illusion correlations were observed in the perception of visual illusions (Cretenoud et al., 2019; Grzeczowski et al., 2017, 2018). Several variants of the Ebbinghaus illusion, which varied in color, size, or texture, highly correlated. There were also strong correlations among different variants of the same illusion with different luminances and orientations but not across different illusions (Cretenoud et al., 2019). Therefore, it seems that each illusion involves one or several specific factor(s).

However, the two factors extracted from a dimensionality reduction technique showed some interdependency following an oblique rotation procedure. Similarly, 35 of 36 attenuated inter-illusion correlations were positive, suggesting the existence of some relationship between the ML and PZ illusion. Indeed, a one-sample t -test indicated that the 36 attenuated inter-illusion correlations were significantly different from zero. Nevertheless, the average strength of these correlations was $r = 0.17$, suggesting that this relationship is rather weak.

Response bias unlikely explains the weak inter-illusion correlations. First, by separating the adjustable line from the context, the possibility that participants used landmarks or strategies to perform the task was reduced. Second, if there was a response bias, test-retest reliabilities would be much higher than what we observed. Last, only a systematic response bias (i.e., either liberal or conservative adjustments in both illusions and both configurations) may underlie the positive inter-illusion correlations, which is very unlikely because the direction of the effects are different across configurations and because participants do not usually know the direction of the illusory effects.

Common factors for visual illusions

Some studies suggested a general common factor for visual illusions (e.g., Thurstone, 1944; see also Aftanas & Royce, 1969; Roff, 1953). To test this, we computed one-sample t -tests on two previous datasets (experiments 2 and 3 from Cretenoud et al., 2019).

First, five illusions were tested with four different orientations to determine whether illusion magnitudes are orientation-specific. In this dataset, inter-illusion correlations were not significantly different from 0 (two-tailed t -test, $t[159] = -0.780$, $p = 0.437$, $d = -0.062$). Second, when 10 illusions were tested with four luminance conditions, inter-illusion correlations were significantly different from zero (two-tailed t -test, $t[719] = 15.093$, $p < 0.001$, $d = 0.562$), even though the effect size was smaller than in the present experiment and the power rather low.

Therefore, as proposed for illusions of linear extent (Coren et al., 1976), we suggest that some subsets of illusions are closely linked. For example, a common mechanism, such as a constancy phenomenon (Gregory, 1963, 1964), seems to account for a small proportion of the variance in the ML and PZ illusions. However, because intra-illusion correlations are stronger and significantly larger than inter-illusion correlations, a large part of the variance seems to specifically tap into unique aspects of each illusion. The two main factors extracted from the factor analysis provide evidence toward this claim.

Individual differences in vision are nowadays widely studied (for reviews, see Mollon, Bosten, Peterzell, & Webster, 2017; Peterzell, 2016; Tulver, 2019). No study has so far found a single factor underlying the structure of visual space, contrary to the g factor of general intelligence (Spearman, 1904a). Only weak- or no-evidence for a general common factor was reported for bistable perception (Brascamp, Becker, & Hambrick, 2018; Cao, Wang, Sun, Engel, & He, 2018; Wexler, 2005), contrast perception (Bosten & Mollon, 2010), eye movements (Bargary et al., 2017), local-global processing (Chamberlain, Van der Hallen, Huygelier, Van de Cruys, & Wagemans, 2017; Milne & Szczerbinski, 2009), change detection (Andermane, Bosten, Seth, & Ward, 2019), face recognition (Verhallen et al., 2017; see also Ćepulić, Wilhelm, Sommer, & Hildebrandt, 2018), hue scaling (Emery, Volbrecht, Peterzell, & Webster, 2017a, 2017b), color matching (Webster & MacLeod, 1988), contrast sensitivity (Peterzell, 2016; Peterzell, Scheffrin, Tregear, & Werner, 2000), binocular disparity sensitivity (Peterzell, Serrano-Pedraza, Widdall, & Read, 2017), and in the use of expectations and knowledge priors (Tulver, Aru, Rutiku, & Bachmann, 2019), suggesting that vision is multifactorial. However, a few studies reported two common factors that are consistent with the activity of magnocellular and parvocellular systems (Dobkins, Gunther, & Peterzell, 2000; Peterzell & Teller, 1996; Simpson & McFadden, 2005; Ward, Rothen, Chang, & Kanai, 2017; but see Goodbourn et al., 2012) but that two-factor distinction seems to play no role in high-level vision, as reflected with visual illusions (Cretenoud et al., 2019).

Results are open to interpretation depending on the criterion for common factors. On the one hand, [Cappe and colleagues \(2014\)](#) tested participants with six basic visual tasks and argued against a common factor of visual performance because of a main PCA component explaining 34% of the total variance. On the other hand, [Halpern, Andrews and Purves \(1999\)](#) claimed that a main PCA component explaining 30% of the total variance gave evidence for a common factor underlying the seven tasks they tested. By comparison, 50% of the variance was suggested to be the lower threshold for a common factor of cognitive abilities such as the *g* factor ([Jensen, 1998](#); [Lubinski, 2000](#)). We computed a factor analysis, which resulted in two main factors together explaining 54% of the variance. Each factor mainly loaded on all variants of the ML or PZ illusion, respectively.

These results and the weak inter-illusion correlations that we observed between the ML and the PZ variants support the idea that there are several specific factors in the perception of illusions and vision in general. [Coren and colleagues \(1976\)](#) found that the inward and outward configurations of 11 different Müller-Lyer variants loaded on two factors, respectively. However, a second-order factor solution revealed only two main factors, among which one highly loaded on illusions of linear extent, including all Müller-Lyer (inward and outward) variants. These results highlight the differences in the interpretation of the results as a function of the methods chosen to compute factor analysis (e.g., to determine the number of factors to retain and how to rotate them).

Factors specific to classes of illusions have already been found almost 50 years ago. [Taylor \(1974\)](#) computed a factor analysis on 21 measures of visual illusions, which revealed three distinct factors. Several Poggendorff variants highly loaded on the first factor. The second factor involved distortions of parallelism, such as in the Zöllner or Hering illusions, whereas the third factor accounted for five illusions requiring length judgments, such as the Müller-Lyer illusion. We like to mention that 10 illusions were not strongly accounted for by any of these factors.

The average inter-illusion correlation in Taylor's study was low ($r = 0.16$), as was the case in a follow-up study, which included 18 illusion measures ([Taylor, 1976](#)). In the follow-up study, a four-factor solution was found with the first three factors that were similar to the ones found in the first study. However, some high loadings did not replicate between the two studies.

When adding 12 perceptual, cognitive or temperament measures to the factor analysis, the illusions still did not group into a general factor but were split into four factors. Contrary to [Thurstone \(1944\)](#), Taylor therefore showed that illusions do not cluster on a single dimension when embedded into a

heterogeneous task collection, suggesting that illusions themselves are heterogeneous perceptual tasks.

Similarly, [Robinson \(1968\)](#) suggested that illusions are too heterogeneous to be explained by a single mechanism. The author made a clear distinction between illusions, which involve a misperception of length or size, and distortions, which imply perceptual bending of lines. Likewise, the two higher-order classes of illusions found by [Coren and colleagues \(1976\)](#), i.e., illusions of linear extent and illusions of direction and area, were shown to interact with spatial abilities in a different way, suggesting that separate mechanisms underlie both classes of illusions ([Coren & Porac, 1987](#)).

Illusion magnitudes

Surprisingly, the inward and outward ML configurations were influenced by the context in a different way. The rich-context variants led to the weakest and strongest effect in the inward and outward configurations, respectively. We speculate that this pattern comes from the different backgrounds used in these two configurations. Contrary to the PZ illusion, where we used the same background picture for both up and down configurations, different mechanisms may come at hand with the different backgrounds in the ML configurations and therefore induce distinct effects. Further investigation is needed to specifically address this point.

The illusion magnitudes of the PZ down variants only resulted in very tiny-almost null-effects, while we expected these variants to be underestimated. Similar results were observed by [Yildiz, Sperandio, Kettle, and Chouinard \(2019\)](#), who tested the effect of linear perspective cues and texture gradients on a Ponzo illusion. A significant perceptual effect was observed in the expected directions when both top and bottom elements of the illusion were presented. However, the bottom element was not perceived differently than its physical size when the top element was presented outside of the background (as in the present experiment). The authors suggested that a higher degree of attention drawn to the upper part of the illusion, where pictorial distance cues take the lead over binocular and oculomotor distance cues, may explain the difference ([Cutting & Vishton, 1995](#)). If indeed participants focused more on the upper part of the contextual cues than on the lower one, then the centration theory by [Piaget \(1963\)](#) also explains the absence of effect observed in the PZ down variants. In addition, the large distance between the reference and the contextual cues may also have caused this absence of effect, contrary to the large effects observed in the PZ up variants where the reference line was closer to the contextual cues.

Experimental concerns

To account for the fact that a previous context may alter the susceptibility of a participant to subsequent contexts (carryover effects), we did not randomly interleaved the order of presentation of the different contexts (see Mollon et al., 2017). Illusion magnitudes varied with the context (Figure 2 and Table 3), specifically arguing against carryover effects. However, test-retest correlations were moderate, which may be due to the small number of repetitions for each variant (i.e., two trials per variant). The weak inter-illusion correlations may result from the moderate test-retest correlations. Disattenuated correlation coefficients (Table 1, lower triangle) show that inter-illusion correlations are indeed stronger when accounting for measurement errors but they are still significantly smaller than intra-illusion correlations. In the ML variants, the adjustable element was a vertical line, whereas a horizontal line was adjusted in the PZ variants. We previously tested several illusions with different orientations (Cretenoud et al., 2019) and observed strong intra-illusion correlations between different orientations. For example, a strong correlation ($r = 0.71$) was found between a Poggendorff illusion rotated at -15° and the same illusion after a 30° rotation, suggesting that illusion magnitudes are not orientation-specific. Therefore, it is unlikely that the different orientations of the adjustable element for the two illusions had a critical impact on the correlation strengths.

Conclusion

In summary, we found no interaction between age and context on the susceptibility to the Müller-Lyer and Ponzo illusions but rather a slight general decrease in illusion magnitudes with age, independently of the context. Accordingly, we previously observed only weak correlations between the magnitudes of different visual illusions (Grzeczowski et al., 2017; 2018) but strong correlations among different variants of a same illusion (Cretenoud et al., 2019), suggesting the existence of factors specific to each illusion. Here, we similarly showed that there are strong correlations among different variants of a same illusion, which varied in contextual cues, but weaker correlations between the two illusions we tested. Together, these results suggest that there are a multitude of specific factors underlying visual illusions.

Keywords: factors, illusions, individual differences, context

Acknowledgments

We thank Marc Repnow for technical support and Professor Pierre Barrouillet, director of the Jean Piaget Archives in Geneva, Switzerland, for helpful insights.

Supported by the project “Basics of visual processing: From elements to figures” (Project number: 320030_176153/1) of the Swiss National Science Foundation (SNSF) and by a National Centre of Competence in Research (NCCR Synapsy) grant from the SNSF (51NF40-185897).

Commercial relationships: none.

Corresponding author: Aline F. Cretenoud.

Email: aline.cretenoud@epfl.ch.

Address: École Polytechnique Fédérale de Lausanne (EPFL), Laboratory of Psychophysics, Brain Mind Institute, 1015 Lausanne, Switzerland.

References

- Aftanas, M. S., & Royce, J. R. (1969). A factor analysis of brain damage tests administered to normal subjects with factor score comparisons across ages. *Multivariate Behavioral Research*, 4(4), 459–481.
- Andermane, N., Bosten, J. M., Seth, A. K., & Ward, J. (2019). Individual differences in change blindness are predicted by the strength and stability of visual representations. *Neuroscience of Consciousness*, 5(1), 1–12, <https://doi.org/10.1093/nc/niy010>.
- Bargary, G., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). Individual differences in human eye movements: An oculomotor signature? *Vision Research*, 141, 157–169, <https://doi.org/10.1016/j.visres.2017.03.001>.
- Binet, A. (1895). La mesure des illusions visuelles chez les enfants. *Revue Philosophique de La France et de l'Étranger*, 40, 11–25.
- Bosten, J. M., Goodbourn, P. T., Bargary, G., Verhallen, R. J., Lawrance-Owen, A. J., Hogg, R. E., . . . Mollon, J. D. (2017). An exploratory factor analysis of visual performance in a large population. *Vision Research*, 141, 303–316, <https://doi.org/10.1016/j.visres.2017.02.005>.
- Bosten, J. M., & Mollon, J. D. (2010). Is there a general trait of susceptibility to simultaneous contrast? *Vision Research*, 50(17), 1656–1664, <https://doi.org/10.1016/j.visres.2010.05.012>.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436.
- Brascamp, J. W., Becker, M. W., & Hambrick, D. Z. (2018). Revisiting individual differences in the time

- course of binocular rivalry. *Journal of Vision*, 18(7), 1–20, <https://doi.org/10.1167/18.7.3>.
- Brislin, R. W. (1974). The Ponzo illusion: Additional cues, age, orientation, and culture. *Journal of Cross-Cultural Psychology*, 5(2), 139–161, <https://doi.org/10.1177/002202217400500201>.
- Brislin, R. W., & Keating, C. F. (1976). Cultural differences in the perception of a three-dimensional Ponzo illusion. *Journal of Cross-Cultural Psychology*, 7(4), 397–412.
- Cao, T., Wang, L., Sun, Z., Engel, S. A., & He, S. (2018). The independent and shared mechanisms of intrinsic brain dynamics: Insights from bistable perception. *Frontiers in Psychology*, 9:589, <https://doi.org/10.3389/fpsyg.2018.00589>.
- Caparos, S., Ahmed, L., Bremner, A. J., de Fockert, J. W., Linnell, K. J., & Davidoff, J. (2012). Exposure to an urban environment alters the local bias of a remote culture. *Cognition*, 122(1), 80–85, <https://doi.org/10.1016/j.cognition.2011.08.013>.
- Cappe, C., Clarke, A., Mohr, C., & Herzog, M. H. (2014). Is there a common factor for vision? *Journal of Vision*, 14(8), 1–11, <https://doi.org/10.1167/14.8.4>.
- Ćepulić, D. B., Wilhelm, O., Sommer, W., & Hildebrandt, A. (2018). All categories are equal, but some categories are more equal than others: The psychometric structure of object and face cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(8), 1254–1268.
- Chamberlain, R., Van der Hallen, R., Huygelier, H., Van de Cruys, S., & Wagemans, J. (2017). Local-global processing bias is not a unitary individual difference in visual processing. *Vision Research*, 141, 247–257, <https://doi.org/10.1016/j.visres.2017.01.008>.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: L. Erlbaum Associates.
- Coren, S., Girgus, J. S., Erlichman, H., & Hakstian, A. R. (1976). An empirical taxonomy of visual illusions. *Perception & Psychophysics*, 20(2), 129–137, <https://doi.org/10.3758/BF03199444>.
- Coren, S., & Porac, C. (1987). Individual differences in visual-geometric illusions: Predictions from measures of spatial cognitive abilities. *Perception & Psychophysics*, 41(3), 211–219, <https://doi.org/10.3758/BF03208220>.
- Cretenoud, A. F., Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, 19(14), 1–21, <https://doi.org/https://doi.org/10.1167/19.14.12>.
- Cutting, J. E., & Vishton, P. M. (1995). Perceived layout: The integration, relative dominance, and contextual use of different information about depth. In: W. Epstein, & S. Rogers (Eds.), *Handbook of Perception and Cognition*. (pp. 69–117). NY: Academic Press.
- de Fockert, J., Davidoff, J., Fagot, J., Parron, C., & Goldstein, J. (2007). More accurate size contrast judgments in the Ebbinghaus Illusion by a remote culture. *Journal of Experimental Psychology: Human Perception and Performance*, 33(3), 738–742, <https://doi.org/10.1037/0096-1523.33.3.738>.
- Dobkins, K. R., Gunther, K. L., & Peterzell, D. H. (2000). What covariance mechanisms underlie green/red equiluminance, luminance contrast sensitivity and chromatic (green/red) contrast sensitivity? *Vision Research*, 40(6), 613–628, [https://doi.org/10.1016/S0042-6989\(99\)00211-4](https://doi.org/10.1016/S0042-6989(99)00211-4).
- Doherty, M. J., Campbell, N. M., Tsuji, H., & Phillips, W. A. (2010). The Ebbinghaus illusion deceives adults but not young children. *Developmental Science*, 13(5), 714–721, <https://doi.org/10.1111/j.1467-7687.2009.00931.x>.
- Dresslar, F. B. (1894). Studies in the psychology of touch. *The American Journal of Psychology*, 6(3), 313–368, <https://doi.org/10.2307/1411644>.
- Duemmler, T., Franz, V. H., Jovanovic, B., & Schwarzer, G. (2008). Effects of the Ebbinghaus illusion on children’s perception and grasping. *Experimental Brain Research*, 186(2), 249–260, <https://doi.org/10.1007/s00221-007-1229-0>.
- Emery, K. J., Volbrecht, V. J., Peterzell, D. H., & Webster, M. A. (2017a). Variations in normal color vision. VI. Factors underlying individual differences in hue scaling and their implications for models of color appearance. *Vision Research*, 141, 51–65, <https://doi.org/10.1016/j.visres.2016.12.006>.
- Emery, K. J., Volbrecht, V. J., Peterzell, D. H., & Webster, M. A. (2017b). Variations in normal color vision. VII. Relationships between color naming and hue scaling. *Vision Research*, 141, 66–75, <https://doi.org/10.1016/j.visres.2016.12.007>.
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74–78, <https://doi.org/10.1016/j.paid.2016.06.069>.
- Goodbourn, P. T., Bosten, J. M., Hogg, R. E., Bargary, G., Lawrence-Owen, A. J., & Mollon, J. D. (2012). Do different “magnocellular tasks” probe the same neural substrate? *Proceedings of the Royal Society B: Biological Sciences*, 279(1745), 4263–4271, <https://doi.org/10.1098/rspb.2012.1430>.
- Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature*, 199, 678–680.

- Gregory, R. L. (1964). Illusory perception as a constancy phenomenon. *Nature*, *204*, 302–303.
- Grzeczowski, L., Clarke, A. M., Francis, G., Mast, F. W., & Herzog, M. H. (2017). About individual differences in vision. *Vision Research*, *141*, 282–292, <https://doi.org/10.1016/j.visres.2016.10.006>.
- Grzeczowski, L., Roinishvili, M., Chkonia, E., Brand, A., Mast, F. W., Herzog, M. H., . . . Shaqiri, A. (2018). Is the perception of illusions abnormal in schizophrenia? *Psychiatry Research*, *270*, 929–939, <https://doi.org/10.1016/j.psychres.2018.10.063>.
- Halpern, S. D., Andrews, T. J., & Purves, D. (1999). Interindividual variation in human visual performance. *Journal of Cognitive Neuroscience*, *11*(5), 521–534.
- Hanisch, C., Konczak, J., & Dohle, C. (2001). The effect of the Ebbinghaus illusion on grasping behaviour of children. *Experimental Brain Research*, *137*(2), 237–245, <https://doi.org/10.1007/s002210000655>.
- Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, *58*(1), 78–79, <https://doi.org/10.1037/0003-066X.58.1.78>.
- Howe, C. Q., & Purves, D. (2004). Size contrast and assimilation explained by the statistics of natural scene geometry. *Journal of Cognitive Neuroscience*, *16*(1), 90–102, <https://doi.org/10.1162/089892904322755584>.
- Howe, C. Q., & Purves, D. (2005). *Perceiving geometry: Geometrical illusions explained by natural scene statistics*. New York, NY: Springer Science & Business Media.
- Iglewicz, B., & Hoaglin, D. (1993). Volume 16: how to detect and handle outliers. In: *The ASQC basic references in quality control: statistical techniques*. Milwaukee, WI: Asq Press.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Káldy, Z., & Kovács, I. (2003). Visual context integration is not fully developed in 4-year-old children. *Perception*, *32*(6), 657–666, <https://doi.org/10.1068/p3473>.
- Leibowitz, H., Brislin, R., Perlmutter, L., & Hennessy, R. (1969). Ponzo perspective illusion as a manifestation of space perception. *Science*, *166*(3909), 1174–1176.
- Leibowitz, H. W., & Heisel, M. A. (1958). L'évolution de l'illusion de Ponzo en fonction de l'âge. *Archives de Psychologie*, *36*, 328–331.
- Leibowitz, H. W., & Pick, H. A. (1972). Cross-cultural and educational aspects of the Ponzo perspective illusion. *Perception & Psychophysics*, *12*(5), 430–432, <https://doi.org/10.3758/BF03205856>.
- Lubinski, D. (2000). Scientific and social significance of assessing individual differences: “Sinking shafts at a few critical points”. *Annual Review of Psychology*, *51*(1), 405–444.
- Milne, E., & Szczerbinski, M. (2009). Global and local perceptual style, field-independence, and central coherence: An attempt at concept validation. *Advances in Cognitive Psychology*, *5*, 1–26, <https://doi.org/10.2478/v10053-008-0062-8>.
- Mollon, J. D., Bosten, J. M., Peterzell, D. H., & Webster, M. A. (2017). Individual differences in visual science: What can be learned and what is good experimental practice? *Vision Research*, *141*, 4–15, <https://doi.org/10.1016/j.visres.2017.11.001>.
- Ninio, J. (2014). Geometrical illusions are not always where you think they are: a review of some classical and less classical illusions, and ways to describe them. *Frontiers in Human Neuroscience*, *8*:856, <https://doi.org/10.3389/fnhum.2014.00856>.
- Osborne, J. W. (2003). Effect sizes and the disattenuation of correlation and regression coefficients: Lessons from educational psychology. *Practical Assessment, Research and Evaluation*, *8*(11), 5–11.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.
- Peterzell, D. H. (2016). Discovering sensory processes using individual differences: A review and factor analytic manifesto. *Electronic Imaging*, 2016(16), 1–11, <https://doi.org/10.2352/ISSN.2470-1173.2016.16HVEI-112>.
- Peterzell, D. H., Scheffrin, B. E., Tregear, S. J., & Werner, J. S. (2000). *Spatial frequency tuned covariance channels underlying scotopic contrast sensitivity. Vision Science and its Applications: OSA Technical Digest*. Washington, D.C.: Optical Society of America.
- Peterzell, D. H., Serrano-Pedraza, I., Widdall, M., & Read, J. C. A. (2017). Thresholds for sine-wave corrugations defined by binocular disparity in random dot stereograms: Factor analysis of individual differences reveals two stereoscopic mechanisms tuned for spatial frequency. *Vision Research*, *141*, 127–135, <https://doi.org/10.1016/j.visres.2017.11.002>.
- Peterzell, D. H., & Teller, D. Y. (1996). Individual differences in contrast sensitivity functions: The lowest spatial frequency channels. *Vision Research*, *36*(19), 3077–3085, [https://doi.org/10.1016/0042-6989\(96\)00061-2](https://doi.org/10.1016/0042-6989(96)00061-2).
- Piaget, J. (1961). *Les Mécanismes perceptifs, Modèles probabilistes, Analyse génétique, Relations avec l'intelligence* (2nd ed.). Paris: Presses universitaires de France.

- Piaget, J. (1963). Le développement des perceptions en fonction de l'âge. In: *Traité de psychologie expérimentale. Vol. VI, La Perception* (pp. 1–162). Paris: Presses universitaires de France.
- Pollack, R. H. (1964). Simultaneous and successive presentation of elements of the Mueller-Lyer figure and chronological age. *Perceptual and Motor Skills*, 19(1), 303–310.
- Pollack, R. H., & Silvar, S. D. (1967). Magnitude of the Mueller-Lyer illusion in children as a function of pigmentation of the fundus oculi. *Psychonomic Science*, 8(2), 83–84, <https://doi.org/10.3758/BF03330678>.
- Pressey, A. W. (1974). Age changes in the Ponzo and filled-space illusions. *Perception & Psychophysics*, 15(2), 315–319.
- Purves, D., & Lotto, R. B. (2003). *Why we see what we do: An empirical theory of vision*. Sunderland, MA: Sinauer Associates. Retrieved from <https://psycnet.apa.org/record/2003-02306-000>.
- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.r-project.org/>.
- Rival, C., Olivier, I., Ceyte, H., & Bard, C. (2004). Age-related differences in the visual processes implied in perception and action: Distance and location parameters. *Journal of Experimental Child Psychology*, 87(2), 107–124, <https://doi.org/10.1016/j.jecp.2003.10.003>.
- Robinson, J. O. (1968). Retinal inhibition in visual distortion. *British Journal of Psychology*, 59(1), 29–36.
- Roff, M. (1953). A factorial study of tests in the perceptual area. *Psychometric Monographs*, 8, 1–41.
- Segall, M. H., Campbell, D. T., & Herskovits, M. J. (1963). Cultural differences in the perception of geometric illusions. *Science*, 139(3556), 769–771.
- Shaqiri, A., Pilz, K. S., Cretienoud, A. F., Neumann, K., Clarke, A., Kunchulia, M., . . . Herzog, M. H. (2019). No evidence for a common factor underlying visual abilities in healthy older people. *Developmental Psychology*, 55(8), 1775–1787, <https://doi.org/10.1037/dev0000740>.
- Silvar, S. D., & Pollack, R. H. (1967). Racial differences in pigmentation of the fundus oculi. *Psychonomic Science*, 7(4), 159–159, <https://doi.org/10.3758/BF03328514>.
- Simpson, W. A., & McFadden, S. M. (2005). Spatial frequency channels derived from individual differences. *Vision Research*, 45(21), 2723–2727, <https://doi.org/10.1016/j.visres.2005.01.015>.
- Spearman, C. (1904a). “General intelligence” objectively determined and measured. *The American Journal of Psychology*, 15(2), 201–292.
- Spearman, C. (1904b). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101, <https://doi.org/10.1037/h0065390>.
- Stewart, V. M. (1973). Tests of the “carpentered world” hypothesis by race and environment in America and Zambia. *International Journal of Psychology*, 8(2), 83–94.
- Taylor, T. R. (1974). A factor analysis of 21 illusions: The implications for theory. *Psychologia Africana*, 15, 137–148.
- Taylor, T. R. (1976). The factor structure of geometric illusions: A second study. *Psychologia Africana*, 16, 177–200.
- Thiéry, A. (1896). Ueber geometrisch-optische Täuschungen. *Philosophische Studien*, 12, 67–126.
- Thurstone, L. L. (1944). *A factorial study of perception*. Chicago, IL, US: The University of Chicago Press.
- Tulver, K. (2019). The factorial structure of individual differences in visual perception. *Consciousness and Cognition*, 73:102762, <https://doi.org/https://doi.org/10.1016/j.concog.2019.102762>.
- Tulver, K., Aru, J., Rutiku, R., & Bachmann, T. (2019). Individual differences in the effects of priors on perception: A multi-paradigm approach. *Cognition*, 187(3), 167–177, <https://doi.org/10.1016/j.cognition.2019.03.008>.
- Verhallen, R. J., Bosten, J. M., Goodbourn, P. T., Lawrance-owen, A. J., Bargary, G., & Mollon, J. D. (2017). General and specific factors in the processing of faces. *Vision Research*, 141, 217–227, <https://doi.org/10.1016/j.visres.2016.12.014>.
- Wagner, D. A. (1977). Ontogeny of the Ponzo illusion: Effects of age, schooling, and environment. *International Journal of Psychology*, 12(3), 161–176, <https://doi.org/10.1080/00207597708247386>.
- Wang, L. L. (2010). Disattenuation of correlations due to fallible measurement. *Newborn and Infant Nursing Reviews*, 10(1), 60–65, <https://doi.org/10.1053/j.nainr.2009.12.013>.
- Ward, J., Rothen, N., Chang, A., & Kanai, R. (2017). The structure of inter-individual differences in visual ability: Evidence from the general population and synaesthesia. *Vision Research*, 141, 293–302, <https://doi.org/10.1016/j.visres.2016.06.009>.
- Webster, M. A., & MacLeod, D. I. A. (1988). Factors underlying individual differences in the color matches of normal observers. *Journal of the*

Optical Society of America, 5(10), 1722–1735,
<https://doi.org/10.1364/josaa.5.001722>.

- Weintraub, D. J. (1979). Ebbinghaus illusion: Context, contour, and age influence the judged size of a circle amidst circles. *Journal of Experimental Psychology: Human Perception and Performance*, 5(2), 353–364.
- Wexler, M. (2005). Anticipating the three-dimensional consequences of eye movements. *Proceedings of the National Academy of Sciences of the United States of America*, 102(4), 1246–1251.
- Yildiz, G. Y., Sperandio, I., Kettle, C., & Chouinard, P. A. (2019). The contribution of linear perspective cues and texture gradients in the perceptual rescaling of stimuli inside a Ponzo illusion corridor. *PLoS One*, 14(10), e0223583, <https://doi.org/10.1371/journal.pone.0223583>.

Appendix

Description of the stimuli

In all Müller-Lyer (ML) variants, the reference line was 15.6° long, shifted by 0.2° to the left compared with the middle of the left half screen and vertically centered. The adjustable line was similarly drawn in the right half screen but vertically displaced by 2.1° to the bottom to avoid any direct horizontal comparison of the two lines. The length of the adjustable line was pseudorandomized at the beginning of each trial between 0° and 21.5°.

In the moderate-context inward ML variant, two oblique lines linked the extremities of the reference line with the left border of the screen, making an angle of 77° compared with the reference line. Similarly, two oblique lines were drawn between the extremities

of the reference line and the vertical screen midline, making an angle of 74.5° compared with the reference line. In the moderate-context outward ML variant, all oblique lines made an angle of 102.2° compared with the reference line. Two vertical lines were added 6.1° to the left and 4.6° to the right of the reference line in the inward variant, and 7.4° to the left and 7° to the right of the reference line in the outward variant, stopping at the intersection with the oblique lines.

In the poor-context ML variants, there were no additional vertical lines and the oblique lines were cut into 3.5° long fins.

In all Ponzo (PZ) variants, the horizontal reference line was 4.4° long and shifted by 0.8° to the left and 5.8° to the top (up configurations) or bottom (down configurations) compared with the middle of the left half screen. The adjustable horizontal line was perfectly centered in the right half screen and its length was pseudorandomly chosen at the beginning of each trial between 0° and 23.7°.

In the poor-context PZ variants, two diagonals centered at (5.7°, 12.8°) and (18.3°, 12.8°) from the upper left corner of the screen, oriented 36.9° clockwise and 40.7° counterclockwise and 18.9° and 19.8° long, were drawn, respectively.

In the moderate-context PZ variants, two diagonals, as well as seven horizontal lines, were added to the elements of the poor-context PZ variants. The two additional diagonals were centered at (5.4°, 7.7°) and (18.6°, 7.7°) from the left corner of the screen, oriented 66.2° clockwise and 69.1° counterclockwise and were 11.8° and 13.3° long, respectively. The horizontal lines were 5.3°, 5.8°, 6.4°, 7°, 8.8°, 12°, and 25.4° away from the top of the screen. Their length was half the width of the screen.

Appendix C

Cretenoud, A. F., Francis, G., & Herzog, M. H. (2020). When illusions merge. *Journal of Vision*, 20(8):12, 1–15, <https://doi.org/10.1167/jov.20.8.12>.

Copyright holder: © ARVO

Published under CC-BY-NC-ND 4.0

When illusions merge

Aline F. Cretenoud

Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland



Gregory Francis

Psychological Sciences, Purdue University, West Lafayette, IN, USA



Michael H. Herzog

Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland



We recently found only weak correlations between the susceptibility to various visual illusions. However, we observed strong correlations among different variants of an illusion, suggesting that the visual space of illusions includes several illusion-specific factors. Here, we specifically examined how factors for the vertical–horizontal, Müller–Lyer, and Ponzo illusions relate to each other. We measured the susceptibility to each illusion separately and to combinations of two illusions, which we refer to as a merged illusion; for example, we tested the Müller–Lyer illusion and the vertical–horizontal illusion, as well as a merged version of both illusions. We used an adjustment procedure in two experiments with 306 and 98 participants, respectively. Using path analyses, correlations, and exploratory factor analyses, we found that the susceptibility to a merged illusion is well predicted from the susceptibilities to the individual illusions. We suggest that there are illusion-specific factors that, by independent combinations, represent the whole visual structure underlying illusions.

For vision, however, there seems to be no strong, unique factor but rather several specific ones (for reviews, see Mollon, Bosten, Peterzell, & Webster, 2017; Peterzell, 2016; Tulver, 2019). For example, more than 1000 participants were tested with 25 visual and auditory measures (Bosten, Goodbourn, Bargary, Verhallen, Lawrance-Owen, Hogg, & Mollon, 2017). Only 20% of the total variance underlying these 25 measures was explained by a unique factor, whereas eight specific factors explained about 57% of the total variance. In addition, Brascamp, Becker, and Hambrick (2018) claimed that there is only weak evidence for a single mechanism (i.e., a common factor) to underlie different forms of bistable perception, such as binocular rivalry (see also Cao, Wang, Sun, Engel, & He, 2018). Similarly, there was only weak evidence for a unique factor in contrast perception (Bosten & Mollon, 2010; Peterzell, 2016; Peterzell, Scheffrin, Tregear, & Werner, 2000), eye movements (Bargary, Bosten, Goodbourn, Lawrance-Owen, Hogg, & Mollon, 2017), face recognition (Verhallen, Bosten, Goodbourn, Lawrance-Owen, Bargary, & Mollon, 2017), hue scaling (e.g., Emery, Volbrecht, Peterzell, & Webster, 2017a; Emery, Volbrecht, Peterzell, & Webster, 2017b), local–global processing (Chamberlain, Van der Hallen, Huygelier, Van de Cruys, & Wagemans, 2017; Milne & Szczerbinski, 2009), color matching (Webster & MacLeod, 1988), stereopsis (e.g., Hibbard, Bradshaw, Langley, & Rogers, 2002; Peterzell, Serrano-Pedraza, Widdall, & Read, 2017), luminance contrast sensitivity (Dobkins, Gunther, & Peterzell, 2000; Peterzell, Chang, & Teller, 2000), and in the use of expectations and knowledge priors (Tulver, Aru, Rutiku, & Bachmann, 2019), suggesting that the structure of visual space is multifactorial. Only weak correlations were also found between performance in six basic visual tasks,

Introduction

Common factors are frequently encountered in everyday life. For example, the Big Five personality traits scale determines five different aspects of personality: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (John & Srivastava, 1999). These five specific factors are supposed to represent the factor structure of personality. Similarly, *g* (general intelligence) is a common factor for intelligence that can be inferred from several specific factors, such as the Wechsler scale (Wechsler, 2003). Likewise, touch and audition have been shown to correlate highly, reflecting a common factor for somatosensation (Frenzel et al., 2012).

Citation: Cretenoud, A. F., Francis, G., & Herzog, M. H. (2020). When illusions merge. *Journal of Vision*, 20(8):12, 1–15, <https://doi.org/10.1167/jov.20.8.12>.



such as visual backward masking and Vernier offset discrimination (Cappe, Clarke, Mohr, & Herzog, 2014).

Likewise, only weak correlations were observed between the susceptibility to different illusions, such as the Ebbinghaus and Müller–Lyer illusions (Grzeczowski, Clarke, Francis, Mast, & Herzog, 2017; Grzeczowski, Roinishvili, Chkonia, Brand, Mast, Herzog, & Shaqiri, 2018), suggesting that the structure of the visual space underlying illusions is multifactorial. Factors specific to classes of illusions have been found in the past. For example, Thurstone (1944) found a factor underlying geometric illusions that, however, did not show strong loadings on other classes of illusions such as brightness or size–weight illusions. Likewise, Coren, Girgus, Erlichman, and Hakstian (1976) suggested that visual illusions can be classified into two classes: illusions of linear extent and illusions of shape or direction (see also Robinson, 1968). Taylor (1974, 1976) found that a set of different illusion measures was best represented by a three- or four-factor model, suggesting that visual illusions are heterogeneous perceptual tasks (but see Aftanas & Royce, 1969; Roff, 1953). Similarly, it has recently been suggested (Bulatov, 2017) that the visual structure underlying illusions of spatial extent, such as the Müller–Lyer illusion, combines several physiological and psychological factors.

Recently, we tested several variants of the Ebbinghaus illusion, differing in color, shape, size, and texture, and found strong correlations between the susceptibility to all variants. Similarly, when testing several illusions with different luminance, orientation, and contextual conditions, strong within-illusion correlations but only weak between-illusion correlations were observed (Cretienoud, Grzeczowski, Bertamini, & Herzog, 2020; Cretienoud, Karimpur, Grzeczowski, Francis, Hamburger, & Herzog, 2019); that is, an individual who is strongly susceptible to one illusion is not necessarily strongly susceptible to a different illusion, but there is a high probability for this individual to also be strongly susceptible to another variant of the same illusion. We suggested that factors are illusion-specific, even though a small proportion of the variance in the illusion magnitudes (especially in some subsets of illusions) may be accounted for by a general—but weak—common factor.

Here, we wondered how these factors relate to each other. To this aim, we merged pairs of illusions and tested whether the susceptibility to a merged illusion can be predicted from the susceptibilities to the non-merged illusions of which it was made.

Experiment 1

Methods

Participants

Participants were students at Purdue University who received course credit in return for their participation.

Among the 310 students that took part in this experiment, 306 were considered for further analysis (four participants were considered as outliers). Ages ranged from 19 to 27 years (mean age, 22 years), and 109 females participated. Two participants did not provide any information about their gender. Procedures were conducted in accordance with the tenets of the Declaration of Helsinki, except for the preregistration, and were approved by the local ethics committee.

Apparatus

Each participant ran the experiment on their own computer in a web browser over the Internet. Due to the online nature of the experiment, the details of a participant's computer and screen are mostly unknown. The experiment software was able to verify that the web browser was on a desktop or laptop computer rather than on a handheld device. Participants trying to access the experiment with a phone or tablet were asked to try again with a desktop or laptop computer.

Stimuli

Figure 1 provides a schematic of the stimuli used in Experiment 1. A horizontal reference line on the bottom right was either 100 or 150 pixels long. An adjustable line on the top left had a length randomly selected between 15 pixels and an upper limit that was itself randomly chosen to be 75 to 100 pixels larger than the length of the reference line, with the restriction that it could not be closer than 15 pixels to the length of the reference line. Each line was 2 pixels thick. The adjustable line was presented either horizontally or vertically, as in the vertical–horizontal illusion. As in the Müller–Lyer illusion, inward and outward wings were added to some conditions in a virtual square with sides of 15 pixels and at a 45° angle relative to the adjustable line. With two reference line lengths (100 and 150 pixels), two adjustable line orientations (horizontal and vertical), and three types of wings (inward, none, and outward) for the adjustable line, there were 12 stimulus conditions.

Because the reference line was always horizontal, we considered the adjustable line in the horizontal no wings conditions (HorNone100 and HorNone150) to be control conditions. In addition to these control conditions, we considered six conditions as non-merged conditions because they referred to one illusory effect only: a vertical–horizontal illusion in which the vertical line is usually perceived as being longer than a horizontal line of the same length (vertical no wings conditions: VerNone100 and VerNone150) or a Müller–Lyer illusion in which a line with inward wings usually looks shorter than the same line with outward wings (horizontal inward wings conditions: HorIn100 and HorIn150; horizontal outward wings conditions: HorOut100 and HorOut150). Four conditions were considered as merged conditions because they included

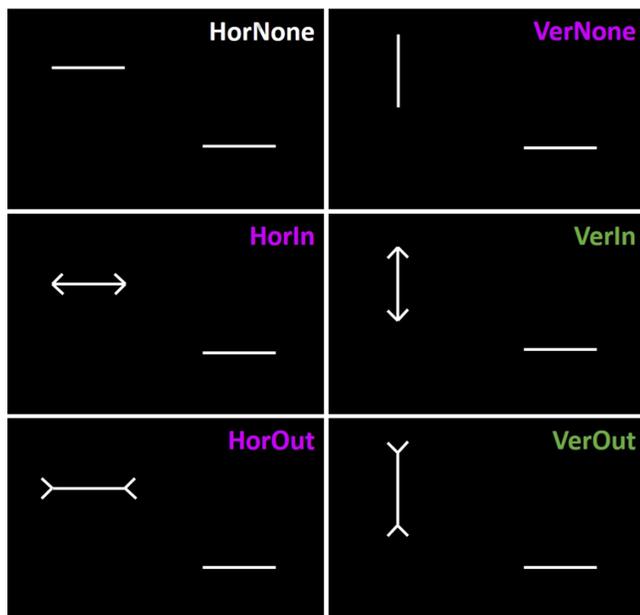


Figure 1. In [Experiment 1](#), 12 conditions were tested. The six conditions illustrated here (2 adjustable line orientations \times 3 types of wings) were tested with two reference line lengths (100 and 150 pixels). Abbreviations: “Hor” = horizontal adjustable line; “Ver” = vertical adjustable line; “None” = no wings; “In” = inward wings; “Out” = outward wings. The participant moved a slider to adjust the length of the top left line so that it appeared to be same length as the reference line on the bottom right. Conditions with purple labels show non-merged conditions, as they refer to one illusory effect only (either vertical–horizontal or Müller–Lyer illusion), whereas conditions with green labels show merged conditions. We considered the comparison of two horizontal lines (HorNone conditions; white label) as purely non-illusory.

both illusory effects: the adjustable line was a vertical line with either inward wings (VerIn100 and VerIn150) or outward wings (VerOut100 and VerOut150).

Procedure

Each trial started when the participant clicked a button. A condition appeared on the screen and the participant moved a slider to the left or right to change the length of the adjustable line. The task was to set the length of the adjustable line so that it appeared to be the same length as the reference line. When the participants were satisfied that the lines looked the same length, they clicked on a “Submit Match” button to end the trial. The first four trials were practice trials that presented all combinations of reference line length and adjustable line orientation. The adjustable line always had no wings during the practice trials. For subsequent trials, each condition was presented eight times in a random order. The final three questions gathered demographic

information (sex, age, and ethnicity). The experiment took less than 30 minutes to complete.

Data analysis

Intrarater reliability was assessed by computing two-way mixed-effects models (intraclass correlations of type [3,1] or $ICC_{3,1}$) for each condition, as suggested in [Shrout and Fleiss \(1979\)](#) and in [Koo and Li \(2016\)](#).

The eight adjustments of each participant were averaged for each condition. To compute illusion strength, the reference line length (either 100 or 150 pixels) was subtracted from the mean adjustments. The results were subsequently divided by the same reference line length, thus turning them into illusion magnitudes as a proportion of the reference line length. Overadjustments and underadjustments are indicated by positive and negative illusion magnitudes, respectively. Analyses were performed in R ([R Core Team, 2018](#)).

To check for outliers, illusion magnitudes were standardized. Rather than the commonly used z -scores, we used modified z -scores, which are based on the median and the median absolute deviation ([Iglewicz & Hoaglin, 1993](#)). Modified z -scores were summed (in absolute values) for each participant across all conditions. Four participants were considered as outliers and removed from the dataset because the sum of their modified z -scores was outside the mean \pm 3 standard deviations (SDs) range.

A three-way repeated measures analysis of variance (ANOVA) was performed on the illusion magnitudes after outlier removal. We used Mauchly’s tests to verify sphericity assumptions and applied Greenhouse–Geisser correction in case of violations of sphericity. In both experiments, the alpha level for statistical significance was set to 0.05.

To determine whether merged illusion magnitudes can be predicted from non-merged illusion magnitudes, a path analysis was computed with the vertical inward wings and vertical outward wings conditions as outcomes (VerIn100, VerIn150, VerOut100, and VerOut150; endogenous variables) and all the other conditions as predictors (exogenous variables). In a path analysis, each predictor is regressed onto each outcome ([Beaujean, 2014](#)). The vertical inward wings (VerIn) conditions were expected to be strongly predicted from the vertical no wings (VerNone) and horizontal inward wings (HorIn) conditions. Likewise, we expected the vertical no wings (VerNone) and horizontal outward wings (HorOut) conditions to strongly predict the vertical outward wings (VerOut) conditions. Importantly, the path model considers the covariances among the exogenous variables (i.e., control and non-merged conditions).

We computed Pearson’s correlations between the illusion magnitudes of each condition and participants’

Test	<i>F</i>	<i>df</i> ₁	<i>df</i> ₂	η_p^2
Orientation	197.687***	1	305	0.393
Reference length	1431.100***	1	305	0.824
Wings	2023.802***	1.997	608.971	0.869
Orientation × reference length	0.267	1	305	<0.001
Orientation × wings	217.404***	1.974	601.993	0.416
Reference length × wings	7.435***	1.992	607.428	0.024
Orientation × reference length × wings	0.528	1.950	594.645	0.002

Table 1. Statistical results from the three-way repeated measures ANOVA for illusion magnitudes in [Experiment 1](#) with adjustable line orientation, reference line length, and type of wings as main factors. *** $p < 0.001$

age. We found no significant effects (see Table S1A in the Supplementary Material). We also computed Pearson's correlations between the illusion magnitudes of each pair of conditions. Finally, an exploratory factor analysis (EFA) was computed to explore the factors underlying the global structure of the dataset using the guidelines outlined in [Preacher, Zhang, Kim, and Mels \(2013\)](#). Because our aim was to identify factors reflecting only the variance shared between conditions (i.e., the common variance), we extracted factors with a common factor analysis. An oblique rotation (promax) was used in the EFA because we had no reason to preclude factors to correlate. Similar loadings are observed with an oblique or orthogonal rotation if the factors are uncorrelated (e.g., [Costello & Osborne, 2005](#)). We should mention that an EFA explores the structure underlying a dataset (i.e., the EFA looks for latent factors), whereas a path analysis explores the linear relationships between measured variables.

Results

Intrarater reliability

All conditions showed significant intraclass correlations even after Bonferroni correction was applied, highlighting consistent adjustments across all eight trials of the same condition ([Table 2B](#), diagonal). We should note, however, that the correlation coefficients indicate a large range of reliabilities from poor (HorNone100 ICC coefficient = 0.119) to large (VerIn150 ICC coefficient = 0.329) according to [Gignac and Szodorai \(2016\)](#) but only poor to moderate according to [Cohen \(1988\)](#).

Magnitudes of the illusions

The illusion magnitudes are illustrated in [Figure 2](#) and summarized in the Supplementary Material ([Table S2A](#)). A three-way repeated measures ANOVA was computed with the main factors of adjustable line orientation (horizontal or vertical), reference line length

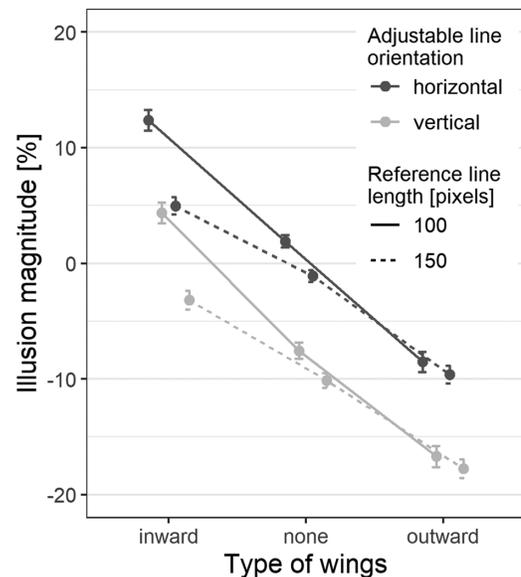


Figure 2. Illusion magnitudes ± 2 SEM in [Experiment 1](#) as a function of the type of wings (inward, none, or outward), the adjustable line orientation (dark gray, horizontal; light gray, vertical), and the reference line length (solid line, 100 pixels; dotted line, 150 pixels). Positive and negative magnitudes indicate overadjustments and underadjustments compared to the reference, respectively.

(100 or 150 pixels), and type of wings (inward, none, or outward). Statistical results are shown in [Table 1](#).

Path analysis

A path analysis was computed to determine whether the illusion magnitudes of the merged conditions can be predicted from the illusion magnitudes of the two non-merged conditions of which it was made. Standardized path coefficients are reported in [Table 2A](#). The two non-merged conditions that made each merged condition (i.e., the expected predictors) showed significant path coefficients. A Welch two-tailed t -test between the path coefficients that were expected to be high ($M = 0.317$, $SD = 0.100$) and the others ($M = 0.049$, $SD = 0.115$) resulted in a significant difference

(A)	HorNone100	HorNone150	HorIn100	HorIn150	HorOut100	HorOut150	VerNone100	VerNone150	r^2
VerIn100	-.044	-.050	.495***	.020	.069	.028	.212***	.200***	.513
VerIn150	-.072	.052	.271***	.253***	-.114*	.164**	.122*	.308***	.544
VerOut100	.054	-.110*	.176**	-.198***	.387***	.212***	.196***	.139**	.548
VerOut150	.001	-.038	.053	.068	.087	.306***	.096	.378***	.488

(B)	HorNone100	HorNone150	HorIn100	HorIn150	HorOut100	HorOut150	VerNone100	VerNone150	VerIn100	VerIn150	VerOut100	VerOut150
HorNone100	.119	.372	.380	.205	.374	.255	.437	.231	.302	.226	.356	.261
HorNone150		.171	.226	.461	.113	.446	.123	.423	.186	.408	.080	.324
HorIn100			.238	.484	.508	.429	.347	.172	.632	.485	.456	.352
HorIn150				.253	.221	.550	.023	.365	.336	.573	.105	.405
HorOut100					.194	.442	.413	.197	.442	.242	.643	.375
HorOut150						.230	.165	.396	.363	.516	.402	.554
VerNone100							.177	.417	.476	.305	.516	.356
VerNone150								.274	.374	.546	.305	.575
VerIn100									.277	.627	.523	.491
VerIn150										.329	.369	.635
VerOut100											.270	.584
VerOut150												.324

Table 2. Experiment 1. (A) Standardized path coefficients (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$) from the path analysis and communalities (i.e., the variance explained, r^2) of each merged condition. Gray shading indicates the non-merged conditions that made up every merged condition (i.e., expected predictors). (B) Diagonal (in gray): Intrarater reliabilities expressed as intraclass correlation coefficients ($ICC_{3,1}$) for each condition. All of them were significant. Triangle: Correlations between each pair of conditions expressed as correlation coefficients (Pearson’s r). A color scale from white to red reflects effect sizes from $r = 0$ to $r = 1$ (no negative effect sizes). Italics and bold font indicate significant results without and with Bonferroni correction, respectively. Black labels indicate control (non-illusory) conditions; purple labels indicate non-merged conditions; green labels indicate merged conditions

($t[13.668] = 6.298, p < 0.001, d = 2.483$). However, some unexpected, non-merged conditions significantly loaded on the merged conditions. For example, not only the VerNone100 and HorIn100 conditions significantly loaded on the VerIn100 condition (standardized path coefficients: VerNone100 = 0.212, $p < 0.001$; HorIn100 = 0.495, $p < 0.001$), but also the VerNone150 condition (standardized path coefficient = 0.200, $p < 0.001$). However, this might be due to the very strong correlations observed between conditions, especially between the two reference line lengths of a same combination of adjustable line orientation and type of wings.

We observed only weak path coefficients from the horizontal no wings conditions (HorNone100 and HorNone150) to the merged conditions, which was expected as these two conditions were considered to be control conditions. Indeed, the adjustable line was presented horizontally (unlike the vertical–horizontal illusion) and with no additional wings (unlike the Müller–Lyer illusion) in both control conditions. Table 2A also shows the communalities of each merged condition (i.e., the variance explained, or r^2). The uniqueness (i.e., the variance not explained by the exogenous variables, or $1 - r^2$) of each merged condition was between 0.452 and 0.512.

	RF1	RF2
HorNone100	0.427	0.099
HorNone150	-0.175	0.696
HorIn100	0.484	0.252
HorIn150	-0.231	0.889
HorOut100	0.759	-0.096
HorOut150	0.134	0.622
VerNone100	0.735	-0.146
VerNone150	0.134	0.523
VerIn100	0.584	0.215
VerIn150	0.120	0.732
VerOut100	0.948	-0.190
VerOut150	0.356	0.471

Table 3. Rotated factor loadings from an EFA after promax rotation for all conditions of [Experiment 1](#). A color scale from blue (negative loadings) to red (positive loadings) is shown. Black indicates control (non-illusory) conditions; purple indicates non-merged conditions; green indicates merged conditions. The two factors seem to relate to the two reference line length conditions

Correlations

Correlations between the illusion magnitudes of each pair of conditions are reported in [Table 2B](#) (triangle). Interestingly, most correlations were significant even after Bonferroni correction was applied for multiple comparisons. The effect sizes of the eight between-illusion correlations—correlations between a condition that refers to the vertical–horizontal illusion only (VerNone conditions) and a condition that refers to the Müller–Lyer illusion only (HorIn and HorOut conditions)—were significantly different from zero (one-sample two-tailed t -test: $t[7] = 5.244$, $p = 0.001$, $d = 1.854$).

Exploratory factor analysis

A two-factor model was suggested by a scree plot inspection (see Figure S1A in the Supplementary Material), whereas a parallel analysis suggested a five-factor model. Because the eigenvalues of factors three to five were well below one (RF1, 4.645; RF2, 0.996; RF3, 0.448; RF4, 0.309; RF5, 0.245), we retained the two-factor model. Loadings from the two-factor model are reported in [Table 3](#). The two factors together explained 49% of the variance (RF1, 25%; RF2, 24%). They were strongly inter-correlated ($r = 0.597$) and seem to relate to the two reference line length conditions.

Experiment 2

Methods

Participants

Participants were 100 visitors at a public event organized in Geneva, Switzerland. Of these, 98 participants were considered for further analysis (there were two outliers), with age ranging from 8 to 81 years (mean age, 36 years; 57 females). Adults signed informed consent and we obtained the assent of children as well as the consent of their parents. Participation was not compensated for in any form. Procedures were conducted in accordance with the tenets of the Declaration of Helsinki, except for the preregistration, and were approved by the local ethics committee.

Apparatus

The experiment was coded in MATLAB (R2014b, 64 bits; MathWorks, Natick, MA) using the Psychophysics Toolbox 3.1 (64 bits; [Brainard, 1997](#); [Pelli, 1997](#)) and was presented on a 24.5-inch BenQ monitor (Taipei, Taiwan) at a resolution of 1920×1080 pixels with a 60-Hz refresh rate. Participants sat approximately 60 cm away from the screen and used a Logitech LS1 computer mouse (Lausanne, Switzerland) to fulfill the task. A Minolta LS-100 luminance meter (Osaka, Japan) was used to calibrate the monitor before the experiment began. A room with artificial light conditions was used to run the experiment.

Stimuli

Stimuli were presented in white (~ 176 cd/m²) on a black background (~ 1 cd/m²). The line width was 4 pixels. Participants were tested with the vertical–horizontal (VH), Müller–Lyer (ML), and Ponzo (PZ) illusions ([Figure 3](#)). In addition, the vertical–horizontal and Müller–Lyer illusions were congruently (VH-ML con.) and incongruently (VH-ML inc.) merged. In the congruent VH-ML illusion, inward and outward wings were added to the horizontal and vertical segments of the vertical–horizontal illusion, respectively. In the incongruent VH-ML illusion, inward and outward wings were drawn at the extremities of the vertical and horizontal segments of the vertical–horizontal illusion, respectively. Similarly, the Müller–Lyer and Ponzo illusions were congruently merged (PZ-ML con.) by adding inward and outward wings to the lower and upper horizontal segments of the Ponzo illusion, respectively, and incongruently merged (PZ-ML inc.) when the contrary was true. In total, seven illusions were tested.

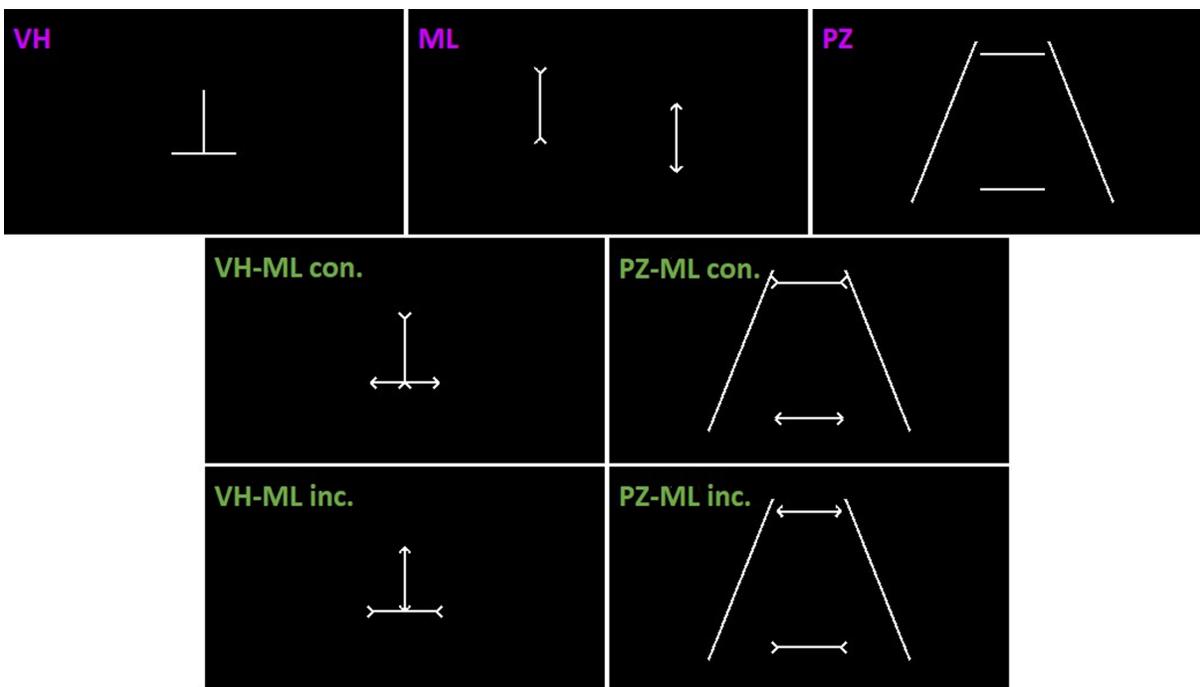


Figure 3. Seven illusions were tested in [Experiment 2](#) with an adjustment procedure: the vertical-horizontal (VH), Müller-Lyer (ML) and Ponzo (PZ) illusions, as well as the vertical–horizontal illusion congruently merged (VH-ML con.) or incongruently merged (VH-ML inc.) with the Müller-Lyer illusion and the Ponzo illusion congruently merged (PZ-ML con.) or incongruently merged (PZ-ML inc.) with the Müller-Lyer illusion. Two configurations were tested for each illusion (e.g., either the horizontal or vertical segment of the VH illusion was adjusted), making up 14 conditions. There were two trials of each condition. Purple indicates non-merged illusions; green indicates merged illusions.

In the VH and the congruent and incongruent VH-ML illusions, participants were asked to adjust the length of the horizontal segment to match the length of the vertical segment or to adjust the length of the vertical segment to match the length of the horizontal segment. In the ML illusion, we asked participants to adjust the length of the segment with inward wings to match the length of the segment with outward wings, and vice versa. In the PZ and the congruent and incongruent PZ-ML illusions, the task was to adjust the upper segment in length to match the lower one in length, and vice versa. Hence, each illusion was tested under two configurations, as one segment was in turn the reference or the adjustable segment. Adjustments were made by moving the computer mouse, and participants validated their adjustments by clicking on the left button of the computer mouse.

In each illusion, the reference segment was 8° long, and the length of the adjustable segment was pseudorandomly chosen to be between 2° and 14° at the beginning of each trial. In the ML, congruent and incongruent VH-ML, and congruent and incongruent PZ-ML illusions, the wings were 1° long and rotated by 45° compared to the orientation of the segments. In the VH and the congruent and incongruent VH-ML illusions, the center of the horizontal segment was

displayed 4° to the bottom of the midscreen, and the vertical segment was always touching it. In the ML illusion, the distance between the two segments was 16.9° . The centers of the segments with outward and inward wings were presented 2° to the top and bottom of the midscreen, respectively. In the PZ and the congruent and incongruent PZ-ML illusions, the two segments were 16.9° apart from each other. The distances between the upper and lower ends of the diagonals were 9° and 25° , respectively, and the total height of the illusion was 20° .

Procedure

The experimenter first explained the task to the participants. Three practice trials were run in order to ensure that the participants correctly understood the task: one ML trial, one VH-ML trial (either congruent or incongruent), and one PZ-ML trial (either congruent or incongruent). Each condition (7 illusions \times 2 configurations = 14 conditions) was tested twice, resulting in 28 trials per participant. The order of presentation of the 28 trials was chosen randomly by the computer. Participants were asked to refrain from any prior knowledge about visual illusions and to rely on their percepts only. The experimenter stayed

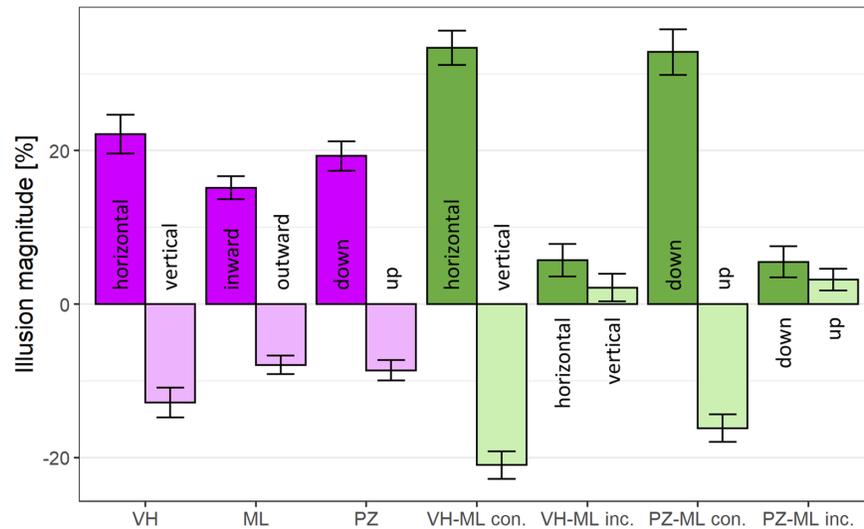


Figure 4. Illusion magnitudes ± 2 SEM in Experiment 2. Purple indicates non-merged conditions; green indicates merged conditions.

in the room during the experiment and answered any questions at any time. At the end of the experiment, the participants were shown their results and debriefed.

Data analysis

Intrarater reliability was assessed by computing intraclass correlations, as in Experiment 1. For each participant and each condition, the adjustments from both trials were averaged. As in Experiment 1, illusion magnitudes were computed. In order to make the magnitudes comparable across illusions, modified z -scores were computed for each illusion. At this stage, two outliers were detected and removed from the dataset. A path analysis was computed to determine whether the illusion magnitudes of the merged conditions (VH-ML con. hor., VH-ML con. ver., VH-ML inc. hor., VH-ML inc. ver., PZ-ML con. down, PZ-ML con. up, PZ-ML inc. down, and PZ-ML inc. up) can be predicted from the illusion magnitudes of the non-merged conditions (VH hor., VH ver., ML in., ML out., PZ down, and PZ up). As in Experiment 1, age did not significantly influence illusion magnitudes (see Table S1B in the Supplementary Material). Finally, we computed pairwise correlations and an EFA as in Experiment 1.

Results

Intrarater reliability

All intraclass correlations were significant even after Bonferroni correction was applied (Table 4B, diagonal).

Magnitudes of the illusions

As expected, all non-merged illusions showed an overadjusted and an underadjusted configuration (Figure 4 and see Table S2B in the Supplementary Material). For example, the horizontal segment of the vertical–horizontal illusion in which both segments are the same length is usually perceived to be shorter than the vertical segment and was therefore overadjusted (VH hor.: $M = 22.152\%$; $SEM = 1.266\%$). The vertical segment of the same illusion was underadjusted (VH ver.: $M = -12.829\%$; $SEM = 0.979\%$) because it is usually perceived to be longer than the horizontal segment. Interestingly, however, the illusion magnitudes are not symmetrical across both configurations. For example, the PZ down condition showed stronger absolute illusion magnitude compared to the PZ up condition (PZ down: $M = 19.301\%$, $SEM = 0.966\%$; PZ up: $M = -8.646\%$, $SEM = 0.661\%$). Stronger illusion magnitudes were observed for congruently merged illusions compared to non-merged illusions, while incongruently merged illusions revealed only very weak effects.

Path analysis

We conducted a path analysis with the non-merged conditions as predictors and observed that the illusion magnitude of a merged condition was strongly predicted by the illusion magnitudes of the non-merged conditions of which it was made (Table 4A). For example, adjusting the vertical segment of the incongruent VH-ML illusion (VH-ML inc. ver.) referred to both the vertical VH (VH ver.) and the inward ML (ML in.) conditions. Both the vertical VH (VH ver.) and the inward ML (ML in.) conditions

(A)	VH hor.	VH ver.	ML in.	ML out.	PZ down	PZ up	r^2
VH-ML con. hor.	.601***	-.113	.175*	.110	-.069	-.151	.452
VH-ML con. ver.	-.122	.606***	.053	.301***	.004	.024	.524
VH-ML inc. hor.	.606***	.037	.122	.093	.073	.050	.455
VH-ML inc. ver.	-.022	.530***	.194*	.100	-.105	.016	.341
PZ-ML con. down	.194*	-.029	.165*	.120	.451***	-.177	.437
PZ-ML con. up	-.003	.083	.119	.260**	-.221*	.484***	.446
PZ-ML inc. down	.235*	.143	-.037	.332***	.328***	-.166	.389
PZ-ML inc. up	.080	-.092	.104	-.008	-.079	.391***	.233

(B)	VH hor.	VH ver.	ML in.	ML out.	PZ down	PZ up	VH-ML con. hor.	VH-ML con. ver.	VH-ML inc. hor.	VH-ML inc. ver.	PZ-ML con. down	PZ-ML con. up	PZ-ML inc. down	PZ-ML inc. up
VH hor.	.496	-.360	.111	.158	.184	.285	.622	-.279	.649	-.191	.274	.118	.245	.220
VH ver.		.317	-.044	.032	-.046	.000	-.330	.656	-.187	.538	-.123	.098	.056	-.121
ML in.			.394	.017	.146	.074	.227	.020	.204	.156	.243	.123	.024	.135
ML out.				.311	.148	.045	.188	.304	.205	.102	.211	.253	.414	.010
PZ down					.647	-.359	.143	-.002	.196	-.096	.594	-.343	.468	-.186
PZ up						.428	.063	.005	.210	.066	-.266	.583	-.204	.450
VH-ML con. hor.							.311	-.283	.549	-.183	.399	.000	.130	.264
VH-ML con. ver.								.529	-.181	.576	-.131	.280	.221	-.256
VH-ML inc. hor.									.343	-.082	.297	.061	.247	.242
VH-ML inc. ver.										.297	-.015	.072	-.060	.006
PZ-ML con. down											.574	-.388	.364	.142
PZ-ML con. up												.486	.096	.097
PZ-ML inc. down													.623	-.320
PZ-ML inc. up														.501

Table 4. Experiment 2. (A) Standardized path coefficients (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$) from the path analysis and communalities (i.e., the variance explained, r^2) of each merged condition. Gray shading indicates the non-merged conditions that made up every merged condition (i.e., expected predictors). (B) Diagonal (in gray): Intrarater reliabilities expressed as intraclass correlation coefficients (ICC_{3,1}) for each condition. All of them were significant. Triangle: Correlations between each pair of conditions expressed as correlation coefficients (Pearson's r). A color scale from blue to red reflects effect sizes from $r = -1$ to $r = 1$ (white corresponds to $r = 0$). Italics and bold font indicate significant results without and with Bonferroni correction, respectively. Purple labels indicate non-merged conditions; green labels indicate merged conditions

showed significant standardized path coefficients (VH ver.: 0.530, $p < 0.001$; ML in.: 0.194, $p = 0.021$) to the vertical incongruent VH-ML condition (VH-ML inc. ver.). A Welch two-tailed t -test between the path coefficients that were expected to be high ($M = 0.351$, $SD = 0.180$) and the others ($M = 0.005$, $SD = 0.113$) resulted in a significant difference ($t[21.127] = 7.045$, $p < 0.001$, $d = 2.307$). The communalities of

each merged condition are reported in Table 4A. The uniqueness of each merged condition was between 0.476 and 0.767.

Correlations

Correlations were computed between the illusion magnitudes of each pair of conditions (Table 4B,

triangle). Only three correlations were significant between the two configurations of the same illusion (significant with Bonferroni correction for VH: $r = -0.360$, $p < 0.001$; PZ: $r = -0.359$, $p < 0.001$; PZ-ML con.: $r = -0.388$, $p < 0.001$).

Significant correlations were observed between some merged conditions and the non-merged conditions of which they were made. For example, the horizontal congruent and incongruent VH-ML conditions (VH-ML con. hor. and VH-ML inc. hor.) significantly correlated with the horizontal VH condition (VH hor.). Interestingly, only weak correlations were observed between the merged conditions and the Müller-Lyer illusion (ML in. and ML out. conditions). Hence, when two illusions merge, one illusion seems stronger than the other. This was also observed from the path analysis (Table 4A), where path coefficients to the merged conditions are weaker from the Müller-Lyer conditions than from the vertical–horizontal or Ponzo conditions.

A one-sample two-tailed t -test on the effect sizes of the 28 between-illusion correlations (VH vs. ML conditions, VH vs. PZ conditions, ML vs. PZ conditions, VH vs. PZ-ML conditions, and PZ vs. VH-ML conditions) was significantly different from zero ($t[27] = 4.029$, $p < 0.001$, $d = 0.761$), suggesting that a small proportion of the variance underlying the VH, ML, and PZ illusions is accounted for by a weak common factor.

Exploratory factor analysis

As in Experiment 1, a common factor analysis was computed to extract factors. A parallel analysis and scree plot inspection (see Figure S1B in the Supplementary Material) suggested a four-factor model explaining 56% of the total variance (RF1, 16%; RF2, 14%; RF3, 14%; RF4, 12%). The first and third factors highly loaded on the different conditions including the horizontal and vertical configurations of the VH illusion, respectively, whereas the second and fourth factors highly loaded on the conditions including the PZ and ML illusions, respectively (Table 5). The interfactor correlations showed small to medium effect sizes (RF1–RF2, $r = 0.024$; RF1–RF3, $r = -0.166$; RF1–RF4, $r = 0.164$; RF2–RF3, $r = 0.098$; RF2–RF4, $r = -0.203$; RF3–RF4, $r = 0.068$).

Discussion

We previously found strong correlations between variants of one illusion but weak correlations between different illusions (Cretienoud et al., 2019; Cretienoud et al., 2020; Grzeczowski et al., 2017; Grzeczowski et al., 2018), arguing for a multitude of factors

	RF1	RF2	RF3	RF4
VH hor.	0.628	0.158	−0.308	0.226
VH ver.	−0.150	0.009	0.699	0.048
ML in.	0.353	−0.054	0.157	−0.032
ML out.	0.182	0.131	0.141	0.460
PZ down	0.181	−0.542	0.000	0.340
PZ up	0.337	0.692	0.046	−0.119
VH-ML con. hor.	0.651	−0.041	−0.218	0.076
VH-ML con. ver.	−0.171	0.153	0.770	0.366
VH-ML inc. hor.	0.649	0.066	−0.120	0.183
VH-ML inc. ver.	0.131	−0.024	0.774	−0.106
PZ-ML con. down	0.547	−0.621	0.090	0.106
PZ-ML con. up	0.021	0.831	0.064	0.330
PZ-ML inc. down	0.008	−0.073	−0.037	0.856
PZ-ML inc. up	0.550	0.152	0.038	−0.422

Table 5. Rotated factor loadings from an EFA after promax rotation for all conditions of Experiment 2. A color scale from blue (negative loadings) to red (positive loadings) is shown. Purple indicates non-merged conditions; green indicates merged conditions. Bold numbers indicate loadings that are expected to be high under the hypothesis that the factors relate to conditions including horizontal VH, PZ, vertical VH, and ML, respectively

underlying illusions. Here, we investigated how these factors interact. To this aim, we merged two illusions.

Our results suggest that there are many illusion-specific factors that combine through combinatorics. Indeed, the results from the path analyses suggest that the illusion magnitudes of the merged conditions can be predicted from the illusion magnitudes of the non-merged conditions of which they were made. Our results support the hypothesis that the factors underlying visual illusions combine independently, which is well captured by the linear regressions of the path analyses. However, we have not tested for nonlinear models.

In Experiment 1, we expected to find two factors, one related to the Müller-Lyer and the other related to the vertical–horizontal illusion. We indeed observed two factors, but they seem to relate to the reference line lengths (i.e., the 100-pixel or 150-pixel reference length conditions). This discrepancy between the results and our expectations may be due to the strong positive correlations observed among all conditions (Table 2B, triangle), which may hide the factorial structure underlying the dataset. Also, the strong correlation between both factors and the large difference between the eigenvalues of the two factors suggest that a one-factor model could be an alternative to represent the underlying structure of this dataset.

In [Experiment 2](#), an exploratory factor analysis suggested a four-factor model to best represent the data. The four factors highly loaded on the different conditions that included the horizontal VH, PZ, vertical VH, and ML configurations. This finding corroborates our previous findings, which suggested that there seems to be one factor (or very few factors) per illusion, no matter the specific features of the illusion. Importantly, it seems that there are no proper factors for merged illusions but that merged conditions are well represented by specific combinations of the factors underlying non-merged conditions. Note that we tested more conditions with VH configurations than conditions with ML or PZ configurations. This may explain why two factors related to the VH illusion, while only one factor related to each of the ML and PZ illusion. Similarly, [Coren and colleagues \(1976\)](#) tested several illusions, including 11 variants of the Müller–Lyer illusion and one variant of the Ebbinghaus illusion. They found that two factors highly loaded on the two configurations of the Müller–Lyer variants, while only one factor highly loaded on the two configurations of the Ebbinghaus illusion.

The experimental conditions were different in both experiments. In [Experiment 1](#), the adjustable element was always compared to a reference line, which was not embedded in the context. In contrast, the reference line was always embedded in the context in [Experiment 2](#). As previously suggested, the effect of the vertical–horizontal illusion is weakened when both segments are not in contact (see, for example, [Hamburger & Hansen, 2010](#)), which may explain the absence of a factor specifically related to the VH illusion in [Experiment 1](#). In addition, unlike most illusion studies, the illusion stimulus was adjustable but the reference line was fixed in [Experiment 1](#), which may also influence the strength of our results. Indeed, the illusion magnitude varies as a function of the size of the illusion (see, for example, [Bulatov, Bertulis, Gutauskas, & Bulatova, 2010](#)).

Contrary to [Experiment 1](#), luminance and noise conditions were better controlled in [Experiment 2](#), and participants were all tested on the same setup. Still, the patterns of results from the path analyses are similar across both experiments; that is, a merged illusion magnitude is strongly predictable from the magnitudes of the two non-merged illusions that made it, suggesting that differences in the experimental designs do not interfere much in the results.

We expected better intrarater reliabilities in [Experiment 1](#) compared to [Experiment 2](#) because there were more repetitions of each condition in [Experiment 1](#) compared to [Experiment 2](#). Nevertheless, intrarater reliabilities were weaker in [Experiment 1](#) ($M = 0.238$, $SD = 0.063$) compared to [Experiment 2](#) ($M = 0.447$, $SD = 0.121$), which may reflect a weakness of the experimental design in [Experiment 1](#) (e.g., lack of

control for external conditions such as noise and other external distractions). Note that the moderate intrarater reliabilities may induce an upper limit on the pairwise correlations, thus leading to underestimated correlations.

In agreement with our previous studies ([Cretenoud et al., 2019](#); [Cretenoud et al., 2020](#)), we observed only weak, but mostly positive correlations between different illusions. For example, in [Experiment 2](#), the non-merged ML conditions only weakly correlated with the non-merged PZ conditions ($M = 0.103$, $SD = 0.052$), suggesting that the space underlying visual illusions is multifactorial. However, one-sample two-tailed t -tests on the effect sizes of the between-illusion correlations in both [Experiments 1](#) and [2](#) were significantly different from zero, suggesting that a small proportion of the variance underlying the illusions is accounted for by a weak common factor.

Contrary to our results, previous studies have often proposed taxonomies for visual illusions. For example, a factor highly loading on the vertical–horizontal, Müller–Lyer, and Ponzo illusions was described as a length factor ([Taylor, 1976](#)). However, more specific factors have been sometimes suggested to underlie specific illusions (such as the Müller–Lyer illusion; see [Coren et al., 1976](#)). The mixed results may come from discrepancies in the data analysis (e.g., how to determine the number of factors to retain), in the experimental design (e.g., whether or not it is a heterogeneous set of tasks, such that perceptual tasks other than illusions are tested in the same battery), or in the interpretation of the results.

Our research aim was to determine to what extent the magnitudes of the merged illusions can be explained by the magnitudes of the individual illusions and their underlying mechanisms. An extreme case could have been that performance could not be predicted at all—for example, because the merged illusions are not perceived as illusory anymore. Or, one illusion could dominate the other one or there could be some highly nonlinear combinations as often occurs in other fields of cue combination. For example, the joint effects of two grouping cues are sometimes quite different than the sum of the individual effects (e.g., [Ben-Av & Sagi, 1995](#); [Huang, 2005](#); but see [Kubovy & Van Den Berg, 2008](#)). Extreme cases are configural superiority effects where small changes in layout strongly change perception ([Pomerantz, 2003](#)). However, factors underlying the illusions tested here seem to combine independently.

We like to highlight that the illusion magnitudes of conditions with inward wings were smaller for the 150-pixel compared to the 100-pixel reference line in [Experiment 1](#) (it even changed sign in the case of the vertical adjustable line; see [Figure 2](#)). Further investigation is needed to understand this specific pattern of results, which contradicts the general assumption that a task becomes harder when

the stimulus size increases (e.g., see Bulatov et al., 2010).

When the VH illusion was merged with the ML illusion in Experiment 2, we observed stronger path loadings from the VH illusion than from the ML illusion. Similarly, stronger path loadings were observed from the PZ illusion than from the ML illusion when both were merged. The ML illusion therefore seems to have a weaker impact on the merged illusions than the VH and PZ illusions, which may come from the weaker susceptibilities to the ML illusion compared to VH and PZ illusions (see Figure 4 and Table S2B in the Supplementary Material).

In both experiments, the two configurations of an illusion were tested separately. The effects of the two configurations may be additive or superadditive; that is, the illusion magnitude is bigger than the sum of the two individual effects. Foster and Franz (2014) suggested that the Ebbinghaus and Müller–Lyer illusions show superadditive effects when a simultaneous adjustment procedure is used; that is, as the size of one target increases, the size of the other target decreases of a similar amount. On the other hand, they also suggested that there is no superadditivity when an independent adjustment procedure is used; that is, one target is adjustable and the other is the reference (see also Gilster & Kutz-Buschbeck, 2010). However, it was shown that the Ebbinghaus illusion magnitude was stronger than the sum of the parts even with an independent adjustment procedure (Grzeczowski et al., 2018; see also Duemmler, Franz, Jovanovic, & Schwarzer, 2008). Here, we observed asymmetrical illusion magnitudes across both configurations of the same illusion. For example, the inward ML condition yielded stronger illusion magnitudes compared to the outward ML condition in absolute values (Figure 4; see also Cretenoud et al., 2020; Yildiz, Sperandio, Kettle, & Chouinard, 2019). In addition, some illusions showed very weak correlations between the two configurations. For example, the correlation between the inward and outward ML conditions (ML in. and ML out.) in Experiment 2 was not significant ($r = 0.017$, $p = 0.871$). Similarly, we observed strong path loadings between the merged illusions and the non-merged illusions of which they were made, but only in the same configuration. For example, the vertical VH (VH ver.) and the outward ML (ML out.) conditions only weakly loaded on the congruently merged VH-ML illusion when the horizontal segment was adjustable (VH-ML con. hor.). These results suggest that both configurations of an illusion (either merged or non-merged) can have distinct effects at the individual level.

To our knowledge, merged illusions have been tested only once, in a study by Deręowski (2015), who tested the Ponzo illusion and merged it congruently and incongruently with the Müller–Lyer illusion. There was a significant difference between the effect

of the inversion phenomenon on the Ponzo illusion compared to the merged Ponzo/Müller–Lyer illusion, suggesting that the factors underlying the Ponzo and Müller–Lyer illusions have different origins, similarly to what we observed here and has been observed previously (Cretenoud et al., 2019; Grzeczowski et al., 2017). However, the author did not investigate the relationships between merged and non-merged illusion magnitudes.

To summarize, we previously observed only weak correlations between the susceptibility to different visual illusions but strong correlations between different variants of the same illusion (Cretenoud et al., 2019), suggesting that the structure of the visual space underlying illusions is multifactorial. Here, we investigated whether and how the factors for illusions interact. We tested merged illusions and observed that the susceptibility to a merged illusion is strongly predicted by the susceptibility to the two illusions that made it, suggesting that the factors for illusions combine independently. In comparison with Lavoisier’s famous chemistry principle, which claims that nothing is lost, nothing is created but everything is transformed (Lavoisier, 1789), we suggest that in the merged illusions, no factor is lost and no factor is created. Factors are not even transformed—they just combine.

Keywords: illusions, factors, individual differences

Acknowledgments

The authors thank Marc Repnow for technical support and Lukasz Grzeczowski for helpful comments. This work was supported by the project “Basics of visual processing: from elements to figures” (project no. 320030_176153/1) of the Swiss National Science Foundation (SNSF) and by a National Centre of Competence in Research (NCCR Synapsy) grant from the SNSF (51NF40-185897).

Commercial relationships: none.

Corresponding author: Aline F. Cretenoud.

Email: aline.cretenoud@epfl.ch.

Address: Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

References

- Aftanas, M. S., & Royce, J. R. (1969). A factor analysis of brain damage tests administered to normal subjects with factor score comparisons across ages. *Multivariate Behavioral Research*, 4(4), 459–481.

- Bargary, G., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). Individual differences in human eye movements: An oculomotor signature? *Vision Research*, *141*, 157–169.
- Beaujean, A. A. (2014). *Latent variable modeling using R: A step-by-step guide*. London: Routledge.
- Ben-Av, M. B., & Sagi, D. (1995). Perceptual grouping by similarity and proximity: Experimental results can be predicted by intensity autocorrelations. *Vision Research*, *35*(6), 853–866.
- Bosten, J. M., Goodbourn, P. T., Bargary, G., Verhallen, R. J., Lawrance-Owen, A. J., Hogg, R. E., . . . Mollon, J. D. (2017). An exploratory factor analysis of visual performance in a large population. *Vision Research*, *141*, 303–316.
- Bosten, J. M., & Mollon, J. D. (2010). Is there a general trait of susceptibility to simultaneous contrast? *Vision Research*, *50*(17), 1656–1664.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.
- Brascamp, J. W., Becker, M. W., & Hambrick, D. Z. (2018). Revisiting individual differences in the time course of binocular rivalry. *Journal of Vision*, *18*(7):3, 1–20, <https://doi.org/10.1167/18.7.3>.
- Bulatov, A. (2017). Weight positioned averaging in the illusions of the Müller-Lyer type. In A. G. Shapiro, & D. Todorović (Eds.), *The Oxford compendium of visual illusions* (pp. 159–163). New York: Oxford University Press.
- Bulatov, A., Bertulis, A., Gutasukas, A., & Bulatova, N. (2010). Stimulus size and the magnitude of the visual illusion of extent. *Human Physiology*, *36*(2), 164–171.
- Cao, T., Wang, L., Sun, Z., Engel, S. A., & He, S. (2018). The independent and shared mechanisms of intrinsic brain dynamics: Insights from bistable perception. *Frontiers in Psychology*, *9*, 589.
- Cappe, C., Clarke, A., Mohr, C., & Herzog, M. H. (2014). Is there a common factor for vision? *Journal of Vision*, *14*(8):4, 1–11, <https://doi.org/10.1167/14.8.4>.
- Chamberlain, R., Van der Hallen, R., Huygelier, H., Van de Cruys, S., & Wagemans, J. (2017). Local-global processing bias is not a unitary individual difference in visual processing. *Vision Research*, *141*, 247–257.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Coren, S., Girgus, J. S., Erlichman, H., & Hakstian, A. R. (1976). An empirical taxonomy of visual illusions. *Perception & Psychophysics*, *20*(2), 129–137.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Education*, *10*(7), 86–99.
- Cretenoud, A. F., Grzeczowski, L., Bertamini, M., & Herzog, M. H. (2020). Individual differences in the Müller-Lyer and Ponzo illusions are stable across different contexts. *Journal of Vision*, *20*(6):4, 86–99, <https://doi.org/10.1167/jov.20.6.4>.
- Cretenoud, A. F., Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, *19*(14):12, 1–21, <https://doi.org/https://doi.org/10.1167/19.14.12>.
- Deregowski, J. B. (2015). Illusions within an illusion. *Perception*, *44*(12), 1416–1421.
- Dobkins, K. R., Gunther, K. L., & Peterzell, D. H. (2000). What covariance mechanisms underlie green/red equiluminance, luminance contrast sensitivity and chromatic (green/red) contrast sensitivity? *Vision Research*, *40*(6), 613–628.
- Duemmler, T., Franz, V. H., Jovanovic, B., & Schwarzer, G. (2008). Effects of the Ebbinghaus illusion on children's perception and grasping. *Experimental Brain Research*, *186*(2), 249–260.
- Emery, K. J., Volbrecht, V. J., Peterzell, D. H., & Webster, M. A. (2017a). Variations in normal color vision. VI. Factors underlying individual differences in hue scaling and their implications for models of color appearance. *Vision Research*, *141*, 51–65.
- Emery, K. J., Volbrecht, V. J., Peterzell, D. H., & Webster, M. A. (2017b). Variations in normal color vision. VII. Relationships between color naming and hue scaling. *Vision Research*, *141*, 66–75.
- Foster, R. M., & Franz, V. H. (2014). Superadditivity of the Ebbinghaus and Müller-Lyer illusions depends on the method of comparison used. *Perception*, *43*(8), 783–795.
- Frenzel, H., Bohlender, J., Pinsker, K., Wohlleben, B., Tank, J., Lechner, S. G., . . . Lewin, G. R. (2012). A genetic basis for mechanosensory traits in humans. *PLoS Biology*, *10*(5), e1001318.
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78.
- Gilster, R., & Kutzt-Buschbeck, J. P. (2010). The Müller-Lyer illusion: Investigation of a center of gravity effect on the amplitudes of saccades. *Journal of Vision*, *10*(1):11, 1–13, <https://doi.org/10.1167/10.1.11>.
- Grzeczowski, L., Clarke, A. M., Francis, G., Mast, F. W., & Herzog, M. H. (2017). About individual differences in vision. *Vision Research*, *141*, 282–292.

- Grzeczowski, L., Roinishvili, M., Chkonina, E., Brand, A., Mast, F. W., Herzog, M. H., . . . Shaqiri, A. (2018). Is the perception of illusions abnormal in schizophrenia? *Psychiatry Research*, *270*, 929–939.
- Hamburger, K., & Hansen, T. (2010). Analysis of individual variations in the classical horizontal-vertical illusion. *Attention, Perception & Psychophysics*, *72*(4), 1045–1052.
- Hibbard, P. B., Bradshaw, M. F., Langley, K., & Rogers, B. J. (2002). The stereoscopic anisotropy: Individual differences and underlying mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 469–476.
- Huang, L. (2005). Grouping by similarity is mediated by feature selection: Evidence from the failure of cue combination. *Psychonomic Bulletin & Review*, *22*(5), 1364–1369.
- Iglewicz, B., & Hoaglin, D. (1993). *Volume 16: How to detect and handle outliers*. Milwaukee, WI: ASQC Press.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin, & O. P. John (Eds.), *Handbook of personality: Theory and research* (Vol. 2, pp. 102–138). New York: Guilford Press.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163.
- Kubovy, M., & Van Den Berg, M. (2008). The whole is equal to the sum of its parts: A probabilistic model of grouping by proximity and similarity in regular patterns. *Psychological Review*, *115*(1), 131–154.
- Lavoisier, A. L. (1789). *Traité élémentaire de chimie (Elementary Treatise on Chemistry)*. Paris: Cuchet.
- Milne, E., & Szczerbinski, M. (2009). Global and local perceptual style, field-independence, and central coherence: An attempt at concept validation. *Advances in Cognitive Psychology*, *5*, 1–26.
- Mollon, J. D., Bosten, J. M., Peterzell, D. H., & Webster, M. A. (2017). Individual differences in visual science: What can be learned and what is good experimental practice? *Vision Research*, *141*, 4–15.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.
- Peterzell, D. H. (2016). Discovering sensory processes using individual differences: A review and factor analytic manifesto. *Electronic Imaging*, *2016*(16), 1–11.
- Peterzell, D. H., Chang, S. K., & Teller, D. Y. (2000). Spatial frequency tuned covariance channels for red-green and luminance-modulated gratings: Psychophysical data from human infants. *Vision Research*, *40*(4), 431–444.
- Peterzell, D. H., Scheffrin, B. E., Tregear, S. J., & Werner, J. S. (2000). Spatial frequency tuned covariance channels underlying scotopic contrast sensitivity. In: *Vision science and its applications*, OSA Technical Digest, paper FC2. Washington, D.C.: Optical Society of America.
- Peterzell, D. H., Serrano-Pedraza, I., Widdall, M., & Read, J. C. A. (2017). Thresholds for sine-wave corrugations defined by binocular disparity in random dot stereograms: Factor analysis of individual differences reveals two stereoscopic mechanisms tuned for spatial frequency. *Vision Research*, *141*, 127–135.
- Pomerantz, J. R. (2003). Wholes, holes, and basic features in vision. *Trends in Cognitive Sciences*, *7*(11), 471–473.
- Preacher, K. J., Zhang, G., Kim, C., & Mels, G. (2013). Choosing the optimal number of factors in exploratory factor analysis: a model selection perspective. *Multivariate Behavioral Research*, *48*(1), 28–56.
- R Core Team. (2018). The R project for statistical computing. Retrieved from <https://www.r-project.org/>.
- Robinson, J. O. (1968). Retinal inhibition in visual distortion. *British Journal of Psychology*, *59*(1), 29–36.
- Roff, M. (1953). *A factorial study of tests in the perceptual area* (pp. 1–41), Psychometric Monographs 8. Madison, WI: Psychometric Society.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428.
- Taylor, T. R. (1974). A factor analysis of 21 illusions: The implications for theory. *Psychologia Africana*, *15*, 137–148.
- Taylor, T. R. (1976). The factor structure of geometric illusions: A second study. *Psychologia Africana*, *16*, 177–200.
- Thurstone, L. L. (1944). *A factorial study of perception*. Chicago, IL: The University of Chicago Press.
- Tulver, K. (2019). The factorial structure of individual differences in visual perception. *Consciousness and Cognition*, *73*, 102762.
- Tulver, K., Aru, J., Rutiku, R., & Bachmann, T. (2019). Individual differences in the effects of priors on perception: A multi-paradigm approach. *Cognition*, *187*(3), 167–177.
- Verhallen, R. J., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Bargary, G., & Mollon,

- J. D. (2017). General and specific factors in the processing of faces. *Vision Research*, *141*, 217–227.
- Webster, M. A., & MacLeod, D. I. A. (1988). Factors underlying individual differences in the color matches of normal observers. *Journal of the Optical Society of America*, *5*(10), 1722–1735.
- Wechsler, D. (2003). *Wechsler intelligence scale for children* (4th ed.). San Antonio, TX: The Psychological Corporation.
- Yildiz, G. Y., Sperandio, I., Kettle, C., & Chouinard, P. A. (2019). The contribution of linear perspective cues and texture gradients in the perceptual rescaling of stimuli inside a Ponzo illusion corridor. *PLoS One*, *14*(10), e0223583.

When illusions merge – Supplementary Material

Aline F. Cretenoud, Gregory Francis, & Michael H. Herzog

Correlations with age

We did not observe any significant correlations between the different conditions and age (Table S1), i.e., age does not seem to influence illusion magnitudes. We however like to mention that the age range was small in Experiment 1 (mean age: 22 years; min: 19 years; max: 27 years) compared to Experiment 2 (mean age: 36 years; min: 8 years; max: 81 years).

Table S1. Correlations between age and the illusion magnitudes of each condition tested in Experiments **(A)** 1 and **(B)** 2. None of them were significant (p values were not corrected for multiple comparisons). Black indicates control (non-illusory) conditions; purple indicates non-merged conditions; green indicates merged conditions.

(A) Exp. 1	HorNone100	HorIn100	HorOut100	VerNone100	VerIn100	VerOut100
r	-0.004	0.017	0.051	0.054	0.025	0.083
p	0.943	0.765	0.374	0.349	0.667	0.152
	HorNone150	HorIn150	HorOut150	VerNone150	VerIn150	VerOut150
r	-0.102	-0.041	0.036	0.057	0.020	0.067
p	0.076	0.483	0.533	0.327	0.727	0.246

(B) Exp. 2	VH hor.	VH ver.	ML in.	ML out.	PZ down	PZ up
r	0.079	-0.181	-0.067	-0.069	-0.095	0.107
p	0.441	0.075	0.514	0.501	0.351	0.295
	VH-ML con. hor.	VH-ML con. ver.	PZ-ML con. down		PZ-ML con. up	
r	0.049	-0.152	-0.094		0.077	
p	0.630	0.135	0.355		0.449	
	VH-ML inc. hor.	VH-ML inc. ver.	PZ-ML inc. down		PZ-ML inc. up	
r	-0.057	-0.145	-0.010		0.022	
p	0.574	0.154	0.919		0.829	

Illusion magnitudes expressed as a proportion of the reference line length

Table S2. Illusion magnitudes (as a percentage of error of the reference line length) summarized as mean (*M*), standard deviation (*SD*), standard error of the mean (*SEM*), and 95% confidence interval (95% *CI*) for each condition of **(A)** Experiment 1 (*n* = 306) and **(B)** Experiment 2 (*n* = 98). Black indicates control (non-illusory) conditions; purple indicates non-merged conditions; green indicates merged conditions.

(A) Experiment 1						
Adjustable line orientation	Type of wings	Reference line length [pixels]	<i>M</i>	<i>SD</i>	<i>SEM</i>	95% <i>CI</i>
horizontal	none	100	1.917	4.573	0.261	0.514
		150	-1.093	4.421	0.253	0.497
horizontal	inward	100	12.383	7.888	0.451	0.887
		150	4.983	6.656	0.381	0.749
horizontal	outward	100	-8.504	7.636	0.437	0.859
		150	-9.617	6.750	0.386	0.759
vertical	none	100	-7.548	6.162	0.352	0.693
		150	-10.115	5.629	0.322	0.633
vertical	inward	100	4.369	7.881	0.451	0.887
		150	-3.182	6.978	0.399	0.785
vertical	outward	100	-16.704	8.070	0.461	0.908
		150	-17.756	7.148	0.409	0.804

(B) Experiment 2					
Illusion	Configuration	<i>M</i>	<i>SD</i>	<i>SE</i>	95% <i>CI</i>
VH	horizontal	22.152	12.536	1.266	2.513
	vertical	-12.829	9.696	0.979	1.944
ML	inward	15.149	7.396	0.747	1.483
	outward	-7.923	6.058	0.612	1.215
PZ	down	19.301	9.560	0.966	1.917
	up	-8.646	6.541	0.661	1.311
VH-ML con.	horizontal	33.404	11.043	1.115	2.214
	vertical	-20.990	9.007	0.910	1.806
VH-ML inc.	horizontal	5.732	10.527	1.063	2.110
	vertical	2.151	8.799	0.889	1.764
PZ-ML con.	down	32.847	14.645	1.479	2.936
	up	-16.175	8.851	0.894	1.774
PZ-ML inc.	down	5.505	9.945	1.005	1.994
	up	3.180	7.108	0.718	1.425

Exploratory Factor Analyses (EFA)

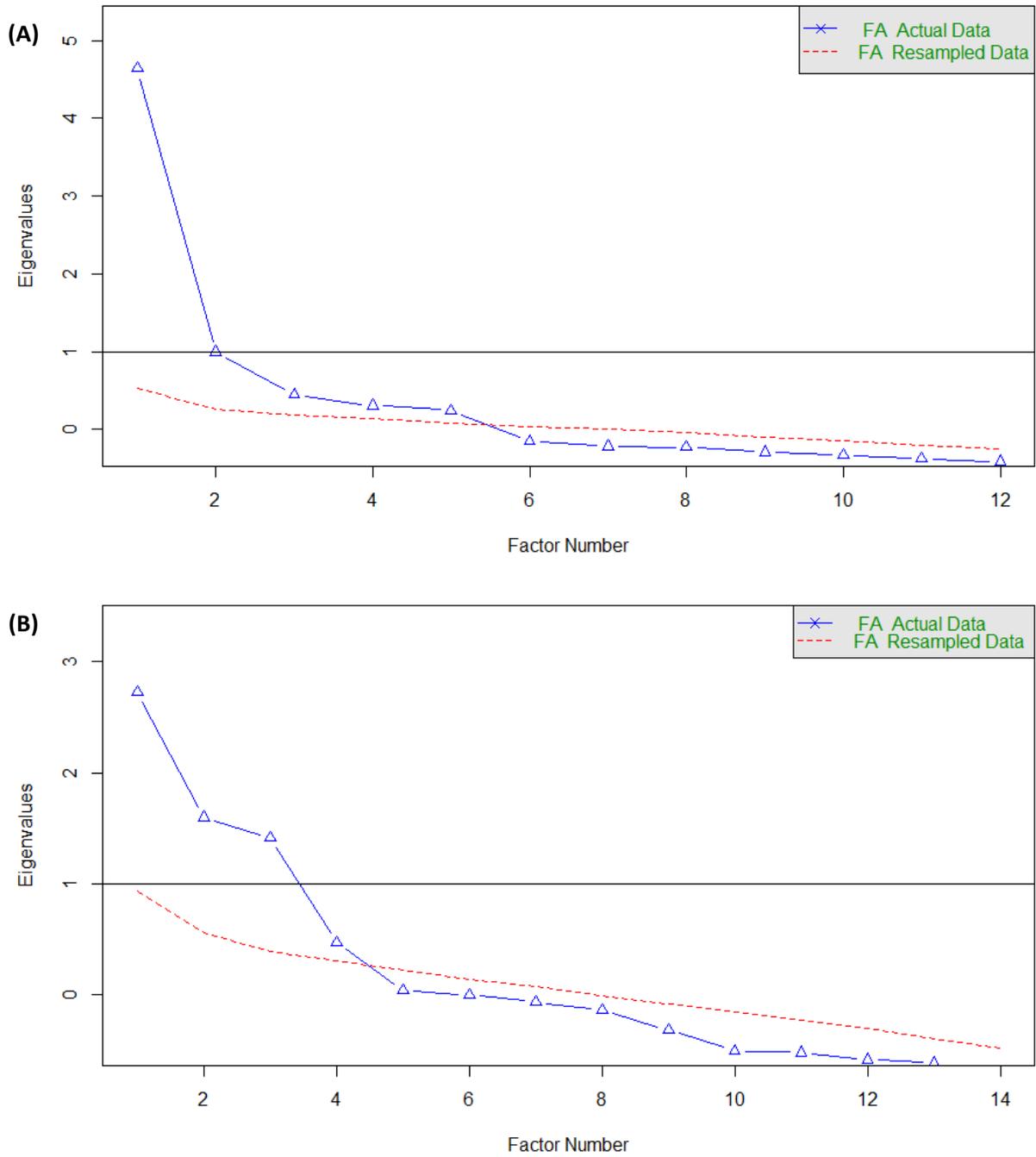


Figure S1. Scree plots from factor analyses of Experiments **(A)** 1 and **(B)** 2. A common factor analysis was computed to extract the factors. Eigenvalues are shown for the actual data (in blue) and for resampled data (in red). In Experiment 1, a two-factor model is suggested by scree plot inspection, while a parallel analysis suggested a five-factor model. In Experiment 2, a parallel analysis and scree plot inspection both suggested a four-factor model.

Appendix D

Cretenoud, A. F., Grzeczowski, L., Kunchulia, M., & Herzog, M. H (submitted). Individual differences in the perception of visual illusions are stable across eyes, time, and measurement methods.

This manuscript is currently under review at *Journal of Vision*.

Copyright: Reprinted with permission from ARVO

Individual differences in the perception of visual illusions are stable across eyes, time, and measurement methods

Aline F. Cretenoud¹, Lukasz Grzeczowski^{1, 2}, Marina Kunchulia^{3, 4} & Michael H. Herzog¹

¹*Laboratory of Psychophysics, Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland*

²*Allgemeine und Experimentelle Psychologie, Department Psychologie, Ludwig-Maximilians-Universität München, Germany*

³*Institute of Cognitive Neurosciences, Free University of Tbilisi, Tbilisi, Georgia*

⁴*Laboratory of Vision Physiology, Ivane Beritashvili Center of Experimental Biomedicine, Tbilisi, Georgia*

Abstract

Vision scientists have tried to classify illusions for more than a century. For example, some studies suggested that there is a unique common factor for all visual illusions. Other studies proposed that there are several subclasses of illusions, e.g., illusions of linear extent or distortions. We previously observed strong within-illusion correlations but only weak between-illusion correlations, arguing in favor of an even higher multifactorial space with – more or less – each illusion making up its own factor. These mixed results are surprising. Here, we examined to what extent individual differences in the perception of visual illusions are stable across eyes, time, and measurements. First, we did not find any significant differences in the magnitudes of the seven illusions tested with monocular or binocular viewing conditions. In addition, illusion magnitudes were not significantly predicted by visual acuity. Second, we observed stable individual differences over time. Last, we compared two illusion measurements, namely an adjustment procedure and a method of constant stimuli, which both led to similar individual differences. Hence, it is unlikely that the individual differences in the perception of visual illusions arise from instability across eyes, time, and measurement methods.

1 Introduction

Common factors are ubiquitous in everyday life. For example, there seems to be a strong common factor for cognition in healthy aging, i.e., cognitive abilities are reliably affected with age (e.g., Baltes & Lindenberger, 1997; Kiely & Anstey, 2017; Lindenberger & Ghisletta, 2009). Age-related changes of different cognitive functions, such as perceptual speed and reasoning skills, were indeed reported to significantly correlate. In analogy, a strong common factor for vision may be expected, i.e., it may be that an individual who performs better in one visual task compared to other individuals also performs better in other visual tasks, suggesting that there is a single monolithic structure underlying vision. However, the space underlying vision seems to be multifactorial, i.e., there seems to be no unique common factor for vision (for reviews, see Mollon et al., 2017; Tulver, 2019). For example, Cappe and colleagues (2014) only observed weak correlations between the performance in six basic visual paradigms, such as visual acuity and contrast detection, suggesting that an individual with good performance in one task does not necessarily show good performance in other tasks. A principal component analysis (PCA) revealed a first component explaining only 34% of the variability in the data (but see Bosten et al., 2017). Likewise, several factors were suggested to underlie individual differences in hue scaling (Emery et al., 2017a, 2017b), oculomotor tasks (Bargary et al., 2017), and binocular rivalry (e.g., Brascamp et al., 2018).

Common factors were proposed for visual illusions. For example, Thurstone (1944) observed a factor underlying geometric illusions. Similarly, Roff (1953) computed a factor analysis on 70 perceptual measures, which resulted in a single factor associated with visual illusions (see also Aftanas & Royce, 1969). However, illusions were also shown to be more heterogeneous. For example, Coren and colleagues (1976) claimed that illusions belong to two classes, namely the illusions of extent and the illusions of shape or direction (see also Robinson, 1968). Likewise, Taylor (1974, 1976) reported that several illusion measures were best represented by a four-factor model including illusions of length judgments and distortions of parallelism.

We previously observed strong within-illusion correlations but only weak between-illusion correlations, suggesting that there are illusion-specific factors. For example, the susceptibility to the Müller-Lyer illusion was only weakly correlated with the susceptibility to the Ponzo illusion (Cretenoud, Francis, et al., 2020; Cretenoud, Grzeczowski, et al., 2020). However, several variants of the Ebbinghaus illusion, differing in color, texture, shape, or contrast, were strongly inter-correlated, arguing in favor of a common mechanism for the Ebbinghaus illusion (Cretenoud et al., 2019). The same was true for nine other illusions tested with different contrast and orientation conditions.

These mixed results, i.e., whether there is one unique or several specific factors for visual illusions, may come from discrepancies in the data analysis and interpretation of the results or in the experimental design. For example, some studies used an adjustment procedure while others used a method of constant stimuli (Peterzell, 2020). Also, it may be that illusion magnitudes are not stable in space and time, leading to different results across studies.

Here, we wondered how stable individual differences in the perception of illusions are. We first investigated whether illusory percepts are reliable interocularly and whether illusion magnitudes differ as a function of visual acuity. Second, the susceptibility to several illusions was measured at different time points within a month to evaluate how stable individual differences in the perception of visual illusions are over time. Last, we tested two illusions, namely the Ebbinghaus and Müller-Lyer illusions, with both an adjustment procedure and a method of constant stimuli (2-AFC) to determine whether reliable individual differences are observed with different measurement methods.

2 Experiment 1

Monocular viewing condition was shown to significantly decrease the magnitude of an actual, real-world Ponzo illusion compared to binocular viewing condition, because of a reduced perception of depth cues following the elimination of stereopsis (Leibowitz et al., 1969). Here, we first examined whether the susceptibility to several two-dimensional visual illusions is different between eyes and for monocular and binocular presentations. Second, we investigated the role of visual acuity in the perception of visual illusions.

2.1 Methods

2.1.1 Participants

Fifteen students and collaborators of the Ecole Polytechnique Fédérale de Lausanne (EPFL, Switzerland) participated in this experiment (5 females, mean age: 24 years, age range: 18-52 years). Participants were naïve to the purpose of the experiment. Prior to the experiment, participants signed informed consent. Both monocular and binocular visual acuities were measured in a random order using the Freiburg visual acuity test (Bach, 1996). Participants were paid 20 Swiss Francs per hour. Procedures were conducted in accordance with the Declaration of Helsinki, except for pre-registration (§ 35), and were approved by the local ethics committee.

2.1.2 Apparatus

A BenQ XL2420T LCD monitor driven by a Windows-PC using Matlab (The MathWorks Inc.) and the Psychophysics toolbox (Brainard, 1997; Pelli, 1997) was used. Stimuli were presented at a 1920 x 1080 pixel resolution with a 60 Hz refresh rate. The distance to the screen was approximately 60 cm. During the monocular testing, participants covered one of their eyes with an eyepatch. Participants adjusted stimuli with a Logitech LS1 computer mouse. The experiment was conducted in the Laboratory of Psychophysics at EPFL, Switzerland.

2.1.3 Stimuli

Each participant was tested with seven illusions (Figure 1): two variants of the Ebbinghaus (EB and EB2), a Müller-Lyer (ML), and four variants of the Ponzo (PZ; a wider variant, PZw; a grid variant, PZg; and a corridor variant, PZc) illusion. Note that the corridor variant of the Ponzo illusion (PZc) has also previously been called a Ponzo “hallway” illusion (Grzeczowski et al., 2017, 2018).

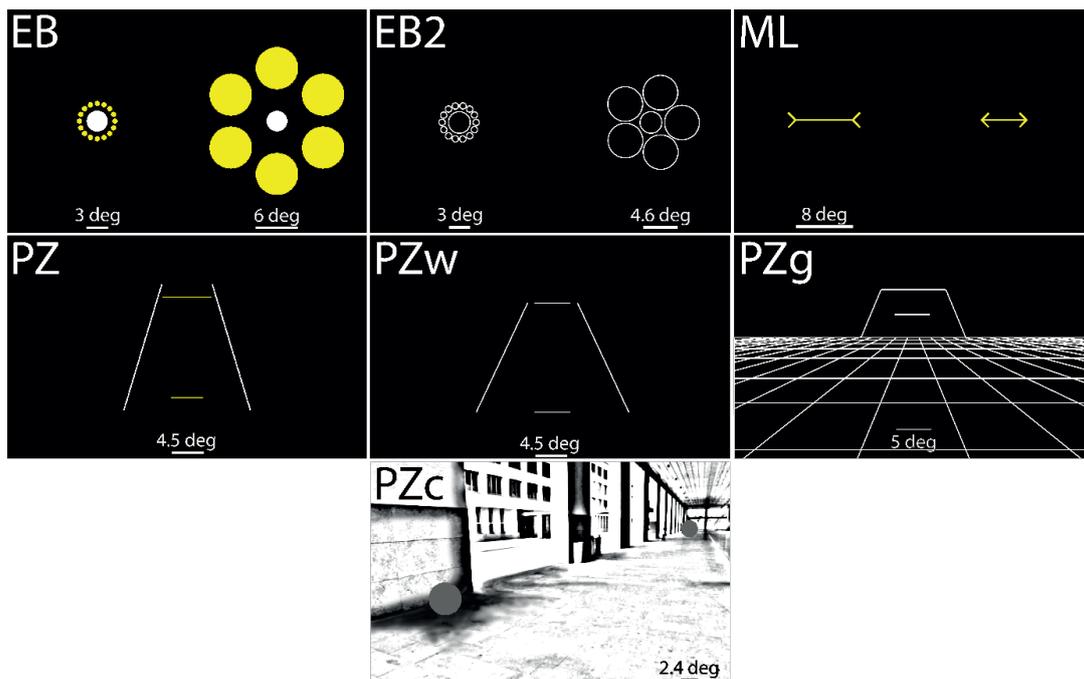


Figure 1. The seven illusions tested in Experiment 1: two variants of the Ebbinghaus (EB and EB2), the Müller-Lyer, and four variants of the Ponzo (PZ, PZw, PZg, and PZc). For each stimulus, participants were asked to adjust the size of a target to match the size of a reference element. For example, in the Ebbinghaus stimuli (EB and EB2), participants adjusted the size of the right central disk to match it with the left central disk using a computer mouse. Each stimulus was tested monocularly (left and right) and binocularly.

Each illusion was tested monocularly (left and right, separately) and binocularly. A method of adjustment was used, i.e., participants adjusted the size of a target to match the size of a reference on the screen by moving the computer mouse on the horizontal axis. In the Ebbinghaus illusions (EB and

EB2), participants adjusted the size of the right central disk to match the left central disk in size. Participants adjusted the length of the right horizontal segment to match the length of the left horizontal segment in the Müller-Lyer illusion (ML). In three variants of the Ponzo illusion (PZ, PZw, and PZg), participants adjusted the length of the upper horizontal segment to match the length of the lower horizontal segment. In the corridor variant of the Ponzo illusion (PZc), participants adjusted the size of the disk in the lower-left quadrant to match the size of the disk in the upper-right quadrant. The luminance was approximately 1 cd/m² for the black background, 30 cd/m² for grey, 146 cd/m² for yellow, and 176 cd/m² for white. For a more detailed description of the different illusions, please refer to Grzeczowski et al. (2017).

2.1.4 Procedure

The experimenter first explained the task to the participants who completed one trial of each illusion binocularly to familiarize with the method of adjustment. The 21 conditions (7 illusions x 3 eye conditions) were then tested twice (42 trials in total). The illusions were presented in the same order to all participants but the order of eye conditions, i.e., left, right, and binocular, was randomly set for each illusion and for each participant. The two trials of the same condition were always presented in a sequential manner, i.e., one trial after the other. There was no time restriction and no feedback.

2.1.5 Data analysis

Analyses were performed in Matlab (The Mathworks Inc.) and R (R Core Team, 2018). We assessed intrarater reliabilities for each illusion, i.e., the within-individual variation of illusion magnitudes across trials, by computing intraclass correlations (ICC) between the six adjustments of each illusion (3 eye conditions x 2 trials). First introduced by Fisher (1992) as an extension of the Pearson correlation coefficient, the concept of ICC was later developed as a measure of reliability within a class of data (Bartko, 1966; Shrout & Fleiss, 1979) rather than between different classes (the correlation between two different classes of data is usually computed as a Pearson correlation). ICCs are based on the analysis of variance (ANOVA) and therefore assume normally distributed data. In short, an ICC is computed as a ratio between the variance of interest (e.g., between-individual variance) and the total variance, i.e., the variance of interest and unwanted variance (e.g., within-individual variance or instrumental variation).

Several types of ICCs were developed to fit different experimental situations (e.g., Koo & Li, 2016; Liljequist et al., 2019). Here, we computed intraclass correlations of type (3,1) or ICC_{3,1}. The first subscript indicates that we computed a two-way mixed effects model (i.e., model 3), in which a random sampling of participants is assumed, while biases are assumed to be fixed (i.e., the only measure of

interest is the illusion magnitude, which was measured several times to assess the test-retest reliability). The second subscript indicates the type of selection used, i.e., each data point either represents a single measurement (i.e., type 1) or an average across several measurements (i.e., type k). We computed a 95% confidence interval around the correlation coefficient r , as suggested in Shrout & Fleiss (1979).

For each participant and each condition (7 illusions x 3 eye conditions = 21 conditions), we then averaged the adjusted values from both trials. The size of the reference element in each illusion was subtracted from the averaged values. Hence, the illusion magnitude is expressed as a size difference compared to the reference with positive and negative values indicating over- and under-adjustments, respectively. Correlations were computed between each pair of conditions. Cohen (1988) considered correlation coefficients of 0.1, 0.3, and 0.5 as small, medium, and large effect sizes, respectively. According to a meta-analysis by Gignac and Szodorai (2016), effect sizes of 0.1, 0.2, and 0.3 are considered as small, medium, and large, respectively.

To account for random variations in the baseline of participants and illusions, we computed mixed effects models (lmer R package). The fixed effects were eye condition, visual acuity, and sex. We accounted for the random effects of participants and illusions (random intercepts). The significance of each predictor in the model was assessed by computing likelihood ratio tests, which express the relative likelihood of the data given two competing models. The effect size was computed as a measure of explained variance with the random effect structure included (MuMIn R package).

2.2 Results

2.2.1 Intrarater reliabilities

For each illusion, we assessed intrarater reliability by computing an intraclass correlation. All ICC coefficients were significant even after correcting for inflated family-wise errors (EB: ICC coef. = 0.823, 95% CI [0.710, 0.913], $p < 0.001$; EB2: ICC coef. = 0.705, 95% CI [0.550, 0.846], $p < 0.001$; ML: ICC coef. = 0.465, 95% CI [0.285, 0.680], $p < 0.001$; PZ: ICC coef. = 0.489, 95% CI [0.309, 0.699], $p < 0.001$; PZw: ICC coef. = 0.628, 95% CI [0.458, 0.797], $p < 0.001$; PZg: ICC coef. = 0.686, 95% CI [0.527, 0.834], $p < 0.001$; PZc: ICC coef. = 0.669, 95% CI [0.506, 0.824], $p < 0.001$). The effect sizes are large according to Gignac and Szodorai (2016) and medium to large according to Cohen (1988), which suggests that the adjustments were consistent across eye conditions.

2.2.2 Illusion magnitudes

Mean illusion magnitudes are shown in Figure 2 and a table summarizing the illusion magnitudes is reported in the Supplementary File (Table S1). There was strong between-illusion differences, while only weak differences were observed between eye conditions. The corridor variant of the Ponzo illusion showed positive illusion magnitudes because the task was to adjust the disk that looks closer to the participant, unlike in the other variants of the Ponzo illusion (PZ, PZw, PZg), where the segment that looks further away (i.e., the top segment) was adjustable.

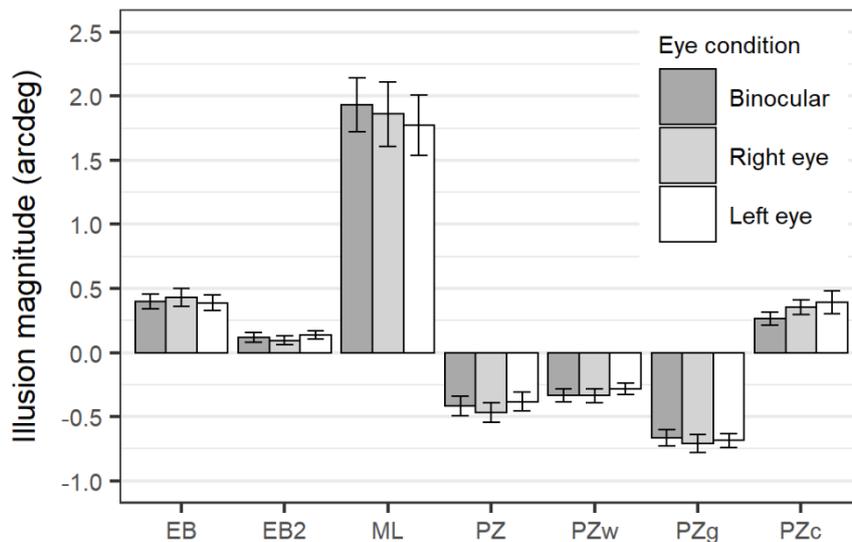


Figure 2. Illusion magnitudes \pm SEM (standard error of the mean) in Experiment 1 as a function of the eye condition (dark grey: binocular; light grey: monocular right; white: monocular left). Positive and negative magnitudes indicate over- and under-adjustments compared to the reference, respectively.

Correlations between pairs of conditions are reported in Table 1. In general, correlations were strong between different eye conditions of the same illusion and between different variants of an illusion (e.g., between the EB and EB2 conditions).

2.2.3 Mixed effects models

We tested for the effect of each predictor (eye condition, visual acuity, and sex) separately. First, and most importantly, a likelihood ratio test showed that the eye condition did not significantly improve the model ($\chi^2(2) = 0.153, p = 0.926$). The eye condition was therefore removed from the model. Second, visual acuity did not significantly improve the model ($\chi^2(1) = 0.140, p = 0.710$) and was hence also removed. Last, sex did not significantly improve the model ($\chi^2(1) = 0.310, p = 0.579$). Hence, the best model did not include any predictor but only random effects for participants and for illusions. This model accounted for 82.6% of the variance in the data ($r^2_c = 0.826$), which suggests that a large part of

the variability in the data is accounted for by individual and between-illusion differences. Our results therefore suggest that individual differences in the perception of visual illusions are stable across eye conditions and are not a function of visual acuity.

Table 1. Correlation coefficients (Pearson’s r) between each pair of conditions in Experiment 1. A color scale from blue to red reflects effect sizes from $r = -1$ to $r = 1$. Italics and bold font indicate significant results without ($\alpha = 0.05$) and with ($\alpha = 0.05/210$) Bonferroni correction, respectively. Binocular: binocular viewing condition; Right: monocular viewing condition with right eye; Left: monocular viewing condition with left eye.

		EB			EB2			ML			PZ			PZw			PZg			PZc		
		Binocular	Right	Left	Binocular	Right	Left	Binocular	Right	Left	Binocular	Right	Left	Binocular	Right	Left	Binocular	Right	Left	Binocular	Right	Left
EB	Binocular		.94	.94	.57	.35	.56	.20	.69	.50	-.20	-.15	.03	-.01	-.07	-.27	-.16	-.24	-.07	.68	.77	.84
	Right			.88	.49	.39	.52	.32	.66	.42	-.14	-.19	-.15	.07	-.13	-.41	-.02	-.23	-.08	.57	.66	.75
	Left				.58	.40	.61	.11	.61	.54	-.15	-.13	-.03	.09	-.12	-.25	-.15	-.29	-.11	.61	.69	.82
EB2	Binocular					.81	.70	.13	.45	.48	-.41	-.22	-.25	-.11	.06	-.25	-.21	-.31	-.25	.44	.62	.43
	Right						.76	.01	.10	.37	-.55	-.40	-.63	-.29	-.30	-.64	-.19	-.55	-.44	.10	.23	.07
	Left							-.36	.04	.36	-.75	-.65	-.32	-.39	-.52	-.66	-.51	-.72	-.44	.44	.50	.40
ML	Binocular								.59	.30	.45	.29	-.16	.32	.46	.26	.47	.43	.28	-.14	.05	.08
	Right									.55	.25	.26	.16	.36	.51	.20	.33	.36	.27	.36	.51	.61
	Left										-.28	-.16	-.04	-.05	-.05	-.03	.09	-.16	-.10	.06	.22	.30
PZ	Binocular											.80	.31	.74	.70	.61	.59	.83	.60	-.28	-.29	-.09
	Right												.47	.63	.74	.57	.59	.81	.55	-.27	-.22	-.10
	Left													.33	.54	.69	.16	.50	.57	.20	.03	.12
PZw	Binocular														.67	.54	.79	.70	.74	-.22	-.21	.14
	Right															.77	.58	.86	.72	-.09	-.06	-.02
	Left																.38	.70	.65	-.11	-.17	-.05
PZg	Binocular																	.71	.71	-.57	-.49	-.22
	Right																		.80	-.38	-.28	-.14
	Left																			-.33	-.30	-.05
PZc	Binocular																				.87	.76
	Right																					.86
	Left																					

3 Experiment 2

Bistable percepts were shown to change over time (Wexler et al., 2015). Likewise, illusion decrement, i.e., a decrease in the susceptibility to an illusion with repeated visual exposure, has been shown for a long time (e.g., Coren & Girgus, 1972; Judd, 1902; Predebon, 2006). Here, we wondered whether the individual differences in the perception of visual illusions vary over time.

3.1 Methods

3.1.1 Participants

Participants were 14 students of the Free University of Tbilisi, Georgia (7 females, mean age: 21 years, age range: 18-27 years). Participants signed informed consent prior to the experiment and were paid 10 GEL per hour. They were naïve to the purpose of the experiment. Procedures were conducted in accordance with the Declaration of Helsinki, except for pre-registration (§ 35), and were approved by the local ethics committee.

3.1.2 Apparatus

The experiment was performed on a Windows-PC with LCD display (ASUS VG248QE; screen resolution: 1920 x 1080 pixels; refresh rate: 60 Hz). Stimuli were generated with Matlab (The MathWorks Inc.) and the Psychophysics toolbox (Brainard, 1997; Pelli, 1997).

Participants were seated at 60 cm from the screen and used a Logitech LS1 computer mouse for the adjustments. The experiment was conducted at the Free University of Tbilisi, Georgia.

3.1.3 Stimuli

Eight illusions were tested in Experiment 2 (Figure 3): the Ebbinghaus (EB), horizontal-vertical (HV), Müller-Lyer (ML), Poggendorff (PD), Ponzo (PZ), two variants of the contrast (CS and CS2), and White (WH) illusions. As in Experiment 1, a method of adjustment was used, i.e., participants had to adjust the size (EB, HV, ML, PZ), position (PD), or shade of grey (CS, CS2, WH) of an element to match the size, position, or shade of grey, respectively, of a reference on the screen by moving the computer mouse. The reference and adjustable elements were the central disks in the Ebbinghaus illusion, the horizontal and vertical segments in the horizontal-vertical illusion, the vertical segments with inward- and outward-pointing arrows in the Müller-Lyer illusion, the left and right parts of the interrupted diagonal in the Poggendorff illusion, the upper and lower horizontal segments in the Ponzo illusion, the inside squares in one variant of the contrast illusion (CS), the two disks in the other variant of the contrast illusion (CS2), and the two columns of rectangles in the White illusion. There were two reference-dependent conditions for each illusion: one element (or series of elements, in the case of the White illusion) was in turn the reference or the adjustable element. For example, in the Ponzo illusion, the task was either to adjust the length of the upper horizontal segment so that it appeared to be the same length as the lower one, or to adjust the length of the lower horizontal segment so that it appeared to be the same length as the upper one. Stimuli were anti-aliased and lines were drawn with a width of

about 0.1 degree. Illusions are described in further detail in the Supplementary File. Black and white had a luminance of $\approx 1 \text{ cd/m}^2$ and $\approx 176 \text{ cd/m}^2$, respectively.

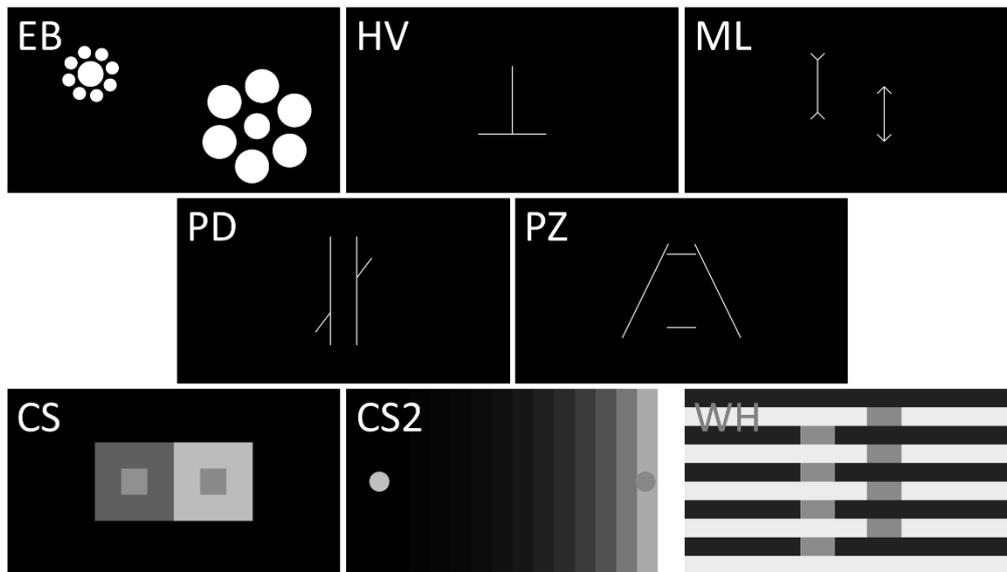


Figure 3. The eight illusions tested in Experiment 2: Ebbinghaus (EB), horizontal-vertical (HV), Müller-Lyer (ML), Poggendorff (PD), Ponzo (PZ), two variants of the contrast (CS and CS2), and White (WH). Each illusion was tested with two reference-dependent conditions (for example, either the horizontal or vertical segment of the HV illusion was adjusted), making up 16 conditions. An adjustment procedure was used and there were two trials of each condition.

3.1.4 Procedure

Participants were tested in 12 sessions with at least a 1-day break between two sessions. Participants were tested at days 1, 2, 3, 4, 5, 8, 9, 10, 11, 12, 19, and 33 at about the same time of the day. Each session lasted for about 10-20 minutes.

On the first day of testing, the experimenter first explained the task to the participants. Then, participants completed eight practice trials (one of each illusion with a reference-dependent condition randomly chosen by the computer) in a random order to familiarize with the adjustment procedure.

Each session consisted of 32 trials since each condition (8 illusions x 2 reference-dependent conditions = 16 conditions) was tested twice. The order of presentation of the 32 trials was randomly chosen by the computer for each participant and each session. Participants were asked to refrain from any prior knowledge about visual illusions. Importantly, no feedback was provided. The experimenter stayed in the experimental room during the experiment and answered any questions.

3.1.5 Data analysis

Intrarater reliabilities were computed as in Experiment 1. As a measure of illusion magnitude for each participant, for both reference-dependent conditions and for each session, the adjustment from both trials were averaged into a mean adjustment, from which the reference was subtracted, similar to Experiment 1. To test for the effect of reference-dependency, we computed Pearson correlations between the magnitudes of the two reference-dependent conditions of each illusion. Since the correlations were significant ($ps < 0.02$), we combined the two reference-dependent conditions of each illusion for further analyses. To estimate the effect of time (i.e., sessions) on the illusion magnitudes, we computed a repeated measures analysis of variance (rmANOVA) with illusion and session as within-subject factors. An rmANOVA catches any deviation from the mean. To examine how individual differences evolve with time (i.e., sessions), we computed Pearson correlations between pairs of sessions for each illusion and compared the correlation coefficients to simulated correlation coefficients.

We both used traditional null statistical hypothesis testing (NHST) and Bayesian statistics. Unlike NHST, inferences about null results are allowed with Bayesian statistics. A Bayes factor (BF₁₀, later referred to as BF) smaller than 0.33 indicates evidence in favor of the null hypothesis, while a BF larger than 3 indicates substantial support for the alternative hypothesis (Jeffreys, 1961). BFs lying between 0.33 and 3 are considered as inconclusive.

3.2 Results

3.2.1 Intrarater reliabilities

We computed intraclass correlations (ICCs) for the two reference-dependent conditions of each illusion across sessions. All ICC coefficients were significant even when applying Bonferroni correction for multiple comparisons (see Table S2 in the Supplementary File). However, we like to mention that the correlation coefficients varied from medium (EB large: ICC coef. = 0.369) to large (EB small: ICC coef. = 0.815) according to Cohen (1988), but all suggesting large effects according to Gignac and Szodorai (2016). The illusion magnitudes of each participant averaged across sessions are shown for all conditions in the Supplementary File (Figure S1).

3.2.2 Illusion magnitudes

Figure 4 shows the illusion magnitudes for both reference-dependent conditions of each illusion (see also Table S3 in the Supplementary File). As expected, all illusions showed an over- and an under-

adjusted condition. For example, the horizontal segment of the horizontal-vertical illusion was over-adjusted when compared to the vertical segment, and vice versa.

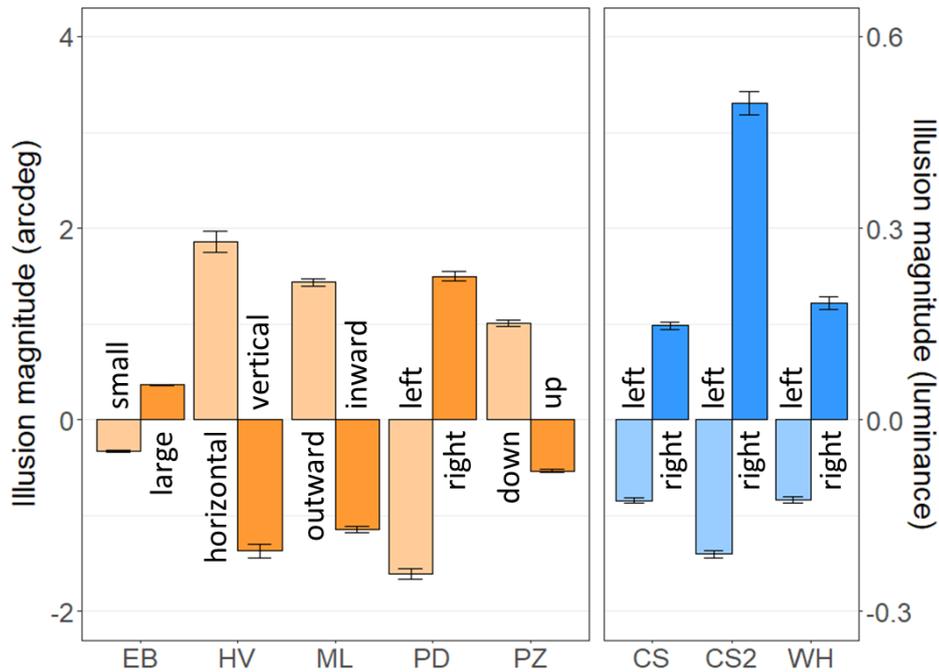


Figure 4. Illusion magnitudes \pm SEM (standard error of the mean) in Experiment 2 for both reference-dependent conditions of each illusion. Positive and negative magnitudes indicate over- and under-adjustments compared to the reference, respectively. Left panel (orange): illusions with adjustment of size, length, or position measured in units of arcdeg; right panel (blue): illusions with contrast adjustment measured in normalized luminance.

3.2.3 Are illusion magnitudes stable over time?

A repeated measures ANOVA was computed with illusion and session as within-subject factors. There was a significant interaction ($F(77, 1001) = 1.565, p = 0.002, \eta_p^2 = 0.107; BF = 3 \times 10^{270}$) and a significant main effect of illusion ($F(7, 91) = 28.14, p < 0.001, \eta_p^2 = 0.684; BF = 1 \times 10^{281}$) but the main effect of session was not significant ($F(11, 143) = 0.707, p = 0.73, \eta_p^2 = 0.052; BF = 7 \times 10^{-6}$). The data from the different illusions were further subjected to separate one-way repeated measures ANOVAs with session as within-subject factor, which revealed a significant main effect of session in the Ebbinghaus illusion and in one variant of the contrast (CS2) illusion (EB: $F(11,143) = 1.893, p = 0.045, \eta_p^2 = 0.127, BF = 0.788$; CS2: $F(11,143) = 2.679, p = 0.004, \eta_p^2 = 0.171, BF = 7.939$) but not in the other illusions (HV: $F(11,143) = 1.687, p = 0.082, \eta_p^2 = 0.115, BF = 0.435$; ML: $F(11,143) = 1.034, p = 0.420, \eta_p^2 = 0.074, BF = 0.073$; PD: $F(11,143) = 1.452, p = 0.156, \eta_p^2 = 0.100, BF = 0.226$; PZ: $F(11,143) = 0.895, p = 0.547, \eta_p^2 = 0.064, BF = 0.050$; CS: $F(11,143) = 0.780, p = 0.659, \eta_p^2 = 0.057, BF = 0.037$; WH: $F(11,143) = 1.473, p = 0.148, \eta_p^2 = 0.102, BF = 0.240$). Note that BFs were inconclusive in the case of the Ebbinghaus and horizontal-vertical illusions. Hence, there were differences in the mean magnitudes of the Ebbinghaus

and one variant of the contrast illusions across sessions (Figure 5), with a general increase in illusion magnitudes over time. However, this does not imply that individual differences are not stable. For example, it may be that all participants undergo a similar change in the susceptibility to one illusion over time (e.g., a 20% increase within a certain amount of time). In this case, individual differences would remain stable, which would show up as strong and significant between-session correlations for each illusion.

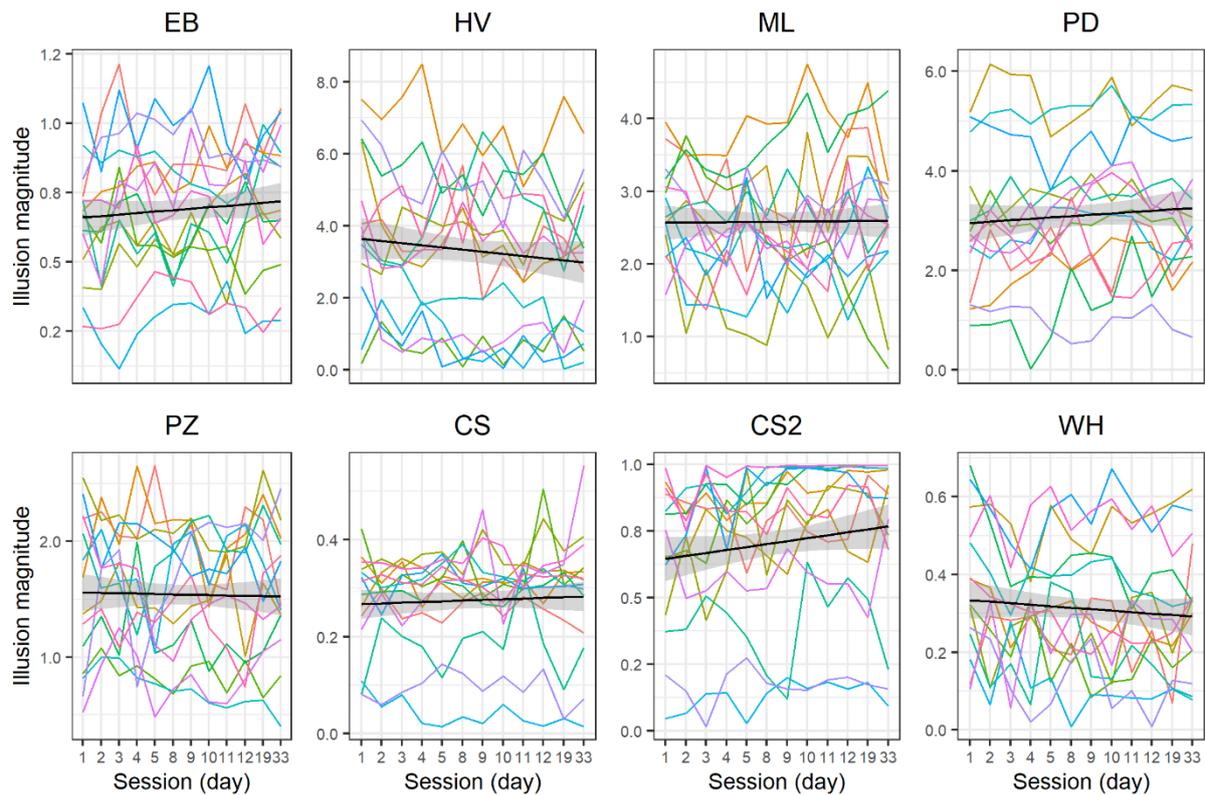


Figure 5. Illusion magnitudes plotted for each illusion (sub-panels) tested in Experiment 2 across sessions. Colors represent individuals. A linear regression (black line) was fitted with 95% confidence interval (shaded area). Illusion magnitudes are plotted in units of arcdeg for the EB, HV, ML, PD, and PZ illusions and in normalized luminance for contrast illusions (CS, CS2, WH).

3.2.4 Are individual differences in the perception of illusion magnitudes stable over time?

To examine how individual differences evolve with time, we computed between-session correlations for each illusion (Figure 6). Correlations were strong in general, and tended to weaken between pairs of sessions that were further apart in time (especially in the Müller-Lyer and Ponzo illusion).

Then, we wondered whether such patterns of correlations may result from stationary time series, i.e., stable individual differences over time. For this purpose, we simulated the individual illusion magnitudes from normal distributions centered on the behavioral data. In other words, for each

participant and each illusion, we computed the mean magnitude and standard deviation across sessions and randomly picked 12 values (for the 12 sessions) from a normal distribution centered and scaled on these values. Between-session correlations were then computed for each illusion and averaged across 10,000 simulations. We show the behavioral and simulated correlation coefficients in Figure 7 as a function of the time-lag (i.e., the time difference in days) between each pair of sessions.

If there were a time effect on the individual differences in the perception of visual illusions, correlations from simulated data should be stronger than correlations from behavioral data and the difference may strengthen with the time-lag (i.e., when two sessions are farther away in time). Paired two-tailed t -tests were computed between the behavioral and simulated correlation coefficients for each illusion. All of them resulted in non-significant differences with BF smaller than 0.33 (EB: $t(65) = 0.796$, $p = 0.429$, $d = 0.098$, BF = 0.183; HV: $t(65) = -0.036$, $p = 0.971$, $d = 0.004$, BF = 0.135; ML: $t(65) = -0.461$, $p = 0.647$, $d = 0.057$, BF = 0.149; PD: $t(65) = 0.069$, $p = 0.945$, $d = 0.009$, BF = 0.135; PZ: $t(65) = -0.303$, $p = 0.763$, $d = 0.037$, BF = 0.141; CS: $t(65) = 0.258$, $p = 0.797$, $d = 0.032$, BF = 0.139; WH: $t(65) = -0.283$, $p = 0.778$, $d = 0.035$, BF = 0.140), except for the second variant of the contrast illusion, which showed inconclusive BF (CS2: $t(65) = 1.879$, $p = 0.065$, $d = 0.231$, BF = 0.705). Overall, these results suggest that individual differences in visual illusions are stable across sessions.

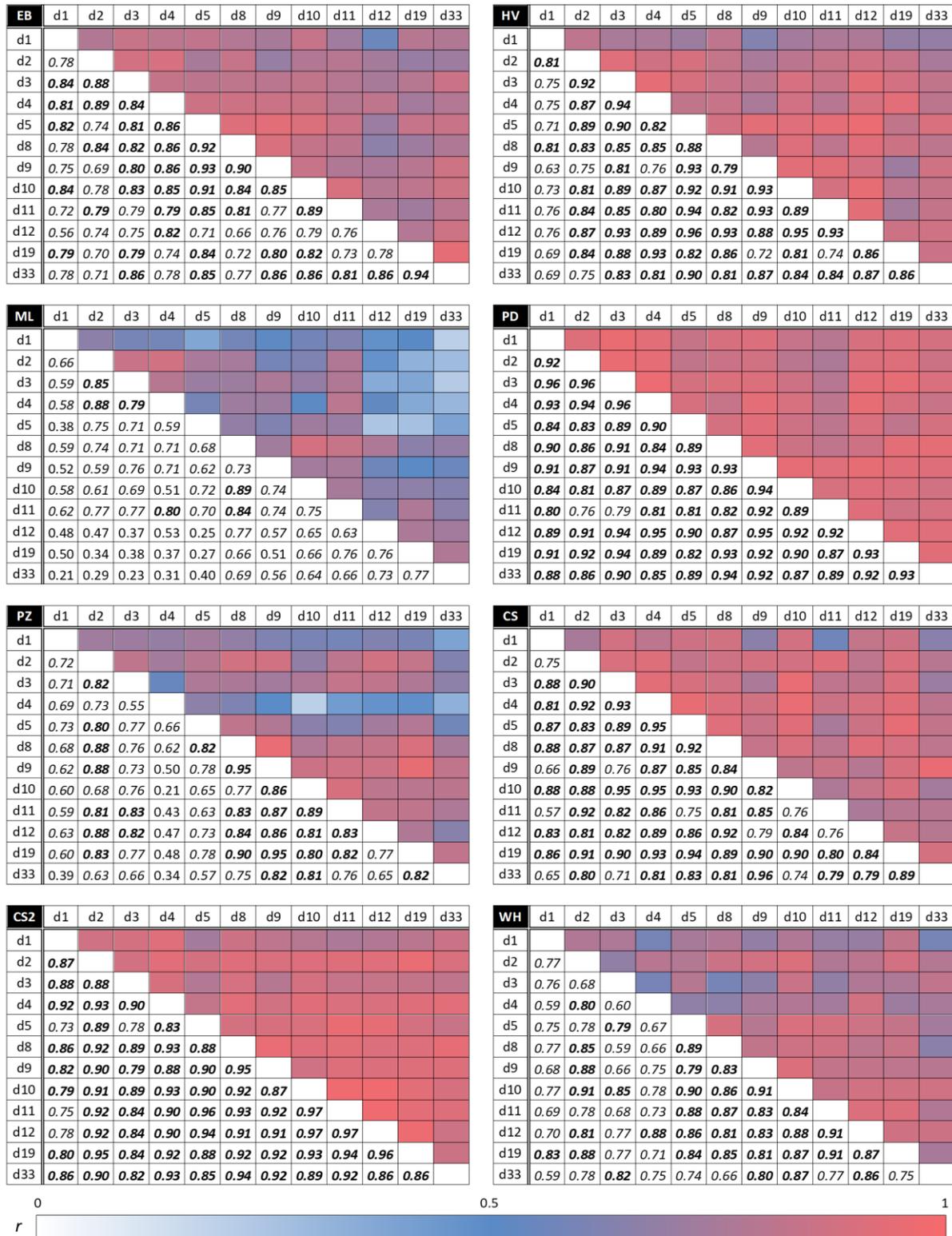


Figure 6. Between-session within-illusion correlation coefficients expressed as Pearson's r . Only positive correlations were observed. Upper triangles: A color scale from white to red reflects effect sizes from $r = 0$ to $r = 1$ (blue corresponds to $r = 0.5$). Lower triangles: Italics and bold font indicate significant correlation coefficients without ($\alpha = 0.05$) and with ($\alpha = 0.05/66$) Bonferroni correction, respectively.

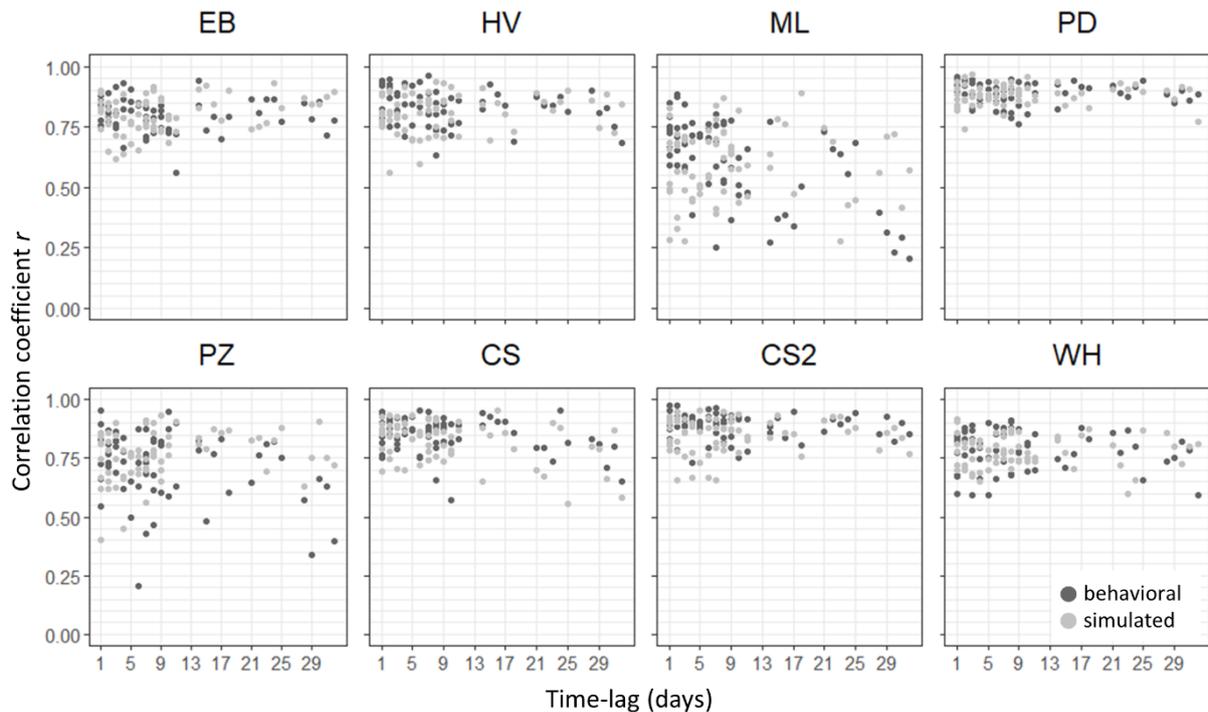


Figure 7. Correlation coefficients (r) as a function of the time-lag between sessions in days (i.e., the time difference) from behavioral data (in dark grey) and simulated data (in light grey) for each illusion. Correlations from simulated data were expected to be stronger than correlations from behavioral data under the hypothesis that individual differences in visual illusions vary with time. However, we did not observe any significant differences between simulated and behavioral correlation coefficients.

4 Experiment 3

We previously measured the susceptibility to visual illusions with an adjustment procedure (Cretenoud et al., 2019; Cretenoud, Francis, et al., 2020; Cretenoud, Grzeczowski, et al., 2020; Grzeczowski et al., 2017, 2018). However, forced choice response modalities, e.g., a method of constant stimuli, are usually thought to be more reliable (e.g., Todorović & Jovanović, 2018). Here, we compared the individual differences when two visual illusions were tested with an adjustment procedure and a method of constant stimuli.

4.1 Methods

4.1.1 Participants

Twenty students from the Ecole Polytechnique Fédérale de Lausanne (EPFL, Switzerland) and Université de Lausanne (UNIL, Switzerland) were tested (7 females, mean age: 23 years, age range: 19-28 years). None of the participants had taken part in Experiments 1 or 2 and all were naïve to the purpose of the experiment. Participants signed informed consent prior to the experiment and were

paid 20 Swiss Francs per hour and an extra amount of 5 Swiss Francs to compensate for the constraints related to Covid-19. Procedures were conducted in accordance with the Declaration of Helsinki, except for pre-registration (§ 35), and were approved by the local ethics committee.

4.1.2 Power analysis

We ran Bayesian simulations as a power analysis to determine the sample size needed for this experiment. Assuming a small effect size of $d = 0.1$, a Bayes factor smaller than 0.33 was computed for $n \geq 20$ (averaged across 1000 simulations). Hence, 20 participants were tested. A sensitivity analysis revealed that an effect size of $r = 0.552$ could be detected with $n = 20$, $\alpha = 0.05$ and 80% power.

4.1.3 Apparatus

Stimuli were presented on a BenQ XL2540 LCD monitor (screen resolution: 1920 x 1080 pixels, refresh rate: 60 Hz) using Matlab (The MathWorks Inc.) and the Psychtoolbox (Brainard, 1997; Pelli, 1997). The distance to the screen was approximately 60 cm. Participants used a Logitech M-BJ58 computer mouse for the adjustments. The experiment was conducted in the Laboratory of Psychophysics at EPFL, Switzerland.

4.1.4 Stimuli

Two illusions were tested: the Ebbinghaus (EB) and Müller-Lyer (ML) illusions (the metrics were the same as in Experiment 2 except when mentioned; see Figure 3 and the Supplementary File).

4.1.5 Procedure

Both illusions were tested with an adjustment procedure, as in Experiments 1 and 2, and with a method of constant stimuli (MCS). In the EB large condition, the central disk surrounded by large flankers was the target, while the central disk surrounded by small flankers was the reference. In the EB small condition, the target and the reference were reversed. Similarly, in the Müller-Lyer illusion, the target was either the vertical segment with inward-pointing arrows (ML inward condition) or the one with outward-pointing arrows (ML outward condition). Hence, one element was in turn the target or reference, i.e., each illusion was tested with two reference-dependent conditions.

For the adjustment procedure, each illusion and reference-dependent condition was tested twice (2 illusions x 2 reference-dependent conditions x 2 trials = 8). For example, in the Müller-Lyer illusion, the vertical segment with inward-pointing arrows was adjusted twice, and similarly for the vertical

segment with outward-pointing arrows. In addition, each illusion was tested twice without flankers (EB) or arrows (ML), i.e., as control conditions (2 illusions x 2 trials = 4).

Participants were asked to adjust the size of the target to match the size of the reference by moving the computer mouse on the vertical axis. At the beginning of each trial, the diameter (EB) and length (ML) of the target was randomly set in the range between 0° and 6.7° and between 2.1° and 19.9°, respectively.

For the method of constant stimuli, participants had to report whether the target was larger or smaller than the reference by using the up or down key on the keyboard, respectively. The target in each illusion and reference-dependent condition was tested 20 times with eight different increments (i.e., eight sizes; 2 illusions x 2 reference-dependent conditions x 8 increments x 20 trials = 640). For each illusion and each reference-dependent condition, the incremental range was centered on the mean adjustment, i.e., the point of subjective equality (PSE), from Experiment 2 and covered three times the absolute difference between the mean adjustment from Experiment 2 and the reference (see Table S4 in the Supplementary File). If a participant shows an illusion bias similar to the average bias observed in Experiment 2, 50% of all responses related to that illusion are expected to be “larger”. Similarly, 50% of the responses are expected to be “larger” if a participant has no illusory bias.

In addition, control conditions were tested with eight increments equally distributed around the reference value and covering the range between the mean adjustments for both reference-dependent conditions in Experiment 2 (2 illusions x 8 increments x 20 trials = 320; see Table S4 in the Supplementary File).

For both methods, the position of the target and reference on the screen was counterbalanced, i.e., the reference was displayed in the right half-screen in half of the trials and in the left half-screen in the other half of the trials. Moreover, the position of the target and reference along the y-axis was randomized with the screen size as limits and with the constraint that the target and reference were never at the same position along the y-axis to avoid any direct horizontal comparison between them. For the method of constant stimuli, a red square cue (0.5° in side, vertically centered) displayed on the left or right of the stimulus indicated which element was the target.

First, the experimenter explained the task to the participants. There were four warming up trials (2 illusions x 2 methods). Then, the experiment was split into eight blocks: each illusion (EB, ML) was tested with both methods (adjustment, MCS) and with and without context. The order of the eight blocks and the order of the trials within a block were randomized across participants. We asked participants to ignore any prior knowledge about illusions. Stimuli were shown until a response was

given, i.e., a mouse click with the adjustment procedure or a key press with the method of constant stimuli. There was no time restriction and no feedback.

4.1.6 Data analysis

As in Experiments 1 and 2, the illusion magnitudes are expressed as a size difference compared to the reference with positive and negative values indicating over- and under-adjustments, respectively.

As in Experiments 1 and 2, we computed intrarater reliabilities for each condition in the method of adjustment and then averaged both trials of each condition into a mean adjustment.

Psychometric curves were fitted for each condition with the method of constant stimuli to determine the PSE, i.e., the size of the target needed so that the participant perceived both the target and reference to be the same size (2 illusions x (2 reference-dependent conditions + 1 control) = 6 conditions). To this end, we summed the reports perceived as “larger” than the reference at each increment. Using a cumulative Gaussian function with a lapse rate of 0.02, we then defined the PSE as the size of the target corresponding to 50% of “larger” responses. We screened the data for invalid psychometric fits, i.e., no PSE was extracted when the mu or sigma parameter of the underlying Gaussian function was at the edge or outside of the search space, which was defined as a function of the incremental range. Based on that criterion, 2.5% of the PSEs were missing.

To check for outliers, the adjustment trials and the PSEs from the psychometric function for each condition (6 conditions) were then standardized by computing modified z-scores, which are more robust than the commonly used z-scores because it makes use of the median and median absolute deviation instead of the mean and standard deviation, respectively. As suggested by Iglewicz and Hoaglin, (1993), modified z-scores larger than 3.5 were considered as outliers. Based on that criterion, 4.2 and 1.7% of the data with the adjustment procedure and method of constant stimuli, respectively, was removed for further analyses. Outlying and missing data points were imputed using the ‘mice’ function from the mice R package with method ‘norm’ (Bayesian linear regression with 20 imputation samples).

To compare the illusion magnitudes from the adjustment procedure and the method of constant stimuli, we first accounted for the bias in the control conditions. Hence, for each illusion and each method, we subtracted the control condition from both reference-dependent conditions, e.g., we subtracted the EB control condition from the EB large and EB small conditions in the adjustment method, and similarly in the method of constant stimuli. Then, we computed paired *t*-tests and

Pearson correlations between the illusion magnitudes from both methods. As in Experiment 2, we both used traditional null statistical hypothesis testing (NHST) and Bayesian statistics.

4.2 Results

4.2.1 Intrarater reliabilities

The non-control conditions showed significant intraclass correlations with medium to large effect sizes, according to Gignac and Szodorai (2016; EB large: ICC coef. = 0.550, 95% CI [0.227, 0.764], $p = 0.005$; EB small: ICC coef. = 0.554, 95% CI [0.233, 0.766], $p = 0.005$; ML outward: ICC coef. = 0.420, 95% CI [0.060, 0.683], $p = 0.029$), except for the ML inward condition, which only showed a small and non-significant effect size (ICC coef. = 0.130, 95% CI [-0.251, 0.476], $p = 0.287$). The control conditions of both illusions showed weaker intraclass correlations (EB control: ICC coef. = 0.038, 95% CI [-0.335, 0.401], $p = 0.434$; ML control: ICC coef. = 0.377, 95% CI [0.010, 0.655], $p = 0.046$), which suggests that there is no strong measurement bias and that the residual bias is mainly due to noise.

4.2.2 Illusion magnitudes

Figure 8 shows the illusion magnitudes for all conditions (for more details, see Table S5 in the Supplementary File). Illusion magnitudes were approximately in the same range as in previous experiments (e.g., Experiment 2; see also Cretenoud, Francis, & Herzog, 2020; Cretenoud, Grzeczowski, Bertamini, & Herzog, 2020; Cretenoud et al., 2019). As expected, the control conditions led to very weak effects.

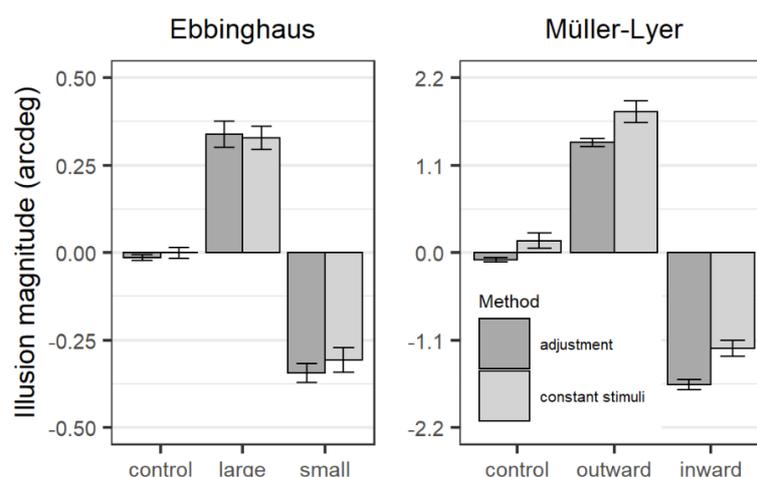


Figure 8. Illusion magnitudes \pm SEM (standard error of the mean) in Experiment 3 for the control condition and both reference-dependent conditions of each illusion (left panel: Ebbinghaus illusion; right panel: Müller-Lyer illusion). Positive and negative magnitudes indicate over- and under-adjustments compared to the reference, respectively. Dard grey: adjustment procedure; light grey: method of constant stimuli.

4.2.3 Comparing both methods

We accounted for the control bias by subtracting the control condition from the two reference-dependent conditions for each illusion in each method. Then, we computed paired t -tests to compare the means between the measures of illusion magnitudes from the adjustment procedure and from the method of constant stimuli. Results are reported in Table 2. None of the four tests revealed a significant difference. We also computed Bayesian paired t -tests, which resulted in Bayes factors in the range of 0.276 to 0.748. There was substantial support for the null hypothesis, i.e., both methods led to similar illusion magnitudes, in the EB small condition. The EB large, ML inward and ML outward conditions were associated with an inconclusive BF.

Table 2. Statistics from paired t -tests and correlations between the measures of illusion magnitudes from the adjustment procedure and method of constant stimuli, after accounting for the bias in the respective control conditions.

		EB large	EB small	ML inward	ML outward
Paired t -tests	t	-0.953	0.617	1.648	1.662
	p	0.353	0.544	0.116	0.113
	BF	0.347	0.276	0.735	0.748
Correlations	r	0.602	0.671	0.324	0.615
	p	0.005	0.001	0.164	0.004
	BF	10.355	27.743	1.021	12.200

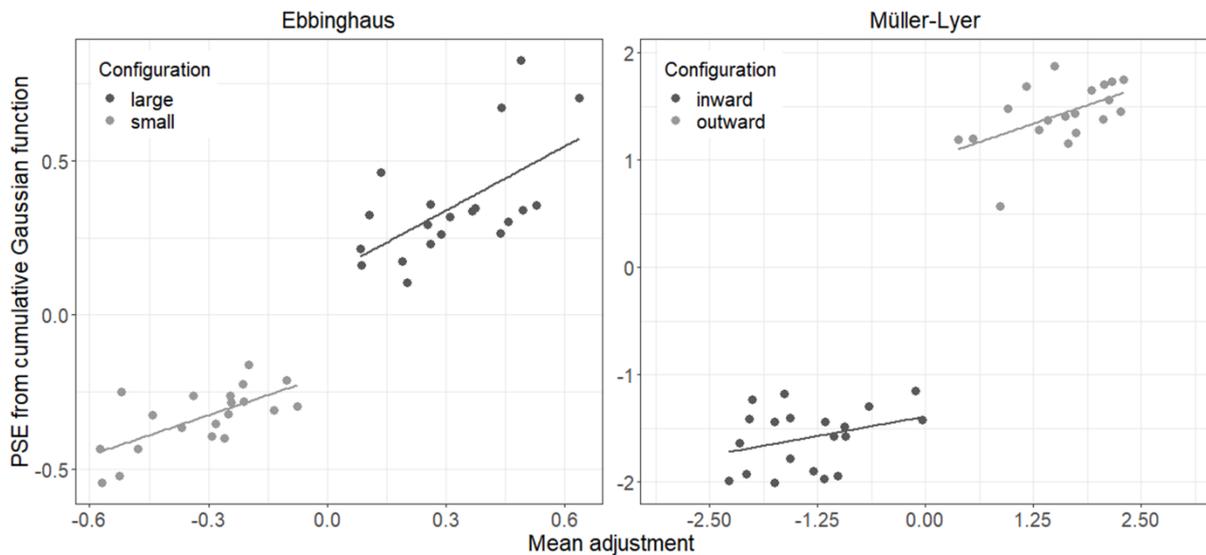


Figure 9. Correlations between the measures from the adjustment procedure and method of constant stimuli in units of arcdeg for both reference-dependent conditions of the Ebbinghaus (left panel) and Müller-Lyer (right panel) illusions, after accounting for the bias in the respective control conditions.

Correlations between the measures of illusion magnitudes from both methods resulted in large and significant effect sizes for the EB large, EB small, and ML outward conditions (Table 2 and Figure 9).

However, the ML inward condition resulted in a smaller and non-significant effect size of $r = 0.324$ with $BF = 1.021$.

5 Discussion

There are large individual differences in visual illusions. Here, we examined how stable individual differences in the perception of visual illusions are. First, we observed that illusory percepts are reliable interocularly and are not a function of visual acuity. Second, individual differences in the perception of visual illusions were reliable over time. Third, the methods of adjustment and constant stimuli showed reliable individual differences, except for one condition. Hence, our results suggest that the mixed results about factors for visual illusions, i.e., there is a unique common or several specific factors underlying illusions, do not result from unstable individual differences in the perception of visual illusions across eyes, time, and measurement methods.

Recent studies suggested a multifactorial structure underlying vision, such as in oculomotor tasks (Bargary et al., 2017), bistable paradigms (Brascamp et al., 2018; Cao et al., 2018; Wexler et al., 2015), local and global visual processing (Chamberlain et al., 2017), and in the use of expectations and knowledge priors (Tulver et al., 2019), which argues against a unique, general factor for vision as proposed previously (Halpern et al., 1999). Similarly, we previously claimed that there are illusion-specific factors (e.g., Cretenoud et al., 2019). For example, we observed strong correlations between different variants of the Ebbinghaus illusion, which differed in color, size, or shape, suggesting that there is a factor specific to the Ebbinghaus illusion. Unlike perceptual learning, which is orientation-specific, illusion magnitudes strongly correlated for different orientations.

Hence, we expected illusion magnitudes to be independent of eye and visual acuity. The results in Experiment 1 further support that claim as we observed no significant differences in the illusion magnitudes between monocular (left vs. right) and binocular viewing conditions. Similarly, visual acuity did not show any significant effect on the illusion magnitudes. Note that in Experiment 1 the two trials of each condition were presented sequentially, i.e., one after the other, which may inflate the intrarater reliabilities. However, the size of the adjustable target was randomly set at the beginning of each trial, which makes two trials (even presented one after the other) hardly comparable.

Stable individual differences were recently reported in the perception of different variants of the Ponzo illusion, which differed in context, e.g., with line-drawing or real-world perspective (Cretenoud, Grzeczowski, et al., 2020), suggesting that similar mechanisms are at hand when geometric or real-world depth cues are presented (but see Leibowitz et al., 1969; Wagner, 1977). In Experiment 1, we

similarly observed strong correlations between the magnitudes of the different variants of the Ponzo illusions. Note, however, that the correlations between the corridor variant (PZc) and the three other variants of the Ponzo illusion (PZ, PZw, and PZg) were weaker than between other pairs of Ponzo variants, especially in the monocular conditions (Table 1). In addition, there were strong correlations between the susceptibilities to the corridor variant of the Ponzo illusion (PZc) and the Ebbinghaus illusion (EB). Full and partial interocular transfer were previously observed in the Ponzo and Ebbinghaus illusions, respectively, suggesting that different mechanisms underlie the two illusions (Song et al., 2011). However, partial interocular transfer was observed in the effect of linear cues (but full transfer for texture gradients) in the perception of the corridor illusion (Yildiz et al., 2021), which suggests that the perceptual rescaling depends on the nature of the cues (i.e., lower- or higher-order features). Hence, in the present investigation, it may be that the corridor and the Ebbinghaus illusions rely on more similar features than the corridor and other variants of the Ponzo illusion.

For more than a century, the susceptibility to visual illusions was suggested to decrease with repeated visual exposures, which was called the illusion decrement (e.g., Coren & Girgus, 1972; Judd, 1902; Predebon, 2006). For example, a Müller-Lyer illusion decrement was observed under different viewing conditions, such as free viewing (e.g., Festinger et al., 1968) or at fixation (e.g., Day, 1962). In Experiment 2, there were no significant differences in the mean illusion magnitudes across time, except for the Ebbinghaus and a variant of the contrast illusion. Importantly, we do not claim that the illusion magnitudes are stable over time. Indeed, it may be that there are cumulative changes over time, as observed in paradigms with bistable stimuli (Wexler, 2018; Wexler et al., 2015). Instead, we wondered whether individual differences in the perception of visual illusions are stable. Hence, we were not interested in the changes in illusion magnitudes *per se*, but rather in comparing these changes across individuals over time. For example, we wondered whether a participant, who showed a stronger susceptibility to the Ebbinghaus illusion compared to other participants in the first testing session, showed this stronger susceptibility also a month later. We indeed found that this is the case, even though we do not preclude that individual differences in the perception of visual illusions may slightly vary over time. It seems that individual differences are largely stable over time, while illusion magnitudes are not. Note that we did not provide feedback to prevent learning. In addition, each illusion was tested only four times in each testing session (2 reference-dependent conditions x 2 trials), resulting in a rather short exposure to the stimuli.

In Experiment 2, the magnitudes of one variant of the Ebbinghaus illusion (EB2) were weaker compared to the other variant of the same illusion (EB), which may come from the close proximity of the flankers compared to the adjustable target (i.e., the size of the adjustable target could not be much increased

without overlapping with the flankers). Despite of this limitation, we observed rather strong correlations between both variants of the Ebbinghaus illusions, suggesting that the EB2 magnitudes are reliable. Interestingly, the magnitudes of the White and both contrast (CS and CS2) illusions strongly – but not always significantly – correlated (e.g., at d9 – CS-WH: $r = 0.473$, $p = 0.088$; CS2-WH: $r = 0.491$, $p = 0.074$; at d10 – CS-WH: $r = 0.615$, $p = 0.019$; CS2-WH: $r = 0.638$, $p = 0.014$), even though the White illusion is phenomenologically different compared to the contrast illusions.

Different methods of illusion measurements have been compared in the past. For example, Coren and Girgus (1972) compared five methods: an adjustment procedure, a method of reproduction, a selection from a graded series, and two subjective methods, i.e., rating scale and magnitude estimation. Unlike the subjective methods, the other three methods showed expected and significant effects in the Ebbinghaus and Müller-Lyer illusions. For example, the inward-pointing Müller-Lyer illusion magnitude was shown to significantly decrease with increasing angle of the fins. Similarly, the apparent size of the Ebbinghaus target significantly decreased when surrounding flankers increased in size. Nevertheless, the authors did not compute correlations between the illusion magnitudes from these different methods. The aim of Experiment 3 was to compare the individual differences in illusion magnitudes (rather than the illusion magnitudes *per se*) from two methods, namely an adjustment method and a method of constant stimuli. The illusion magnitudes measured with both methods were strongly correlated (with large associated BFs) in three out of four conditions. Only the ML inward condition did not show a significant correlation (BF = 1.021; inconclusive). Overall, the results suggest that individual differences are largely reliable across both methods. Similarly, Schwarzkopf and colleagues (2011) observed strong correlations between an adjustment procedure (with 8 trials per condition) and a method of constant stimuli in the estimation of the susceptibility to the Ebbinghaus and Ponzo corridor illusions.

Most illusion studies used forced choice responses (see King, Hodgekins, Chouinard, Chouinard, & Sperandio, 2017), such as the method of constant stimuli used in Experiment 3. However, there are several weaknesses associated with that method. First, the method of constant stimuli requires a large number of trials. For example, in Experiment 3, there were 960 versus 12 trials in the method of constant stimuli versus adjustment procedure. Repeated exposures may lead to fatigue or learning effects, as mentioned earlier. Note that in Experiment 3, the cue was presented together with the stimulus to decrease the proportion of trials answered wrongly (i.e., the lapse rate) due to a lack of attention following the large number of trials in the method of constant stimuli.

Illusions are a matter of perceptual bias. In addition, illusion magnitudes may be contaminated by decisional bias, for example in the method of constant stimuli. Indeed, individual differences in the

illusion magnitudes observed with that method may not only reflect a difference in the sensitivity to visual illusions but also in the participants' subjective, decisional criterion, e.g., a participant may always report that the reference stimulus is larger when he or she is unsure. This type of decisional bias is weakened in the adjustment procedure since participants are not asked to pick one out of two elements (target or reference). Note, however, that we do not claim that the adjustment procedure is completely free from response bias. For example, participants may use different strategies when they have to adjust an element that is obviously larger versus smaller than the reference.

Decisional biases may especially affect the validity of conclusions made when comparing a clinical population to a control group. Using a 2-AFC roving pedestal method, which reduces decisional bias (there is no clue about which element is the reference), Manning, Morgan, Allen, and Pellicano (2017) reported no substantial difference in the susceptibility to the Ebbinghaus and Müller-Lyer between autistic and typically developing children. However, using an adjustment procedure, the authors observed a weak evidence in favor of a group difference in the Müller-Lyer illusion (but not in the Ebbinghaus illusion), with autistic children showing larger illusion magnitudes compared to typically developing children. The authors suggested that the adjustment procedure is susceptible to atypical decisional bias, e.g., in clinical populations. In addition, they claimed that the Müller-Lyer illusion is more prone to decisional bias compared to the Ebbinghaus illusion, which may explain why we observed a weaker correlation between the Müller-Lyer inward illusion magnitudes when tested with an adjustment procedure and a method of constant stimuli (Experiment 3). However, Manning and colleagues (2017) did not test the method of constant stimuli (2-AFC) as we did. Note that negative relationships between autistic-like traits and susceptibility to visual illusions were previously reported (e.g., Chouinard et al., 2013; Happé, 1996; but see Chouinard et al., 2016), while similar patterns of between-illusion correlations were observed in patients with schizophrenia and healthy controls (Grzeczowski et al., 2018). In contrast, the personality dimension of schizotypy was shown to correlate with the likelihood of an individual to see meaning in a noisy, meaningless image (Partos et al., 2016).

More than being playful (e.g., when testing children), the adjustment procedure as we used here and previously (Cretenoud et al., 2019; Cretenoud, Francis, et al., 2020; Cretenoud, Grzeczowski, et al., 2020; Grzeczowski et al., 2017) also has weaknesses. For example, it may be argued that two trials of each condition are too few. Indeed, a mis-click or poor attention paid to one of the trials biases the average, therefore leading to a poor estimate of the illusion magnitude. However, intrarater reliabilities were strong in general in the adjustment procedure. In addition, the results from Experiment 3 showed that individual differences in the perception of visual illusions are in general similar across both methods, emphasizing that an adjustment procedure is amenable to replace a

method of constant stimuli when testing healthy participants. Note that both methods were non-speeded, i.e., the stimuli were shown until a response was given.

In previous publications, we found that there are only weak correlations between the susceptibility to different illusions. For example, out of 15 between-illusion correlations, only one was significant when 144 participants were tested (Grzeczowski et al., 2017). Similarly, we previously reported only 65 out of 720 significant between-illusion correlations (Experiment 2 in Cretenoud et al., 2019). Effect sizes were in most cases rather small (e.g., correlation coefficients r ranged between -0.12 and 0.23 in Grzeczowski et al., 2017). These results are surprising since in the latter publication, all illusions were spatial illusions, which are often implicitly or explicitly assumed to share the same mechanism (e.g., Coren et al., 1976; Ninio, 2014; Piaget, 1961). However, a shared mechanism should have led to substantial correlations. The weak correlations cannot be explained by large intraobserver variability because intrarater reliabilities were mostly significant, as in the present study. In addition, in Cretenoud et al. (2019), we found large within-illusion correlations for different luminance conditions including isoluminant ones, for 19 different variants of the Ebbinghaus illusion including two conditions with rotating flankers, and when illusions were tested under different orientations. Finally, individual differences were stable when illusions were presented as line drawings or within natural scenes like train tracks (Cretenoud, Grzeczowski, et al., 2020). It seems that each illusion makes up its own factor.

Here, we showed that individual differences in a monocular and binocular viewing conditions are robust, showing that differences between eyes are of little relevance. Hence, the mechanisms underlying illusions seem to occur after binocular fusion. In addition, we backed up our previous results by showing that individual differences in the perception of visual illusions are independent of the methods used and over time. Our results provoke the questions to what extent taxonomies of illusions are useful and what we can learn from detailed analysis of mechanisms underlying visual illusions, since these mechanisms seem to be rather idiosyncratic.

Our findings about illusions nicely fit into a larger picture about common factors for vision in general because there are not only many factors for illusions but also for many visual paradigms. For example, individual differences in contrast sensitivity (Peterzell, 2016; Peterzell et al., 2000), hue scaling (e.g., Emery et al., 2017a, 2017b), color matching (Webster & MacLeod, 1988), stereopsis (Peterzell et al., 2017), and in the effects of priors (Tulver et al., 2019) were suggested to rely on several specific factors. All these studies clearly argue against the widespread intuition that there are only a few mechanisms behind vision, which just need to be unearthed. It rather seems that we are dealing with a plethora of idiosyncratic mechanisms, which provokes the question to what extent the study of detailed mechanisms can lead to a unified theory of visual processing.

To summarize, we showed that the individual differences in the perception of visual illusions are reliable interocularly (Experiment 1), over time (Experiment 2), and when measured with an adjustment procedure or a method of constant stimuli (Experiment 3). Hence, the mixed results previously reported, i.e., there is a unique or several specific common factors for illusions, are unlikely related to unstable individual differences across eyes, time, and measurement methods. Future studies may examine the effect of other differences in the experimental design of the studies, such as speeded versus non-speeded tasks, and in the statistical methods used to determine the number of factors to extract.

6 Acknowledgments

We thank Marc Reppow for technical support and Pornbhussorn Kanchanakanok for helping with data collection (Experiment 1). This work was supported by the project “Basics of visual processing: from elements to figures” (project no. 320030_176153/1) of the Swiss National Science Foundation (SNSF) and by a National Centre of Competence in Research (NCCR Synapsy) grant from the SNSF (51NF40-185897).

7 References

- Aftanas, M. S., & Royce, J. R. (1969). A factor analysis of brain damage tests administered to normal subjects with factor score comparisons across ages. *Multivariate Behavioral Research*, *4*(4), 459–481.
- Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging*, *12*(1), 12–21. <https://doi.org/10.1037/0882-7974.12.1.12>
- Bargary, G., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). Individual differences in human eye movements: An oculomotor signature? *Vision Research*, *141*, 157–169. <https://doi.org/10.1016/j.visres.2017.03.001>
- Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological Reports*, *19*(1), 3–11. <https://doi.org/10.2466/pr0.1966.19.1.3>
- Bosten, J. M., Goodbourn, P. T., Bargary, G., Verhallen, R., Lawrance-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). An exploratory factor analysis of visual performance in a large population. *Vision Research*, *141*, 303–316. <https://doi.org/10.1016/j.visres.2017.02.005>
- Brascamp, J. W., Becker, M. W., & Hambrick, D. Z. (2018). Revisiting individual differences in the time course of binocular rivalry. *Journal of Vision*, *18*(7), 1–20. <https://doi.org/10.1167/18.7.3>
- Cao, T., Wang, L., Sun, Z., Engel, S. A., & He, S. (2018). The independent and shared mechanisms of intrinsic brain dynamics: insights from bistable perception. *Frontiers in Psychology*, *9*:589. <https://doi.org/10.3389/fpsyg.2018.00589>
- Cappe, C., Clarke, A., Mohr, C., & Herzog, M. H. (2014). Is there a common factor for vision? *Journal of Vision*, *14*(8), 1–11. <https://doi.org/10.1167/14.8.4>
- Chamberlain, R., Van der Hallen, R., Huygelier, H., Van de Cruys, S., & Wagemans, J. (2017). Local-global processing bias is not a unitary individual difference in visual processing. *Vision Research*, *141*, 247–257. <https://doi.org/10.1016/j.visres.2017.01.008>
- Chouinard, P. A., Noulty, W. A., Sperandio, I., & Landry, O. (2013). Global processing during the Müller-Lyer illusion is distinctively affected by the degree of autistic traits in the typical population. *Experimental Brain Research*, *230*(2), 219–231. <https://doi.org/10.1007/s00221-013-3646-6>

- Chouinard, P. A., Unwin, K. L., Landry, O., & Sperandio, I. (2016). Susceptibility to Optical Illusions Varies as a Function of the Autism-Spectrum Quotient but not in Ways Predicted by Local–Global Biases. *Journal of Autism and Developmental Disorders*, 46(6), 2224–2239. <https://doi.org/10.1007/s10803-016-2753-1>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.
- Coren, S., & Girgus, J. S. (1972). Illusion decrement in intersecting line figures. *Psychonomic Science*, 26(2), 108–110. <https://doi.org/10.3758/BF03335451>
- Coren, S., Girgus, J. S., Erlichman, H., & Hakstian, A. R. (1976). An empirical taxonomy of visual illusions. *Perception & Psychophysics*, 20(2), 129–137. <https://doi.org/10.3758/BF03199444>
- Cretenoud, A. F., Francis, G., & Herzog, M. H. (2020). When illusions merge. *Journal of Vision*, 20(8), 1–15. <https://doi.org/10.1167/JOV.20.8.12>
- Cretenoud, A. F., Grzeczowski, L., Bertamini, M., & Herzog, M. H. (2020). Individual differences in the Müller-Lyer and Ponzo illusions are stable across different contexts. *Journal of Vision*, 20(6), 1–14. <https://doi.org/10.1167/JOV.20.6.4>
- Cretenoud, A. F., Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, 19(14), 1–21. <https://doi.org/10.1167/19.14.12>
- Day, R. H. (1962). The effects of repeated trials and prolonged fixation on error in the Müller-Lyer figure. *Psychological Monographs: General and Applied*, 76(14), 1–19. <https://doi.org/10.1037/h0093815>
- Emery, K. J., Volbrecht, V. J., Peterzell, D. H., & Webster, M. A. (2017a). Variations in normal color vision. VI. Factors underlying individual differences in hue scaling and their implications for models of color appearance. *Vision Research*, 141, 51–65. <https://doi.org/10.1016/j.visres.2016.12.006>
- Emery, K. J., Volbrecht, V. J., Peterzell, D. H., & Webster, M. A. (2017b). Variations in normal color vision. VII. Relationships between color naming and hue scaling. *Vision Research*, 141, 66–75. <https://doi.org/10.1016/j.visres.2016.12.007>
- Festinger, L., White, C. W., & Allyn, M. R. (1968). Eye movements and decrement in the Müller-Lyer illusion. *Perception & Psychophysics*, 3(5), 376–382. <https://doi.org/10.1007/BF03396511>

- Fisher, R. A. (1992). Statistical Methods for Research Workers. In Kotz S., Johnson N.L. (eds) *Breakthroughs in Statistics*. (pp. 66–70). Springer, New York, NY. https://doi.org/10.1007/978-1-4612-4380-9_6
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78. <https://doi.org/10.1016/j.paid.2016.06.069>
- Grzeczowski, L., Clarke, A. M., Francis, G., Mast, F. W., & Herzog, M. H. (2017). About individual differences in vision. *Vision Research*, *141*, 282–292. <https://doi.org/10.1016/j.visres.2016.10.006>
- Grzeczowski, L., Roinishvili, M., Chkonia, E., Brand, A., Mast, F. W., Herzog, M. H., & Shaqiri, A. (2018). Is the perception of illusions abnormal in schizophrenia? *Psychiatry Research*, *270*, 929–939. <https://doi.org/10.1016/j.psychres.2018.10.063>
- Halpern, S. D., Andrews, T. J., & Purves, D. (1999). Interindividual variation in human visual performance. *Journal of Cognitive Neuroscience*, *11*(5), 521–534.
- Happé, F. G. E. (1996). Studying weak central coherence at low levels: Children with autism do not succumb to visual illusions. A research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *37*(7), 873–877. <https://doi.org/10.1111/j.1469-7610.1996.tb01483.x>
- Jeffreys, H. (1961). *Theory of probability*. Oxford University Press.
- Judd, C. H. (1902). Practice and its effects on the perception of illusions. *Psychological Review*, *9*(1), 27–39. <https://doi.org/10.1037/h0073071>
- Kiely, K., & Anstey, K. (2017). Common Cause Theory in Aging. *Encyclopedia of Geropsychology*, 559–569.
- King, D. J., Hodgekins, J., Chouinard, P. A., Chouinard, V. A., & Sperandio, I. (2017). A review of abnormalities in the perception of visual illusions in schizophrenia. *Psychonomic Bulletin and Review*, *24*(3), 734–751. <https://doi.org/10.3758/s13423-016-1168-5>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Leibowitz, H., Brislin, R., Perlmutter, L., & Hennessy, R. (1969). Ponzo perspective illusion as a manifestation of space perception. *Science*, *166*(3909), 1174–1176.

- Liljequist, D., Elfving, B., & Skavberg Roaldsen, K. (2019). Intraclass correlation – A discussion and demonstration of basic features. *PLOS ONE*, *14*(7), e0219854. <https://doi.org/10.1371/journal.pone.0219854>
- Lindenberger, U., & Ghisletta, P. (2009). Cognitive and Sensory Declines in Old Age: Gauging the Evidence for a Common Cause. *Psychology and Aging*, *24*(1), 1–16. <https://doi.org/10.1037/a0014986>
- Manning, C., Morgan, M. J., Allen, C. T. W., & Pellicano, E. (2017). Susceptibility to Ebbinghaus and Müller-Lyer illusions in autistic children: a comparison of three different methods. *Molecular Autism*, *8*, 16. <https://doi.org/10.1186/s13229-017-0127-y>
- Mollon, J. D., Bosten, J. M., Peterzell, D. H., & Webster, M. A. (2017). Individual differences in visual science: What can be learned and what is good experimental practice? *Vision Research*, *141*, 4–15. <https://doi.org/10.1016/j.visres.2017.11.001>
- Ninio, J. (2014). Geometrical illusions are not always where you think they are: a review of some classical and less classical illusions, and ways to describe them. *Frontiers in Human Neuroscience*, *8*:856. <https://doi.org/10.3389/fnhum.2014.00856>
- Partos, T. R., Cropper, S. J., & Rawlings, D. (2016). You Don't See What I See: Individual Differences in the Perception of Meaning from Visual Stimuli. *PLOS ONE*, *11*(3), e0150615. <https://doi.org/10.1371/journal.pone.0150615>
- Peterzell, D. H. (2016). Discovering Sensory Processes Using Individual Differences : A Review and Factor Analytic Manifesto. *Electronic Imaging*, *2016*(16), 1–11. <https://doi.org/10.2352/ISSN.2470-1173.2016.16HVEI-112>
- Peterzell, D. H. (2020). Individual Differences in Perceptual Organization: Reanalyzing Thurstone's classic (1944) data and rediscovering factors for geometrical illusions, perceptual switching, and holistic 'Gestalt' closure. *Journal of Vision*, *20*(11), 140. <https://doi.org/10.1167/jov.20.11.140>
- Peterzell, D. H., Scheffrin, B. E., Tregear, S. J., & Werner, J. S. (2000). Spatial frequency tuned covariance channels underlying scotopic contrast sensitivity. In *Vision Science and its Applications: OSA Technical Digest*. Optical Society of America.
- Peterzell, D. H., Serrano-Pedraza, I., Widdall, M., & Read, J. C. A. (2017). Thresholds for sine-wave corrugations defined by binocular disparity in random dot stereograms: Factor analysis of

- individual differences reveals two stereoscopic mechanisms tuned for spatial frequency. *Vision Research*, 141, 127–135. <https://doi.org/10.1016/j.visres.2017.11.002>
- Piaget, J. (1961). *Les Mécanismes perceptifs, Modèles probabilistes, Analyse génétique, Relations avec l'intelligence* (2nd ed.). Presses universitaires de France.
- Predebon, J. (2006). Decrement of the Müller-Lyer and Poggendorff illusions: The effects of inspection and practice. *Psychological Research*, 70(5), 384–394. <https://doi.org/10.1007/s00426-005-0229-6>
- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/>
- Robinson, J. O. (1968). Retinal inhibition in visual distortion. *British Journal of Psychology*, 59(1), 29–36.
- Roff, M. (1953). A Factorial Study of Tests in the Perceptual Area. *Psychometric Monographs*, No. 8, 1–41.
- Schwarzkopf, D. S., Song, C., & Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nature Neuroscience*, 14(1), 28–30. <https://doi.org/10.1038/nn.2706>
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428.
- Song, C., Schwarzkopf, D. S., & Rees, G. (2011). Interocular induction of illusory size perception. *BMC Neuroscience*, 12(1), 1–9. <https://doi.org/10.1186/1471-2202-12-27>
- Taylor, T. R. (1974). A factor analysis of 21 illusions: The implications for theory. *Psychologia Africana*, 15, 137–148.
- Taylor, T. R. (1976). The factor structure of geometric illusions: A second study. *Psychologia Africana*, 16, 177–200.
- Thurstone, L. L. (1944). *A factorial study of perception*. The University of Chicago Press.
- Todorović, D., & Jovanović, L. (2018). Is the Ebbinghaus illusion a size contrast illusion? *Acta Psychologica*, 185, 180–187. <https://doi.org/10.1016/j.actpsy.2018.02.011>

- Tulver, K. (2019). The factorial structure of individual differences in visual perception. *Consciousness and Cognition*, 73:102762. <https://doi.org/10.1016/j.concog.2019.102762>
- Tulver, K., Aru, J., Rutiku, R., & Bachmann, T. (2019). Individual differences in the effects of priors on perception: A multi-paradigm approach. *Cognition*, 187(3), 167–177. <https://doi.org/10.1016/j.cognition.2019.03.008>
- Wagner, D. A. (1977). Ontogeny of the Ponzo illusion: Effects of age, schooling, and environment. *International Journal of Psychology*, 12(3), 161–176. <https://doi.org/10.1080/00207597708247386>
- Webster, M. A., & MacLeod, D. I. A. (1988). Factors underlying individual differences in the color matches of normal observers. *Journal of the Optical Society of America A*, 5(10), 1722–1735. <https://doi.org/10.1364/JOSAA.5.001722>
- Wexler, M. (2018). Multidimensional internal dynamics underlying the perception of motion. *Journal of Vision*, 18(5), 1–12. <https://doi.org/10.1167/18.5.7>
- Wexler, M., Duyck, M., & Mamassian, P. (2015). Persistent states in vision break universality and time invariance. *Proceedings of the National Academy of Sciences of the United States of America*, 112(48), 14990–14995. <https://doi.org/10.1073/pnas.1508847112>
- Yildiz, G. Y., Sperandio, I., Kettle, C., & Chouinard, P. A. (2021). Interocular transfer effects of linear perspective cues and texture gradients in the perceptual rescaling of size. *Vision Research*, 179, 19–33. <https://doi.org/10.1016/j.visres.2020.11.005>

8 Supplementary File

8.1 Supplementary text

In the Ebbinghaus illusion (EB), the left and right central disk (i.e., left and right targets) were surrounded by 8 small (2° in diameter) and 6 large (5.2° in diameter) flankers, respectively. The reference target was 4° in diameter. The centers of the small flankers were 3.4° away compared to the center of the left target, while the center-to-center distance between the right target and the large flankers was 6.1° . The left and right targets were 12.4° to the left and right and 4° to the top and bottom compared to the center of the screen, respectively. At the beginning of each trial, the adjustable target was displayed with a random size in the range of 0 to 4.5° in diameter when the left target was adjustable, and in the range of 0 to 6.7° in diameter when the right target was adjustable.

In the horizontal-vertical illusion (HV), the length of the reference segment was 10.4° . The initial length of the adjustable segment was randomized between 0° and 17° at the beginning of each trial. The horizontal segment was displaced by 5.2° to the bottom compared to the horizontal screen midline. The vertical segment was always touching the center of the horizontal segment, no matter of its size.

In the Müller-Lyer illusion (ML), the left vertical segment (i.e., the left shaft) was presented with inward-pointing arrows, while the right vertical segment (i.e., the right shaft) was presented with outward-pointing arrows. The reference shaft was 8° long and the length of the adjustable shaft was randomly set at the beginning of each trial between 0 and 20.9° . The fins were made of 1.5° long lines oriented at 45° compared to the vertical. The shaft with inward-pointing arrows was centered 2.1° to the top and 5° (or 12.4° in Experiment 3) to the left compared to the middle of the screen. The shaft with outward-pointing arrows was centered 2.1° to the bottom and 5° (or 12.4° in Experiment 3) to the right compared to the middle of the screen.

In the Poggendorff illusion (PD), two vertical main streams, 16.6° long, and 4° away from each other were centrally displayed. Both parts of the interrupted diagonal were 3.8° long and tilted by 37° compared to the vertical. When the position of the left part of the interrupted diagonal was adjusted, the right part was touching the right main stream 6.3° away compared to the top of the main stream. When the position of the right part of the interrupted diagonal was adjusted, the left part was touching the left main stream 5° away compared to the bottom of the main stream. At the beginning of each trial, the adjustable element was positioned anywhere between the top and bottom of the corresponding main stream.

In the Ponzo illusion (PZ), the reference was a 4.5° long, horizontal segment. The length of the adjustable segment was randomly chosen by the computer at the beginning of each trial but it never exceeded 12° . Both reference and adjustable segments were 5.7° away from the horizontal screen midline. To induce a trapezoid-like perspective in the illusion, two converging lines were shown with 4° separating them at the apex and 18° at the base. The total height of the imaginary trapezoid was 14.4° and the illusion was overall centrally displayed.

In one variant of the contrast illusion (CS), the outside, large squares were 12° in side, while the inside, small squares were 4° in side. The couple of outside squares was centrally displayed and one inside square was displayed in the middle of each outside square. The luminance of the left and right outside squares was approximately 44 and 77 cd/m^2 , while the reference inside square was $\approx 52 \text{ cd/m}^2$ in luminance.

In another variant of the contrast illusion (CS2), the adjustable and reference disks were 3° in diameter. They were 19.9° to the left and right compared to the middle of the screen and vertically centered. The luminance of the reference disk was approximately 52 cd/m^2 . The background was made of 10 vertical stripes of equal size but with different shades of grey (from black to white).

In the White illusion (WH), 10 alternating dark (18 cd/m^2) and light (87 cd/m^2) horizontal, 2.8° wide stripes composed the background. The reference and adjustable elements were two columns of 5.2° wide and 2.8° tall rectangles, centered on the stripes at 5° to the left and right compared to the vertical midline. The left column of rectangles appeared in dark stripes, while the right column of rectangles appeared in light stripes. The luminance of the reference rectangles was $\approx 52 \text{ cd/m}^2$.

In the CS, CS2, and WH illusions, the initial luminance of the adjustable element(s) was randomly assigned at the beginning of each trial between white and black.

8.2 Supplementary figure

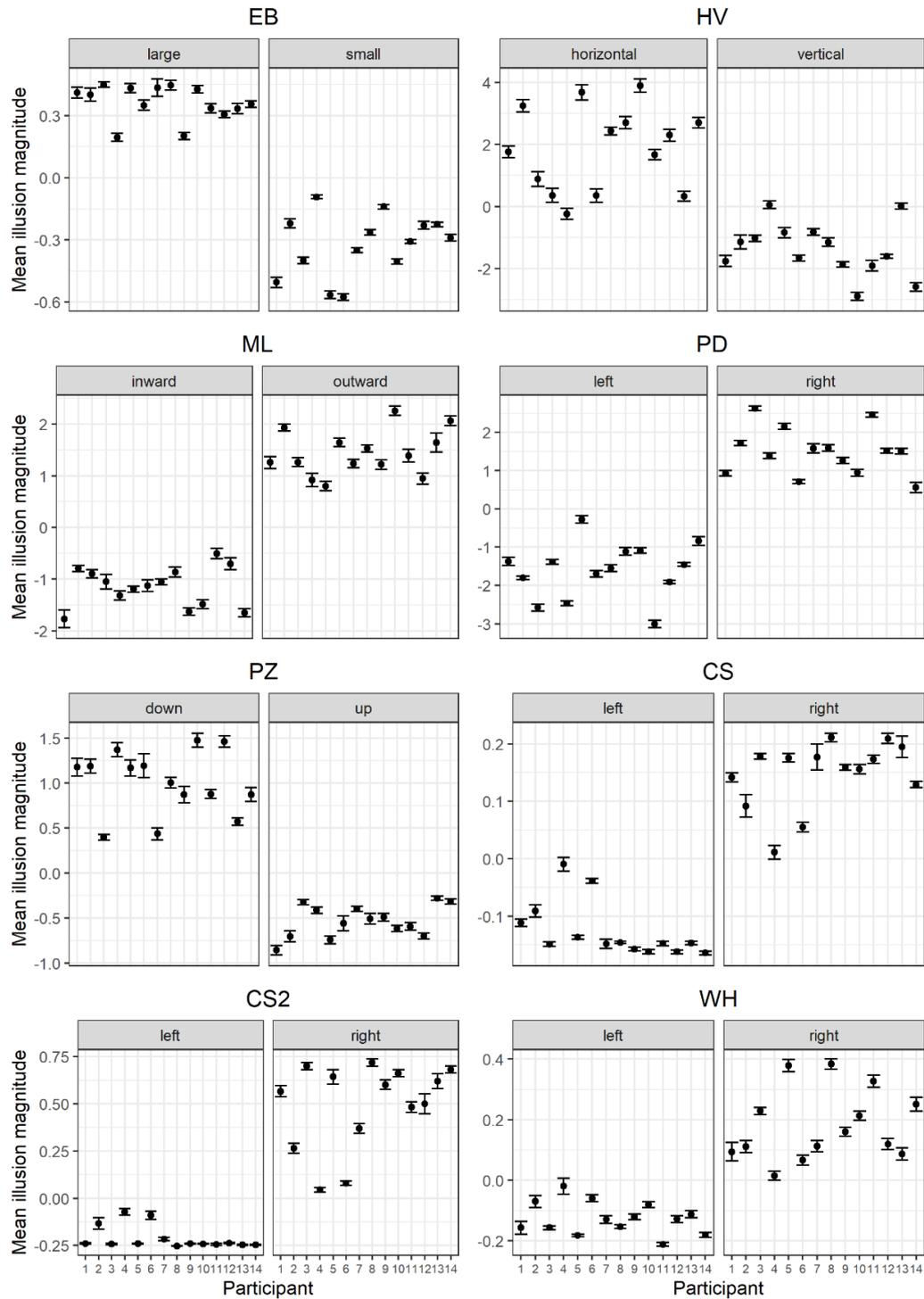


Figure S1. Individual illusion magnitudes \pm SEM (standard error of the mean) averaged across sessions for both reference-dependent conditions of each illusion (sub-panels) tested in Experiment 2. Inter-individual differences are strong, while intra-individual differences (across sessions) seem weaker. Illusion magnitudes are plotted in units of arcdeg for the EB, HV, ML, PD, and PZ illusions and in normalized luminance for contrast illusions (CS, CS2, WH).

8.3 Supplementary tables

Table S1. Illusion magnitudes (expressed as a size difference compared to the reference) summarized as mean (*M*), standard deviation (*SD*), standard error of the mean (*SEM*), and 95% confidence interval (95% CI) for each illusion and each eye condition in Experiment 1 (*n* = 15; values in units of arcdeg).

Illusion	Eye condition	<i>M</i>	<i>SD</i>	<i>SEM</i>	95% CI
EB	binocular	0.398	0.220	0.057	0.122
	monocular - right	0.428	0.269	0.070	0.149
	monocular - left	0.387	0.234	0.060	0.129
EB2	binocular	0.117	0.146	0.038	0.081
	monocular - right	0.095	0.134	0.035	0.074
	monocular - left	0.138	0.118	0.030	0.065
ML	binocular	1.930	0.812	0.210	0.449
	monocular - right	1.861	0.976	0.252	0.541
	monocular - left	1.774	0.915	0.236	0.506
PZ	binocular	-0.415	0.297	0.077	0.164
	monocular - right	-0.466	0.301	0.078	0.167
	monocular - left	-0.384	0.280	0.072	0.155
PZw	binocular	-0.335	0.199	0.051	0.110
	monocular - right	-0.337	0.211	0.054	0.117
	monocular - left	-0.285	0.176	0.045	0.097
PZg	binocular	-0.663	0.248	0.064	0.137
	monocular - right	-0.710	0.275	0.071	0.152
	monocular - left	-0.686	0.216	0.056	0.119
PZc	binocular	0.264	0.196	0.051	0.109
	monocular - right	0.354	0.217	0.056	0.120
	monocular - left	0.393	0.341	0.088	0.189

Table S2. Intrarater reliabilities expressed as intraclass coefficients (ICCs) for both reference-dependent conditions of each illusion across all sessions tested in Experiment 2 ($n = 14$; 12 sessions). All coefficients were significant after Bonferroni correction was applied.

Illusion	Condition	ICC coef.	95% CI	df	$p_{Bonferroni}$
EB	large	0.369	[0.240, 0.575]	[13, 299]	< 0.001
	small	0.815	[0.714, 0.908]	[13, 299]	< 0.001
HV	horizontal	0.713	[0.583, 0.848]	[13, 299]	< 0.001
	vertical	0.652	[0.513, 0.809]	[13, 299]	< 0.001
ML	inward	0.416	[0.280, 0.621]	[13, 299]	< 0.001
	outward	0.458	[0.317, 0.659]	[13, 299]	< 0.001
PD	left	0.796	[0.688, 0.897]	[13, 299]	< 0.001
	right	0.758	[0.639, 0.876]	[13, 299]	< 0.001
PZ	down	0.480	[0.337, 0.678]	[13, 299]	< 0.001
	up	0.406	[0.272, 0.612]	[13, 299]	< 0.001
CS	left	0.753	[0.632, 0.872]	[13, 299]	< 0.001
	right	0.531	[0.386, 0.720]	[13, 299]	< 0.001
CS2	left	0.606	[0.462, 0.776]	[13, 299]	< 0.001
	right	0.785	[0.673, 0.891]	[13, 299]	< 0.001
WH	left	0.471	[0.329, 0.670]	[13, 299]	< 0.001
	right	0.691	[0.558, 0.834]	[13, 299]	< 0.001

Table S3. Illusion magnitudes (expressed as a difference compared to the reference) summarized as mean (M), standard deviation (SD), standard error of the mean (SEM), and 95% confidence interval (95% CI) for both reference-dependent conditions of each illusion across all sessions tested in Experiment 2 ($n = 14$, 12 sessions).

Illusion	Condition	M	SD	SEM	95% CI	unit	
EB	large	0.363	0.113	0.009	0.017	arcdeg	
	small	-0.326	0.154	0.012	0.023		
HV	horizontal	1.861	1.456	0.112	0.222		
	vertical	-1.370	0.938	0.072	0.143		
ML	inward	-1.145	0.496	0.038	0.076		
	outward	1.434	0.546	0.042	0.083		
PD	left	-1.609	0.750	0.058	0.114		
	right	1.496	0.652	0.050	0.099		
PZ	down	1.005	0.428	0.033	0.065		
	up	-0.536	0.228	0.018	0.035		
CS	left	-0.126	0.050	0.004	0.008		normalized luminance
	right	0.147	0.068	0.005	0.010		
CS2	left	-0.211	0.071	0.005	0.011		
	right	0.495	0.236	0.018	0.036		
WH	left	-0.125	0.068	0.005	0.010		
	right	0.182	0.130	0.010	0.020		

Table S4. Incremental ranges tested in Experiment 3. For each illusion and each reference-dependent condition, the incremental range was centered on the mean adjustment from Experiment 2 and covered three times the absolute difference between the mean adjustment from Experiment 2 and the reference (the reference was 2° and 8° in the Ebbinghaus and Müller-Lyer illusion, respectively). The control conditions were tested with eight increments equally distributed around the reference value and covering the range between the mean adjustments for both reference-dependent conditions in Experiment 2. Correct and expected answers are shown for each increment. “L”: larger; “S”: smaller.

	Increments (arcdeg)							
EB control	1.828	1.877	1.926	1.975	2.025	2.074	2.123	2.172
Correct answer°	S	S	S	S	L	L	L	L
Expected answer*	S	S	S	S	L	L	L	L
EB large	1.819	1.974	2.130	2.285	2.441	2.596	2.752	2.908
Correct answer°	S	S	L	L	L	L	L	L
Expected answer*	S	S	S	S	L	L	L	L
EB small	1.185	1.325	1.464	1.604	1.744	1.884	2.023	2.163
Correct answer°	S	S	S	S	S	S	L	L
Expected answer*	S	S	S	S	L	L	L	L
ML control	7.355	7.539	7.724	7.908	8.092	8.276	8.461	8.645
Correct answer°	S	S	S	S	L	L	L	L
Expected answer*	S	S	S	S	L	L	L	L
ML outward	7.283	7.898	8.512	9.127	9.741	10.356	10.970	11.585
Correct answer°	S	S	L	L	L	L	L	L
Expected answer*	S	S	S	S	L	L	L	L
ML inward	5.138	5.628	6.119	6.610	7.100	7.591	8.082	8.573
Correct answer°	S	S	S	S	S	S	L	L
Expected answer*	S	S	S	S	L	L	L	L

° Answers expected from a participant, who has no illusory bias. * Answers expected from a participant, whose illusory bias is equal to the average bias observed in Experiment 2.

Table S5. Illusion magnitudes (averaged across two trials in the method of adjustment; expressed as the PSE from a psychometric function in the method of constant stimuli) summarized as mean (*M*), standard deviation (*SD*), standard error of the mean (*SEM*), and 95% confidence interval (95% CI) for the control and both reference-dependent conditions of each illusion tested in Experiment 3 (*n* = 20; values in units of arcdeg).

Illusion	Condition	Method	<i>M</i>	<i>SD</i>	<i>SEM</i>	95% CI
EB	Control	adjustment	0.010	0.060	0.014	0.028
		MCS	-0.015	0.036	0.008	0.017
	Large	adjustment	0.330	0.152	0.034	0.071
		MCS	0.338	0.166	0.037	0.078
	Small	adjustment	-0.306	0.156	0.035	0.073
		MCS	-0.347	0.121	0.027	0.057
ML	Control	adjustment	0.150	0.421	0.094	0.197
		MCS	-0.094	0.123	0.027	0.057
	Outward	adjustment	1.819	0.543	0.121	0.254
		MCS	1.359	0.270	0.060	0.127
	Inward	adjustment	-1.212	0.457	0.102	0.214
		MCS	-1.681	0.292	0.065	0.137

Appendix E

Shaqiri, A., Pilz, K. S., Cretenoud, A. F., Neumann, K., Clarke, A., Kunchulia, M., & Herzog, M. H. (2019). No evidence for a common factor underlying visual abilities in healthy older people. *Developmental Psychology*, 55(8), 1775–1787, <http://dx.doi.org/10.1037/dev0000740>.

The following article is a postprint of the published article. This article may not exactly replicate the authoritative document published in the APA journal. It is not the copy of record.

Copyright holder: © APA

No Evidence for a Common Factor Underlying Visual Abilities in Healthy Older People

Albulena Shaqiri¹, Karin S. Pilz², Aline F. Cretenoud¹, Konrad Neumann³, Aaron Clarke⁴, Marina Kunchulia^{5,6}, & Michael H. Herzog¹

¹Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

²Department of Experimental Psychology, University of Groningen, The Netherlands

³Institut für Biometrie und klinische Epidemiologie, Charité - Universitaetsmedizin Berlin, Germany

⁴Department of Psychology\Department of Neuroscience, Bilkent University, Ankara, Turkey

⁵Laboratory of Vision Physiology, Beritashvili Center of Experimental Biomedicine, Tbilisi, Georgia

⁶Institute of Cognitive Neurosciences, Agricultural University of Georgia, Tbilisi, Georgia

Abstract

The world's population is aging at an increasing rate. Even in the absence of neurodegenerative disorders, healthy aging affects perception and cognition. In the context of cognition, common factors are well established. Much less is known about common factors for vision. Here, we tested 92 healthy older and 104 healthy younger participants in 19 visual tests (including visual search and contrast sensitivity) and three cognitive tests (including verbal fluency and digit span). Unsurprisingly, younger participants performed better than older participants in almost all tests. Surprisingly, however, the performance of older participants was mostly uncorrelated between visual tests, and we found no evidence for a common factor.

Keywords: aging, perception, common factor, cognition

1. Introduction

The world's population is aging. For the first time in history, people aged 65 and over will soon outnumber children under the age of 5 (NIA, 2015). These demographic changes impact all aspects of society, including the health care system. It has long been suggested that there is a common factor underlying cognitive abilities as, for example, performance strongly correlates between memory and verbal tests in the young adult population (Conway, Cowan, Bunting, Theriault, & Minkoff, 2002; McCabe et al., 2010). Cognitive abilities significantly change with age, affecting professional skills and well-being (Li & Lindenberger, 2002; Park & Reuter-Lorenz, 2009; Salthouse, 2009). Evidence suggests a common factor underlying these cognitive changes in healthy aging (for a review, see Kiely & Anstey, 2017). For example, Baltes and Lindenberger (1997) found that age-related changes of five cognitive functions were correlated. Salthouse and Czaja (2000) found more than eleven cognitive variables that shared a common variance dependent on age.

Contrary to cognition, there seems to be no common factor for adult vision (Bargary et al., 2017; Bosten & Mollon, 2010; Cappe, Clarke, Mohr, & Herzog, 2014; Chamberlain, Van der Hallen, Huygelier, Van de Cruys, & Wagemans, 2017; Emery, Volbrecht, Peterzell, & Webster, 2017; Peterzell, 2016). As in cognition, vision is affected by age-related changes. Decline has been found in contour integration (Roudaia, Bennett, & Sekuler, 2013), motion perception (Arena, Hutchinson, & Shimozaki, 2012; Bennett, Sekuler, & Sekuler, 2007; Billino, Bremmer, & Gegenfurtner, 2008; Pilz, Miller, & Agnew, 2017), biological motion (Agnew, Phillips, & Pilz, 2016; Billino et al., 2008; Norman, Payton, Long, & Hawkes, 2004; Pilz, Bennett, & Sekuler, 2010; Spencer, Sekuler, Bennett, Giese, & Pilz, 2016), vernier acuity (Garcia-Suarez, Barrett, & Pacey, 2004; Li, Edwards, & Brown, 2000) and spatial and temporal processing (Pilz, Kunchulia, Parkosadze, & Herzog, 2015; Roudaia et al., 2013).

Aging may affect certain people more than others, which may, for instance, be due to differences in optical decline and life style. For this reason, even though there is no unique common factor for vision in young people, it is possible that such a factor emerges during aging. In most studies investigating visual perception in healthy aging, performance of younger and older people is compared using only a single or very few perceptual tests, which prevents drawing specific conclusions on common factors (e.g., Andersen, 2012; Bennett et al., 2007; Billino et al., 2008; McBain, Norton, & Chen, 2010; Owsley, McGwin, & Seder, 2011; Pilz et al., 2010, 2015, 2017; Spear, 1993). Here, we employed a much larger battery of mainly perceptual tasks compared with previous studies and investigated performance in younger and

older participants. The tests are not all purely perceptual but are all somehow representative of various visual functions.

We also included three cognitive tests in our battery to assess a potential relationship between cognitive and perceptual functions. Visual sensation and perception precede cognition and, hence, cognitive deficits may at least partially be related to sensory and perceptual deficits, e.g., attentional changes might in fact be changes in visual filtering (Lee, Itti, Koch, & Braun, 1999). Indeed, moderate to strong correlations were found between sensory and cognitive abilities (Anstey, Hofer, & Luszcz, 2003; Ghisletta & Lindenberger, 2005; Lindenberger & Baltes, 1994; Lindenberger & Ghisletta, 2009; for a review, see Kiely & Anstey, 2017).

As in previous studies, we found strong performance differences between younger and older adults and only weak correlations between tests in the younger population. Surprisingly, we found only weak correlations between tests in the older population, i.e., performance in one test did not predict performance in other tests.

2. Methods

2.1. Participants

We tested 105 younger and 131 older participants. Older participants were recruited from the general population in Tbilisi, Georgia, and from a course about refreshing driving skills in Lausanne, Switzerland. For the young group, we recruited university students from both Lausanne and Tbilisi Universities. All participants had normal or corrected-to-normal visual acuity, as assessed via the Freiburg visual acuity test (FrACT; Bach, 1996) and had no known history of eye diseases, such as macular degeneration or cataract. All older adults were screened for mild cognitive impairments with the Montreal Cognitive Assessment (MoCa; Nasreddine et al., 2005). Since we wanted to test only healthy participants, we excluded 36 older participants because their scores were lower than 26, which is the threshold for a potential mild cognitive impairment. In addition, data of one young and three older participants were excluded because of excessive data loss. Therefore, we report the results for the remaining subset of 92 older participants (60-90 years, $M = 70$ years, $SD = 6.89$) and 104 younger adults (18-31 years, $M = 21.8$, $SD = 2.66$). Participants were tested either at the École Polytechnique Fédérale de Lausanne, Switzerland ($N = 79$), or at the Beritashvili Center of Experimental Biomedicine in Tbilisi, Georgia ($N = 117$).

This study was approved by the ethics committee of the Canton de Vaud in Lausanne, Switzerland (study title: “Aspects fondamentaux de la reconnaissance des objets: approche psychophysique et électrophysiologique”, approval number: 384/2011), and the ethics committee of the Beritashvili Center of Experimental Biomedicine in Tbilisi, Georgia. All participants gave written informed consent, were reimbursed for their participation, and were informed that they could quit the experiment at any time.

2.2 Procedure and Analysis

All participants performed 19 visual tests, of which some tests were versions of the same paradigm. For example, we tested biological motion perception with four conditions. In addition, participants performed three cognitive tests (verbal fluency, digit span forward and backward and the Wisconsin Card Sorting Test) and answered three questionnaires: the Montreal Cognitive Assessment, the Geriatric Depression Scale and the Autism-spectrum Quotient questionnaire. The 19 visual tests tested basic to complex visual processing: vernier discrimination (duration and offset), visual backward masking (with a 5 and a 25 elements grating), visual acuity, orientation discrimination, contrast sensitivity, motion direction discrimination, biological motion perception (for 200 ms and 800 ms, upward and inverted), simple reaction time, visual search (for three numbers of distractors; we determined two measures: slope and reaction time) and a Simon test.

Data transformation. We first rescaled the data to approximate normal distributions so that a 3-sigma outlier criterion and Pearson correlations were less affected by skewness. For rescaling, we used a power transformation whose exponent λ was optimized for each test in order to maximize normality according to the Shapiro-Wilk test (the values of λ for each test are given in the Supplemental Material, Table S1). For each test, the optimization procedure minimized the average of the Shapiro-Wilk test statistic values for the two age groups (younger and older). That is, the optimized exponent λ was test-specific but not group-specific. For tests that used the same measure, the same exponent was applied (i.e., $\lambda_{\text{Bio200}} = \lambda_{\text{Bio800}} = \lambda_{\text{Bio200i}} = \lambda_{\text{Bio800i}}$, $\lambda_{\text{BMask5}} = \lambda_{\text{BMask25}}$, and $\lambda_{\text{DigitF}} = \lambda_{\text{DigitB}}$). Finding the optimal exponent for rescaling was interleaved with outlier identification in order to alleviate the problem of both steps, rescaling and outlier identification, spoiling each other.

This procedure was modified for the vernier duration and the biological motion tests, which showed strong floor or ceiling effects at least for some group or conditions. For vernier duration, thresholds above 500 ms were classified as outliers and removed, and the scores were log-transformed. For the biological motion tests, the few outliers were identified and removed just after rescaling.

Finally, we z-transformed the data for each test and flipped the sign when better performance was indicated by lower values. Hence, higher scores indicate better performance for all tests.

Merging task repetitions. For the visual backward masking tasks, performance was determined twice for both the 5 and the 25 elements grating. In the factor analysis, we used the mean values of the two measurements.

Data imputation. We removed data from participants who had more than 25% of the scores missing, e.g., due to outlier rejection. This was the case for one younger and three older participants. For the factor analysis, we imputed missing scores using the function 'mice' from the R package 'mice' with method 'norm' (Bayesian linear regression). Since in total only 2.9% of the scores were missing, we used simple imputation with only one imputation sample.

Factor analysis. To determine the optimal number of factors for each group, we ran exploratory and confirmatory factor analyses for F factors and chose the model with the minimal Akaike criterion. For the exploratory factor analysis (EFA), we extracted F factors (F ranging from 1 to 6) with 'varimax' rotation using the R function 'principal' from the R package 'psych'. In order to simplify the factor structure and to arrive at a suitable hypothesis for the subsequent confirmatory factor analysis, we followed the approach by Jöreskog (1978). First, we identified for each factor F_i the variable V_i which loaded maximally on the factor F_i . Then we performed a suitable linear transformation (oblique rotation) such that the V_i has an even larger loading on the rotated factor F_i and zero or very small loadings on all other rotated factors. Hence after the rotation, these variables mainly loaded on one factor only. For the confirmatory analysis (CFA), we constructed a path model from the simplified model by including only connections whose loadings were > 0.3 (or < -0.3 , respectively). So, the null hypothesis for the CFA is that all other loadings are exactly zero. The coefficients of this path model were computed with the 'sem' function of the R package 'psych'. The correlation matrices that were used to compute the factor analysis for the young and old population are reported in the Supplemental Material (Table S2).

For the young population, the 3-factor model had the smallest Akaike criterion, for the older population it was the 4-factor model, which we report in detail.

2.3. Tests

2.3.1. Perceptual tests

2.3.1.1. *Vernier discrimination and backward masking*

We adapted the shine-through paradigm (Herzog, Kopmann, & Brand, 2004; Roinishvili, Chkonia, Stroux, Brand, & Herzog, 2011). Vernier stimuli were presented from a distance of 500 cm in a dimly illuminated room. The stimuli were white on a black background. In a first step, we presented vernier stimuli consisting of two vertical bars which were offset in the horizontal direction (Figure 1a). Participants indicated via button press the offset direction of the lower bar compared to the upper bar (left or right). Errors were indicated by an auditory signal. For each observer, we determined the individual vernier duration (VernDur) required to reach 75% correct responses using a staircase procedure (for details, see Chkonia et al., 2010). The staircase procedure started with a duration of 150 ms. In addition to the individual vernier duration, we also determined the vernier offset thresholds (VernOfs) for the 150 ms condition.

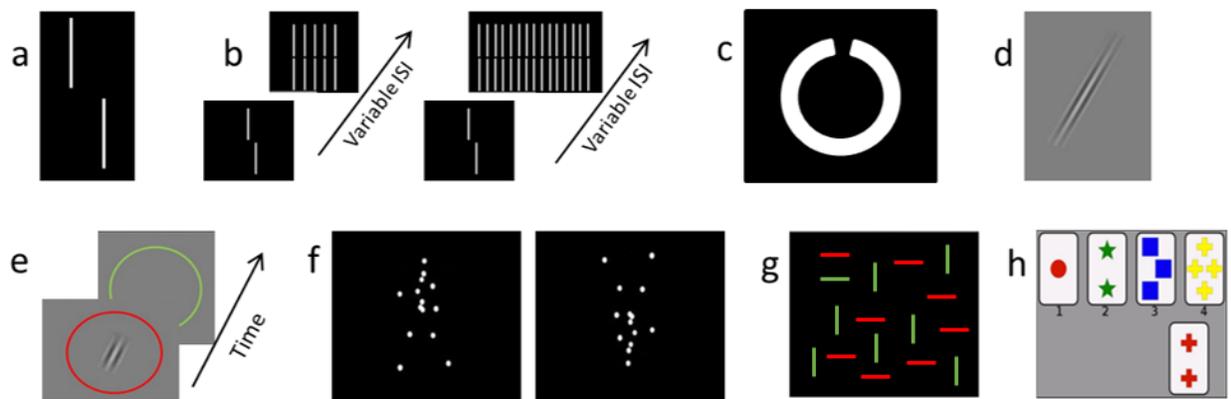


Figure 1. A few examples of the perceptual and cognitive tasks used. (a) vernier discrimination; (b) visual backward masking with the 5 and the 25 elements grating; (c) Freiburg Visual Acuity Test; (d) orientation discrimination; (e) contrast sensitivity; (f) upward and inverted biological motion perception; (g) visual search; (h) Wisconsin Card Sorting Test.

In the backward masking conditions, the vernier was followed by a variable inter-stimulus interval (ISI), i.e., a blank screen, and then a grating for 300 ms. We used the individual vernier duration for each observer as determined in the vernier discrimination task and we varied the ISI adaptively using a staircase procedure (Taylor & Creelman, 1967). The grating consisted of either 5 (BMask5) or 25 (BMask25) aligned elements of the same length as the target vernier (Figure 1b). We reported the stimulus onset asynchrony

(SOA = individual VernDur + ISI, in ms). The five elements grating leads to stronger masking than the 25 elements grating even though the five elements grating is contained within the 25 elements grating (Herzog & Koch, 2001). This difference in masking strength indicates that a substantial part of the masking power is not of retinal origin because retinal processing is mainly determined by the sheer amount of light, e.g., the number of grating elements presented. No feedback was given.

2.3.1.2. Freiburg Visual Acuity Test (FrACT)

This test has been developed and described in detail by Bach (1996) and has been validated in various studies (Schulze-Bonsel, Feltgen, Burau, Hansen, & Bach, 2006; Wesemann, 2002). Landolt-C optotypes with randomized gap orientations were presented on a computer monitor (Figure 1c). Participants indicated the direction of the gap ('up', 'down', 'left', 'right') by verbalizing their responses and the experimenter operated the input device. The size of each optotype changed adaptively following the best-PEST algorithm (Treutwein, 1995).

2.3.1.3. Orientation discrimination (Orient)

This is an adapted version of the orientation discrimination test as used previously (Tibber, Guedes, & Shepherd, 2006). Participants were seated 200 cm from the computer screen and were asked whether a Gabor patch was oriented clockwise or anticlockwise (Figure 1d). Auditory feedback was given for incorrect responses. The outcome measure was the perceptual threshold in degrees from vertical.

2.3.1.4. Contrast sensitivity (Contrast)

We measured contrast sensitivity similarly to Lahav et al. (2011). Stimuli were presented from a distance of 200 cm in a dimly illuminated room. Participants indicated in which of two subsequently presented circles (first red, second green, Figure 1e) a Gabor patch was presented. Participants had to press the button corresponding to the color of the circle in which the Gabor was presented (i.e., a red or green button). A staircase method was used to determine the contrast threshold level at which participants reached 75% correct responses. Auditory feedback was given for incorrect responses.

2.3.1.5. Motion direction discrimination (Motion)

This test measures global motion perception. Stimuli and procedures were based on work by Banton, Dobkins, and Bertenthal (2001). Participants were seated 200 cm in front of the screen, and had to judge the motion of dot patterns. On each trial, a certain number of dots moved either to the right or to the left

(target), and a certain number of dots moved independently from each other (distractors). The percentage of target dots varied according to a staircase procedure (PEST). The starting value was 20%.

2.3.1.6. Biological motion perception

Participants were seated 60 cm from the screen and indicated the walking direction of a point-light walker via a button press (left or right, Figure 1f). The walker did not move across the screen, but appeared to walk in place as if on a treadmill. The walker's direction of motion was either rightward or leftward, and the walker was either presented upright or inverted (conditions were not intermixed) for either 200 or 800 ms (upright: Bio200, Bio800; inverted: Bio200i, Bio800i). We determined accuracy (in %). Auditory feedback was given. For details, see (Pilz et al., 2010).

2.3.1.7. Simple reaction time (SimpRT)

Participants were seated at a distance of 200 cm from the computer screen and were instructed to press a button immediately after a white square (3 arc degrees size) appeared on the black background. The inter-trial interval (ITI) was varied randomly to prevent participants from predicting when the square appears. There was a minimal ITI of 1500 ms. Reaction time (in ms) was measured. This test was a modified version of the classic Hick-paradigm (Hick, 1952).

2.3.1.8. Visual search test

Participants searched for a green horizontal line within an array of distractors (green vertical and red horizontal lines, Figure 1g). They were seated at a distance of 200 cm from the computer screen. The displays consisted of 4, 9 or 16 lines and were presented in random order. Participants had to indicate whether the displays contained the green horizontal line, which was the case in 50% of the trials. Both average reaction time for the correctly answered trials (VSrchRT) and slope of the regression line for reaction time over number of elements (VSrchSL) were assessed.

2.3.1.9. Simon test (Simon)

We used a modified version of the visual Simon test (Castel, Balota, Hutchison, Logan, & Yap, 2007). Participants were seated at a distance of 200 cm from the computer screen and were presented with arrows. Participants were instructed to respond with the right hand to a right-pointing arrow and respond with the left hand to a left-pointing arrow. Arrows were presented at three locations on the screen (left, right or center). In congruent trials, the direction of the arrow matched its location (e.g., left-pointing

arrow on the left side of the screen), whereas this was not the case for incongruent trials (e.g., left-pointing arrow on the right side of the screen). The magnitude of response conflict (Simon effect) was defined by the difference of the response times for the congruent and the incongruent conditions, divided by the average response time. Only valid responses were taken into account.

2.3.2. Cognitive tests

2.3.2.1. *Verbal fluency (VbFlu)*

The Controlled Oral Word Association Test (COWAT) was originally developed by Bechtold and colleagues (Bechtoldt, Benton, & Fogel, 1962) and is a measure of left frontal lobe functioning (Newman, Trivedi, Bendlin, Ries, & Johnson, 2007; Wood, Saling, Abbott, & Jackson, 2001). Participants produced as many words as they could that belonged to the categories of animals and fruit/vegetables. The total number of words from each category was recorded.

2.3.2.2. *Digit span test*

Participants were asked to repeat a series of random numbers, which were presented at a rate of one per second. In the digit span forward test (DigitF), participants were asked to repeat the numbers in a forward order. In the digit span backward test (DigitB), participants were asked to report the numbers in reverse order. In DigitF two sets of three digits were presented first, and an additional digit was added if at least one out of two trials had been correctly recalled. For DigitB two sets of two digits were given first. The longest correctly repeated digit length (based on at least one correctly recalled set; maximum of 9 digits for DigitF and DigitB) was reported. The test was adapted from Elliott and colleagues (2011).

2.3.2.3. *Wisconsin Card Sorting Test (WCST)*

We administered a computerized version of the WCST (Berg, 1948; part of the the PEBL Psychological Test Battery) which contains 64 cards of which four were displayed at the top of the computer screen at a time (Figure 1h). Each card contained items that differed from the other three cards on three dimensions, namely in shape (crosses, triangles, circles or stars), number (1-4 items) and color (red, blue, yellow, green). A fifth card was then presented on the bottom of the screen, and participants matched the fifth card according to the shape, number of items or color by clicking on one of the four cards displayed at the top of the screen. Participants were not told according to which strategy they were supposed to match the cards. Feedback about whether or not cards were successfully matched was given after each response.

Once the participant had learned the correct strategy, the dimension according to which cards had to be matched changed. Hence, participants were required to constantly adjust their strategies. The total number of perseverative errors was assessed.

2.3.3. Questionnaires

2.3.3.1. *Montreal Cognitive Assessment*

Participants performed the Montreal Cognitive Assessment (MoCA), which tests for mild cognitive impairments (Nasreddine et al., 2005). A score between 30 and 26 corresponds to healthy participants, a score between 26 and 16 corresponds to a mild cognitive impairment and a score of below 16 corresponds to dementia.

2.3.3.2. *Geriatric Depression Scale*

The Geriatric Depression Scale (GDS; Yesavage et al., 1982) is a 30-item questionnaire that tests for depressive symptoms in older subjects. This questionnaire is generally used in geriatric standardized evaluations or for research in depression or cognitive decline.

2.3.3.3. *Autism-Spectrum Quotient questionnaire*

The Autism-Spectrum Quotient questionnaire (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) is a 50-item questionnaire designed to detect symptoms of autism or other symptoms from the autism spectrum.

3. Results

3.1. Performances

Mean values for each test and both groups are reported in Table 1. Older participants performed significantly worse in all tests except for the verbal fluency test (Table 2). Scores on the Autism-Spectrum Quotient (AQ) and Geriatric Depression Scale (GDS) did not differ between older and younger participants, indicating the good mental state of the older participants. These results are in line with previous studies (Strauss, Sherman, Spreen, & Spreen, 2006).

Table 1. Means of untransformed data after removal of outliers and of participants with more than 25% of missing data.

	Measure	Young			Old		
		<i>M</i>	<i>SE</i>	<i>N</i>	<i>M</i>	<i>SE</i>	<i>N</i>
FrACT	Visual acuity (decimal)	1.73	0.048	102	1.16	0.041	95
VbFlu	Number of words	19.8	0.48	105	20.4	0.41	94
DigitF	Number of digits	6.31	0.117	104	5.98	0.105	93
DigitB	Number of digits	4.83	0.131	103	4.36	0.114	94
WCST	Number of perseverative errors	17.6	0.90	104	27.4	1.59	93
SimpRT	Response time [ms]	260	4.0	103	280	5.2	94
Simon	Relative percent correct delta [%]	-3.94	0.356	104	-8.55	0.529	94
VernDur	Vernier duration [ms]	36.4	6.06	104	99.5	12.70	87
VernOfs	Vernier offset [arcsec]	23.9	3.21	90	40.3	4.52	88
BMask25	Mask onset asynchrony [ms]	31.7	3.64	82	91.1	10.54	84
BMask5	Mask onset asynchrony [ms]	88.0	4.82	91	183	8.7	84
Contr	Michelson contrast [%]	1.76	0.066	102	6.22	0.613	94
Motion	Motion direction coherence [%]	20.0	0.83	105	25.8	0.98	92
Bio200i	Percent correct [%]	80.8	1.75	104	57.9	1.66	92
Bio800i	Percent correct [%]	86.3	1.77	104	60.0	2.08	92
Bio200	Percent correct [%]	98.5	0.25	103	82.8	1.80	92
Bio800	Percent correct [%]	99.2	0.18	102	88.4	1.84	92
Orient	Angular orientation delta [deg]	0.657	0.0428	105	1.08	0.075	93
VSrchSL	Response time slope [ms/item]	21.5	1.24	104	36.0	1.67	94
VSrchRT	Average response time [ms]	864	18.9	104	1080	20	93

Table 2. *T*-tests comparing older and younger participants.

	<i>t</i>	df	<i>p</i>	Effect size (Cohen's <i>d</i>)	95% IC (Cohen's <i>d</i>)	
					lower bound	upper bound
FrACT	8.95	188.49	< 0.001	1.29	0.97	1.59
VbFlu	-0.90	192.13	0.368	0.13	-0.41	0.15
DigitF	2.10	191.98	0.037	0.30	0.02	0.58
DigitB	2.75	192.00	0.007	0.39	0.11	0.68
WCST	5.63	180.15	< 0.001	0.82	0.51	1.10
SimpRT	3.00	185.28	0.003	0.43	0.15	0.71
Simon	7.25	182.15	< 0.001	1.05	0.74	1.34
VernDur	5.60	129.90	< 0.001	0.84	0.51	1.09
VernOfs	6.17	169.91	< 0.001	0.93	0.59	1.18
BMask25	7.31	163.00	< 0.001	1.14	0.75	1.34
BMask5	11.56	170.04	< 0.001	1.75	1.33	1.98
Contr	15.77	187.52	< 0.001	2.28	1.90	2.61
Motion	4.83	190.40	< 0.001	0.70	0.40	0.98
Bio200i	9.92	193.50	< 0.001	1.41	1.10	1.73
Bio800i	10.11	189.67	< 0.001	1.45	1.13	1.76
Bio200	9.57	97.22	< 0.001	1.41	1.06	1.68
Bio800	6.31	94.06	< 0.001	0.93	0.61	1.20
Orient	5.93	192.81	< 0.001	0.85	0.55	1.14
VSrchSL	8.21	189.86	< 0.001	1.17	0.87	1.48
VSrchRT	9.35	175.00	< 0.001	1.33	1.03	1.65

Positive *t*-values indicate that older participants were worse than younger participants. Effect size is expressed as Cohen's *d*. Freiburg Visual Acuity Test (FrACT); COWAT verbal fluency (VbFlu); digit span forward (DigitF); digit span backward (DigitB), Wisconsin Card Sorting Test (WCST), simple reaction time (SimpRT), Simon test (Simon), vernier duration (VernDur), vernier offset (VernOfs), visual backward masking with 25 and 5 elements grating (BMask25 and BMask5); contrast sensitivity (Contr), motion direction discrimination (Motion), inverted biological motion perception for 200 ms and 800 ms (Bio200i and Bio800i), upward biological motion perception for 200 ms and 800 ms (Bio200 and Bio800), orientation discrimination (Orient), visual search slope and intercept (VSrchSL and VSrchRT).

3.2. Correlations

We computed Pearson correlations for the young and old populations (Figures 2 and 3, respectively). Correlations between most tests are rather low, except for correlations between versions of one paradigm (e.g., between the digit span forward and backward tasks).

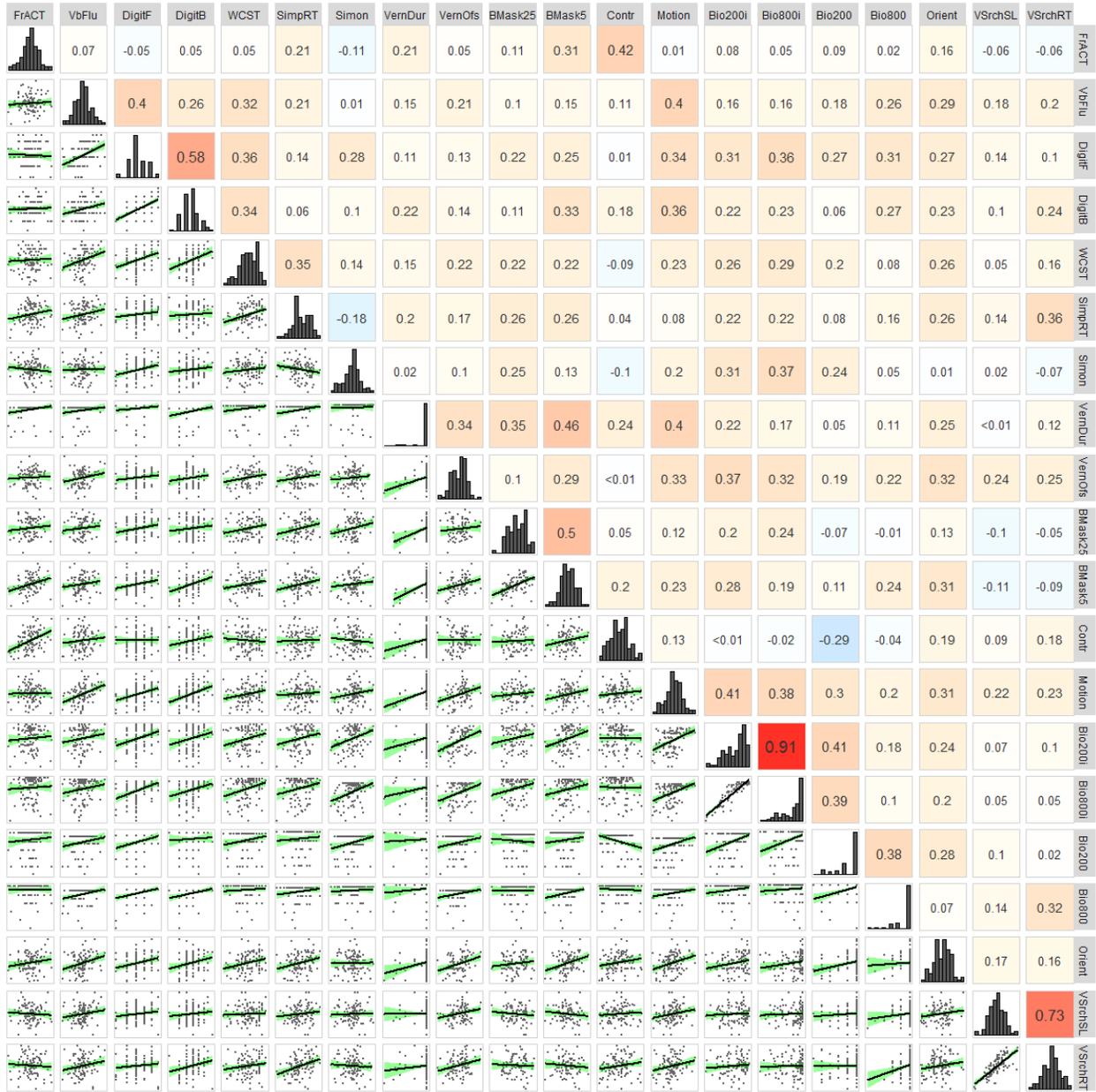


Figure 2. Correlations for the young population. Values in the upper triangular matrix show Pearson correlation coefficients (r) with a color code (blue: negative values; red: positive values). In the lower triangular matrix, we show the scatter plots of the individual performances together with regression lines plus 95% confidence intervals. The diagonal elements show histograms of the scores. Data were power-transformed to approximate normality, after removing outliers and participants with too many missing scores, but before data imputation. Vernier duration (VernDur) and upward biological motion perception (Bio200 and Bio800) show ceiling effects.

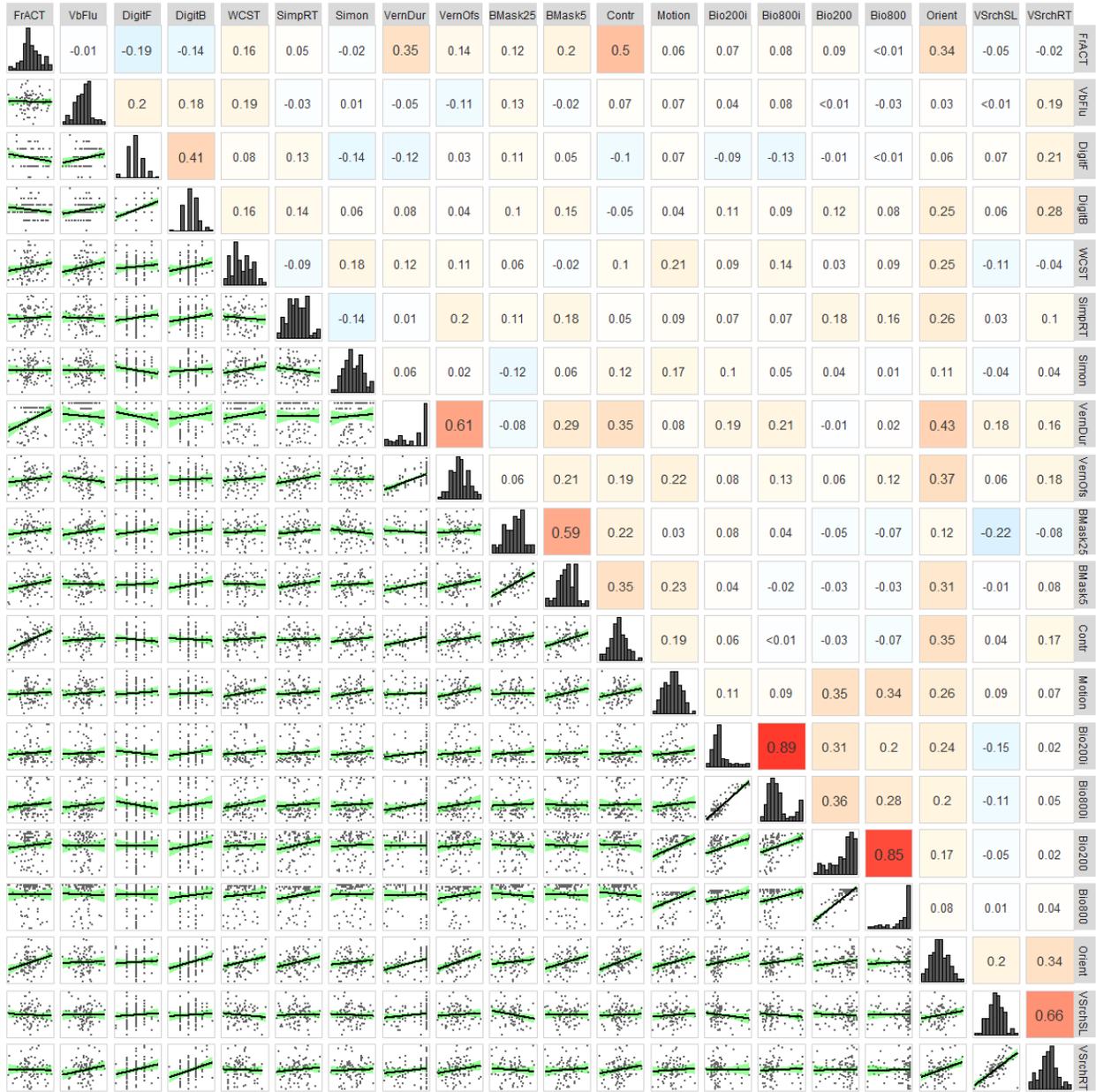


Figure 3. Correlations for the older population. Same format as in Figure 2.

3.3. Test-retest reliability, Power & Factor analysis

The null results cannot be explained by poor test-retest reliability or low statistical power. First, our test-retest correlations were high. For example, we tested performance in the visual backward masking conditions with the 5 and 25 elements grating twice and found test-retest correlations of $r_{\text{BMask5,young}} = 0.76$ [N = 85], $r_{\text{BMask5,old}} = 0.65$ [N = 73], $r_{\text{BMask25,young}} = 0.67$ [N = 72], $r_{\text{BMask25,old}} = 0.65$ [N = 80], respectively (all $p < 0.001$), which are large according to Cohen (1988) and Gignac and Szodorai (2016). In the visual search test, participants searched with 4, 9 and 16 distractors. The corresponding correlation coefficients were 0.922, 0.865 and 0.957 (young) and 0.840, 0.721 and 0.922 (older) with p values < 0.001 . In addition, versions of the same paradigm, such as Bio200 and Bio800, also showed high correlations.

Second, with about 100 participants in each group, we had 80% power to detect effect sizes with r of 0.27, which is close to a medium effect size according to Cohen (1988) and a large effect size according to Gignac and Szodorai (2016).

3.4. Factor analysis

First, the values of the Kaiser-Meyer-Olkin's (KMO) Measure of Sampling Adequacy (MSA) were 0.69 for the young and 0.62 for old participants, indicating a rather weak factor structure (according to the KMO's classification of MSA values, the suitability of the data for factor analysis is "mediocre"). Second, we computed a factor analysis for 1 to 6 factors. The best fitting model according to the Akaike criterion was obtained with 3 factors for the young population and 4 factors for the old population (Figures 4 and 5, respectively). The best fitting models do not reveal a clear factor structure, where single factors load on specific variables. Variables rather load on many factors (for the detailed model parameter statistics, see Tables S3 and S4 in the Supplemental Material). In addition, loadings are often not high. Third, for most tasks, the uniqueness u^2 is larger than the communalities h^2 , which means that the explained variance by the factors is smaller than the unexplained one (Figures 4 and 5).

factors	AIC	χ^2	df	p		F1	F2	F3	h^2	u^2	e^2
1	228.5	178.5	80	$1.7 \cdot 10^{-9}$			0.41		0.17	0.83	0.83
2	189.6	123.6	72	$1.5 \cdot 10^{-4}$		0.55		0.16	0.32	0.68	0.68
3	181.2	97.2	63	$3.7 \cdot 10^{-3}$		0.45		0.46	0.41	0.59	0.60
4	196.9	110.9	62	$1.7 \cdot 10^{-4}$		0.40		0.30	0.25	0.75	0.76
5		sem() failed				0.32	0.37		0.23	0.77	0.75
6		sem() failed						0.63	0.40	0.60	0.60
						0.38	0.58	0.16	0.50	0.50	0.48
							0.80	0.41	0.81	0.19	0.22
							0.34		0.11	0.89	0.89
						0.53	0.11	0.37	0.43	0.57	0.58
						0.28	0.17	0.56	0.42	0.58	0.59
						0.42		0.27	0.25	0.75	0.76
						0.43	0.31	0.14	0.31	0.69	0.68
						0.50			0.25	0.75	0.75

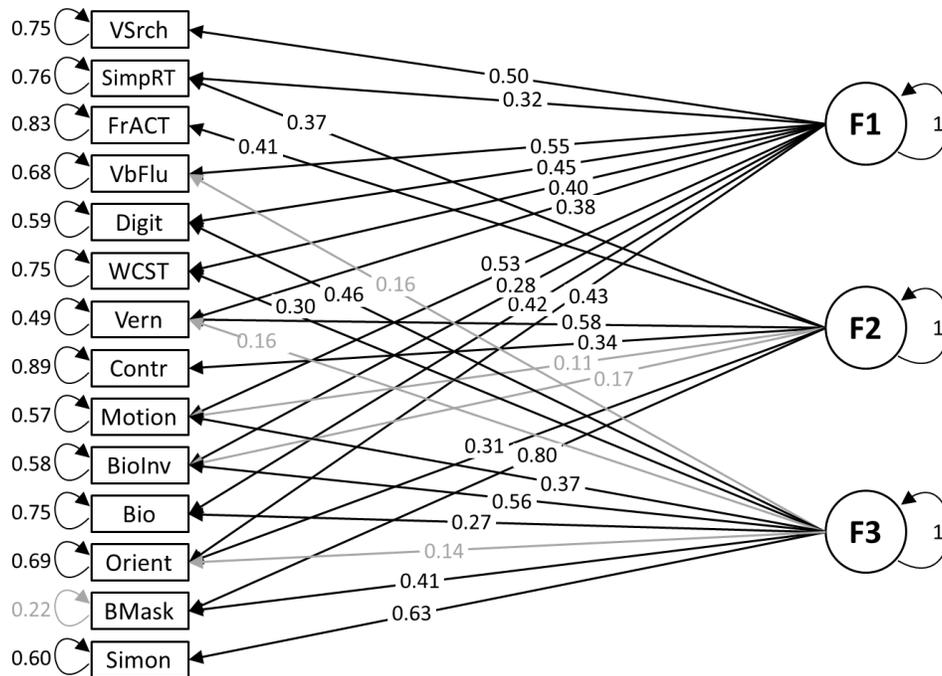


Figure 4. Confirmatory factor analysis for the young population. *Left table.* Number of factors in the model and the corresponding fit statistics including the Akaike (AIC) and χ^2 values, degrees of freedom, and p values for the one-tailed χ^2 test. The Akaike criterion is lowest for 3 factors. “Sem() failed” indicates that the computation of the CFA did not converge properly. *Right table.* For this 3-factor model, the table shows the estimated factor loadings, communalities (h^2) and uniqueness values ($u^2 = 1 - \text{communalities}$) as calculated from the loadings, and the estimated error variances (e^2), which should ideally be identical with the uniqueness values. *Bottom.* Factor structure from the 3-factor model. Grayed-out values are not significantly different from zero ($p > 0.05$).

factors	AIC	χ^2	df	p		F1	F2	F3	F4	h^2	u^2	e^2
1	160.26	112.26	81	0.012	FrACT	0.67				0.45	0.55	0.55
2	145.79	83.79	74	0.204	VbFlu			0.24	0.22	0.11	0.89	0.89
3	138.21	74.21	73	0.439	Digit			0.56		0.31	0.69	0.69
4	134.70	62.70	69	0.690	WCST				0.47	0.22	0.78	0.78
5		sem() failed			SimpRT	0.15	0.28	0.31	-0.41	0.37	0.63	0.67
6					Simon				0.36	0.13	0.87	0.87
					Vern	0.47	0.24			0.28	0.72	0.71
					BMask	0.28		0.14		0.10	0.90	0.91
					Contr	0.69				0.47	0.53	0.53
					Motion		0.48			0.23	0.77	0.77
					Biolnv		0.42			0.18	0.82	0.82
					Bio		0.66			0.43	0.57	0.57
					Orient	0.55	0.32	0.45		0.61	0.39	0.40
					VSrch			0.44		0.20	0.80	0.80

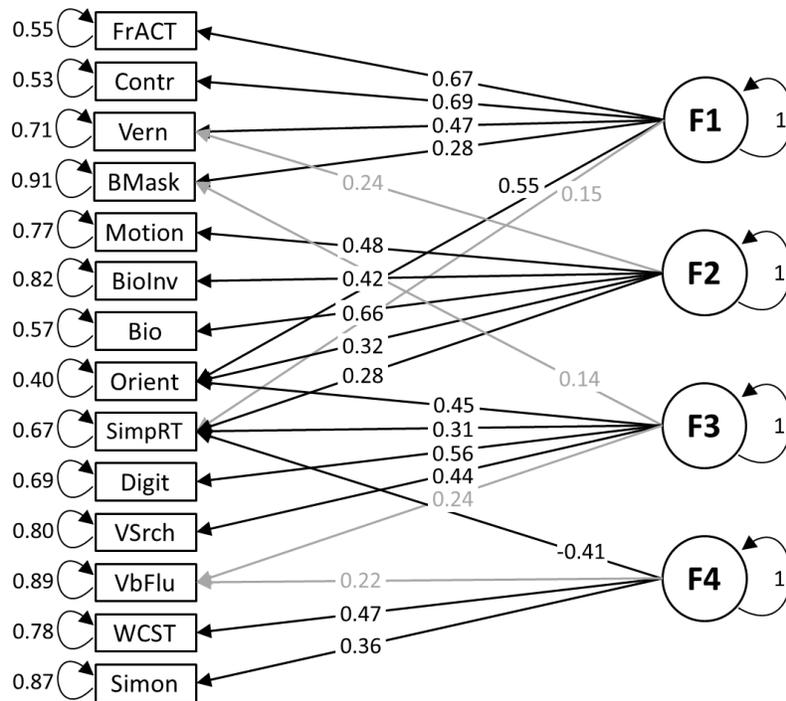


Figure 5. Confirmatory factor analysis for the old population. Same denotations as for young population (Figure 4). The Akaike criterion is lowest for 4 factors.

4. Discussion

Previous studies have shown strong correlations between cognitive tests in younger adults and have provided evidence for common factors underlying age-related cognitive deterioration (Christensen, Mackinnon, Korten, & Jorm, 2001; Salthouse, Hancock, Meinz, & Hambrick, 1996; for a review, see Kiely & Anstey, 2017). Here, we investigated to what extent there are common factors for visual perception.

First, apart from test-retest comparisons or comparisons between similar tasks, we found mainly weak correlations between visual tests in younger adults. A 3-factor model explained the data best without over-fitting. However, the factor structure did not reveal a clear pattern and loadings were mainly low. The 3 factors explained 34.8% of the total variance. As a comparison, the 1-factor model explained 20.4% of the total variance. Our results are in accordance with previous studies (Bosten & Mollon, 2010; Cappe et al., 2014; Grzeczowski, Clarke, Francis, Mast, & Herzog, 2017; Ward, Rothen, Chang, & Kanai, 2017). For example, Halpern, Andrews, & Purves, (1999) tested orientation discrimination, motion direction discrimination, vernier discrimination and contrast sensitivity and found a factor explaining 30% of the variance. Bosten et al. (2017) used 25 visual and auditory tests and found a factor that explained 19.9% of the total variance. Ward and colleagues (2017) tested performance for glass pattern, visual search and motion perception, amongst others. They found two factors explaining 18% and 19% of the variance, respectively. Across these studies, the explained variance is roughly in the same range. Interpretations differ between studies because different criteria are used to determine how much variance a common factor needs to explain.

Secondly, we found that young people perform much better at all tasks except for the verbal fluency test, which is very much in line with previous findings (Agnew et al., 2016; Bennett et al., 2007; Billino et al., 2008; Garcia-Suarez et al., 2004; Norman et al., 2004; Pilz et al., 2017).

Thirdly, most visual functions share common optical, retinal, and neural mechanisms (Thorpe & Fabre-Thorpe, 2001), which strongly change with age, but to a different extent for each individual (Raz, Perlman, Percicot, Lambrou, & Ofri, 2003; Spear, 1993). For example, lens clouding may affect most visual functions, and stronger lens clouding may lead to stronger deficits. For this reason, we expected that some older participants would perform generally better than others. However, this is not what we found as correlations between tests were rather weak. In addition, factor analysis revealed that a 4-factor model explained the data best, however, with no obvious relationship between factors and variables and low factor loadings. For almost all tests, the communalities were smaller than the uniqueness (Figure 5),

indicating that there is more unexplained than explained variance in the model. The 4 factors explained all together 29.1% of the total variance. As a comparison, the 1-factor model explains 13.1% of the total variance. Taken together these results provide no evidence for a common factor for vision.

Our results can be explained neither by a lack of power, given the large number of participants (N = 196 participants), nor by poor test-retest reliability. The visual backward masking and visual search tests showed high test-retest reliability, and the correlations between the different biological motion conditions and between different set-sizes of the visual search tests were also high.

Because of the cross-sectional design of our study, we do not draw conclusions on how perception changes with age. We only claim that we do not find evidence for a common factor for vision in older participants.

We can only speculate why there is more evidence for common factors for cognition than for vision. One reason may be that cognition developed much later in human evolution than vision, which may have led to a much higher degree of specialization in vision (for a review, see Johnson, 1990). Another reason might be that more brain areas interact when performing cognitive compared to visual tasks. The increase in dedifferentiation with age may lead to higher correlations between cognitive compared to visual tests (Cox et al., 2016; Tucker-Drob, 2009). In addition, age-related volumetric changes are widespread across the brain, with the exception of the primary visual cortex (Giorgio et al., 2010; Raz et al., 2005).

It is important to note that we only tested *healthy* participants, i.e., we sampled only from participants with good visual acuity (restricted range sampling), which is the vast majority of the population. There are strong correlations between visual tasks in people with optical or neurological pathologies because, for example, a person close to blindness performs poorly in all tests (Kurtenbach, Langrová, Messias, Zrenner, & Jägle, 2013).

The perceptual tests we chose are representative of visual functions and commonly used to understand visual processing. Gabors, for example, mimic receptive field profiles in primary visual cortex V1 (Jones & Palmer, 1987). Some of our tests are not considered *purely* perceptual because they also tap into other aspects such as inhibition and attention; one such example is the Simon task (Bialystok, Craik, Klein, & Viswanathan, 2004). Our goal was to have a battery of tests which would span a wide range of visual function complexity, from very basic to more complex. Future research is needed to confirm our results with different tests.

We were able to detect the absence of a unique common factor only because we used a large battery of tests. Previous research with older participants has usually employed only one or a few different paradigms and, thus, could not address a common factor hypothesis. Still, most studies support our notion of a lack of a unique common factor for perception because they find that visual abilities age differently and are not equally prone to age-related decline (e.g., Billino et al., 2008; Pilz et al., 2010).

We expected high correlations between perceptual and cognitive tests because differences in cognitive functions may strongly influence perception and vice-versa. For example, increased processing loads come with strong perceptual deficits and, hence, higher attentional and other cognitive demands (Pichora-Fuller, 2003). However, we found very low correlations in both younger and older participants between cognitive and perceptual tests (but see Anstey et al., 2003; Christensen et al., 2001; Lindenberger & Ghisletta, 2009). We also found only weak correlations between the WCST (measuring executive functions) and verbal fluency or digit span tests, most likely because these tests are measuring very different cognitive functions. These results are in accordance with previous studies (Fabiani & Friedman, 1997; Parkin & Lawrence, 1994; Stratta et al., 1997). However, McCabe et al. (2010), found that these three cognitive tests were linked under a common factor representing higher cognitive functions. Our data does not contribute to this discussion because we had only three cognitive tests and thus, cannot address the question of common factors for cognition.

Conclusion. We found only weak correlations between visual tests in both younger and older adults and no evidence for a common factor. This is good news because age-related deficits in one visual function do not imply deficits in other visual functions.

Competing financial interests: The author(s) declare no competing financial interests.

Acknowledgements: We thank Marc Repnow, Vitaly Chicherov and Janir da Cruz for their help with data analysis, Britt Anderson for his comments on the manuscript and Maya Jastrzebowska for proof reading of the manuscript. This work was generously supported by the Velux Foundation. Aaron Clarke was supported by a TÜBİTAK 2236 Co-Funded Brain Circulation Scheme grant, Project No. 116C037. We are sad to report that Dr. Aaron Clarke has passed away during the publication process.

Author Contributions Statement: KP and MH designed the study. MK and AS tested the participants. AFC, KN, AC and MH analysed the data. AS, KP, AK and MH wrote the manuscript. All authors reviewed the manuscript.

5. References

Agnew, H. C., Phillips, L. H., & Pilz, K. S. (2016). Global form and motion processing in healthy ageing. *Acta Psychologica, 166*, 12–20. <https://doi.org/10.1016/J.ACTPSY.2016.03.005>

Andersen, G. J. (2012). Aging and vision: changes in function and performance from optics to perception. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*(3), 403–410. <https://doi.org/10.1002/wcs.1167>

Anstey, K. J., Hofer, S. M., & Luszcz, M. A. (2003). A Latent Growth Curve Analysis of Late-Life Sensory and Cognitive Function Over 8 Years: Evidence for Specific and Common Factors Underlying Change. *Psychology and Aging, 18*(4), 714–726. <https://doi.org/10.1037/0882-7974.18.4.714>

Arena, A., Hutchinson, C. V., & Shimozaki, S. S. (2012). The effects of age on the spatial and temporal integration of global motion. *Vision Research, 58*, 27–32. <https://doi.org/10.1016/J.VISRES.2012.02.004>

Bach, M. (1996). The Freiburg Visual Acuity Test-Automatic Measurement of Visual Acuity. *Optometry and Vision Science, 73*(1), 49–53.

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging, 12*(1), 12–21. <https://doi.org/10.1037/0882-7974.12.1.12>

Banton, T., Dobkins, K., & Bertenthal, B. I. (2001). Infant direction discrimination thresholds. *Vision Research, 41*(8), 1049–1056. [https://doi.org/10.1016/S0042-6989\(01\)00027-X](https://doi.org/10.1016/S0042-6989(01)00027-X)

Bargary, G., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). Individual differences in human eye movements: An oculomotor signature? *Vision Research, 141*, 157–169. <https://doi.org/10.1016/j.visres.2017.03.001>

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5–17. <https://doi.org/10.1023/A:1005653411471>

Bechtoldt, H. P., Benton, A. L., & Fogel, M. L. (1962). An application of factor analysis in neuropsychology. *The Psychological Record, 12*, 147–156.

Bennett, P. J., Sekuler, R., & Sekuler, A. B. (2007). The effects of aging on motion detection and direction identification. *Vision Research*, 47(6), 799–809. <https://doi.org/10.1016/j.visres.2007.01.001>

Berg, E. A. (1948). A Simple Objective Technique for Measuring Flexibility in Thinking. *The Journal of General Psychology*, 39(1), 15–22. <https://doi.org/10.1080/00221309.1948.9918159>

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, Aging, and Cognitive Control: Evidence From the Simon Task. *Psychology and Aging*, 19(2), 290–303. <https://doi.org/10.1037/0882-7974.19.2.290>

Billino, J., Bremmer, F., & Gegenfurtner, K. R. (2008). Differential aging of motion processing mechanisms: Evidence against general perceptual decline. *Vision Research*, 48(10), 1254–1261. <https://doi.org/10.1016/j.visres.2008.02.014>

Bosten, J. M., Goodbourn, P. T., Bargary, G., Verhallen, R. J., Lawrance-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). An exploratory factor analysis of visual performance in a large population. *Vision Research*, 141, 303–316. <https://doi.org/10.1016/j.visres.2017.02.005>

Bosten, J. M., & Mollon, J. D. (2010). Is there a general trait of susceptibility to simultaneous contrast? *Vision Research*, 50(17), 1656–1664. <https://doi.org/10.1016/J.VISRES.2010.05.012>

Cappe, C., Clarke, A., Mohr, C., & Herzog, M. H. (2014). Is there a common factor for vision? *Journal of Vision*, 14(8), 4–4. <https://doi.org/10.1167/14.8.4>

Castel, A. D., Balota, D. A., Hutchison, K. A., Logan, J. M., & Yap, M. J. (2007). Spatial attention and response control in healthy younger and older adults and individuals with Alzheimer's disease: Evidence for disproportionate selection impairments in the simon task. *Neuropsychology*, 21(2), 170–182. <https://doi.org/10.1037/0894-4105.21.2.170>

Chamberlain, R., Van der Hallen, R., Huygelier, H., Van de Cruys, S., & Wagemans, J. (2017). Local-global processing bias is not a unitary individual difference in visual processing. *Vision Research*, 141, 247–257. <https://doi.org/10.1016/j.visres.2017.01.008>

Chkonia, E., Roinishvili, M., Makhatadze, N., Tsverava, L., Stroux, A., Herzog, M. H., & Brand, A. (2010). The Shine-Through Masking Paradigm Is a Potential Endophenotype of Schizophrenia, 5(12). <https://doi.org/10.1371/journal.pone.0014268>

Christensen, H., Mackinnon, A. J., Korten, A., & Jorm, A. F. (2001). The “common cause hypothesis” of cognitive aging: evidence for not only a common factor but also specific associations of age with vision and grip strength in a cross-sectional analysis. *Psychology and Aging, 16*(4), 588–599. <https://doi.org/10.1037/0882-7974.16.4.588>

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, New Jersey: L. Erlbaum Associates.

Conway, A. R., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence, 30*(2), 163–183. [https://doi.org/10.1016/S0160-2896\(01\)00096-4](https://doi.org/10.1016/S0160-2896(01)00096-4)

Cox, S. R., Ritchie, S. J., Tucker-Drob, E. M., Liewald, D. C., Hagenaars, S. P., Davies, G., ... Deary, I. J. (2016). Ageing and brain white matter structure in 3,513 UK Biobank participants. *Nature Communications, 7*, 13629. <https://doi.org/10.1038/ncomms13629>

Elliott, E. M., Cherry, K. E., Brown, J. S., Smitherman, E. A., Jazwinski, S. M., Yu, Q., & Volaufova, J. (2011). Working memory in the oldest-old: evidence from output serial position curves. *Memory & Cognition, 39*(8), 1423–1434. <https://doi.org/10.3758/s13421-011-0119-7>

Emery, K. J., Volbrecht, V. J., Peterzell, D. H., & Webster, M. A. (2017). Variations in normal color vision. VII. Relationships between color naming and hue scaling. *Vision Research, 141*, 66–75. <https://doi.org/10.1016/j.visres.2016.12.007>

Fabiani, M., & Friedman, D. (1997). Dissociations between memory for temporal order and recognition memory in aging. *Neuropsychologia, 35*(2), 129–141. [https://doi.org/10.1016/S0028-3932\(96\)00073-5](https://doi.org/10.1016/S0028-3932(96)00073-5)

Garcia-Suarez, L., Barrett, B. T., & Pacey, I. (2004). A comparison of the effects of ageing upon vernier and bisection acuity. *Vision Research, 44*(10), 1039–1045. <https://doi.org/10.1016/J.VISRES.2003.11.018>

Ghisletta, P., & Lindenberger, U. (2005). Exploring structural dynamics within and between sensory and intellectual functioning in old and very old age: Longitudinal evidence from the Berlin Aging Study. *Intelligence, 33*(6), 555–587. <https://doi.org/10.1016/J.INTELL.2005.07.002>

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74–78. <https://doi.org/10.1016/j.paid.2016.06.069>

Giorgio, A., Santelli, L., Tomassini, V., Bosnell, R., Smith, S., De Stefano, N., & Johansen-Berg, H. (2010). Age-related changes in grey and white matter structure throughout adulthood. *NeuroImage*, *51*(3), 943–951. <https://doi.org/10.1016/J.NEUROIMAGE.2010.03.004>

Grzeczowski, L., Clarke, A. M., Francis, G., Mast, F. W., & Herzog, M. H. (2017). About individual differences in vision. *Vision Research*, *141*, 282–292. <https://doi.org/10.1016/j.visres.2016.10.006>

Halpern, S. D., Andrews, T. J., & Purves, D. (1999). Interindividual variation in human visual performance. *Journal of Cognitive Neuroscience*, *11*(5), 521–534.

Herzog, M. H., & Koch, C. (2001). Seeing properties of an invisible object: feature inheritance and shine-through. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(7), 4271–4275. <https://doi.org/10.1073/pnas.071047498>

Herzog, M. H., Kopmann, S., & Brand, A. (2004). Intact figure-ground segmentation in schizophrenia. *Psychiatry Research*, *129*(1), 55–63. <https://doi.org/10.1016/j.psychres.2004.06.008>

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*(1), 11–26. <https://doi.org/10.1080/17470215208416600>

Johnson, M. H. (1990). Cortical Maturation and the Development of Visual Attention in Early Infancy. *Journal of Cognitive Neuroscience*, *2*(2), 81–95. <https://doi.org/10.1162/jocn.1990.2.2.81>

Jones, J. P., & Palmer, L. A. (1987). An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, *58*(6).

Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika*, *43*(4), 443–477. <https://doi.org/10.1007/BF02293808>

Kiely, K., & Anstey, K. (2017). Common Cause Theory in Aging. *Encyclopedia of Geropsychology*, 559–569.

Kurtenbach, A., Langrová, H., Messias, A., Zrenner, E., & Jägle, H. (2013). A comparison of the performance of three visual evoked potential-based methods to estimate visual acuity. *Documenta Ophthalmologica. Advances in Ophthalmology*, *126*(1), 45–56. <https://doi.org/10.1007/s10633-012-9359-5>

Lahav, K., Levkovitch-Verbin, H., Belkin, M., Glovinsky, Y., & Polat, U. (2011). Reduced mesopic and photopic foveal contrast sensitivity in glaucoma. *Archives of Ophthalmology*, *129*(1), 16–22. <https://doi.org/10.1001/archophthalmol.2010.332>

Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Nature Neuroscience*, *2*(4), 375–381. <https://doi.org/10.1038/7286>

Li, K. Z., & Lindenberger, U. (2002). Relations between aging sensory/sensorimotor and cognitive functions. *Neuroscience & Biobehavioral Reviews*, *26*(7), 777–783. [https://doi.org/10.1016/S0149-7634\(02\)00073-8](https://doi.org/10.1016/S0149-7634(02)00073-8)

Li, R. W., Edwards, M. H., & Brown, B. (2000). Variation in vernier acuity with age. *Vision Research*, *40*(27), 3775–3781. [https://doi.org/10.1016/S0042-6989\(00\)00212-1](https://doi.org/10.1016/S0042-6989(00)00212-1)

Lindenberger, U., & Baltes, P. B. (1994). Sensory functioning and intelligence in old age: A strong connection. *Psychology and Aging*, *9*(3), 339–355.

Lindenberger, U., & Ghisletta, P. (2009). Cognitive and sensory declines in old age: Gauging the evidence for a common cause. *Psychology and Aging*, *24*(1), 1–16. <https://doi.org/10.1037/a0014986>

McBain, R., Norton, D., & Chen, Y. (2010). Is Visual Motion Processing Vulnerable to Early Aging ? *The Open Behavioral Science Journal*, *4*, 26–30. <https://doi.org/10.2174/1874230001004010026>

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., Hambrick, D. Z., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology*, *24*(2), 222–243. <https://doi.org/10.1037/a0017619>

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>

Newman, L. M., Trivedi, M. A., Bendlin, B. B., Ries, M. L., & Johnson, S. C. (2007). The Relationship Between Gray Matter Morphometry and Neuropsychological Performance in a Large Sample of Cognitively Healthy Adults. *Brain Imaging and Behavior*, *1*(1–2), 3–10. <https://doi.org/10.1007/s11682-007-9000-5>

NIA. (2015). National Institute of Aging.

Norman, J. F., Payton, S. M., Long, J. R., & Hawkes, L. M. (2004). Aging and the Perception of Biological Motion. *Psychology and Aging, 19*(1), 219–225. <https://doi.org/10.1037/0882-7974.19.1.219>

Owsley, C., McGwin, G., & Seder, T. (2011). Older drivers' attitudes about instrument cluster designs in vehicles. *Accident Analysis & Prevention, 43*(6), 2024–2029. <https://doi.org/10.1016/j.aap.2011.05.021>

Park, D. C., & Reuter-Lorenz, P. (2009). The Adaptive Brain: Aging and Neurocognitive Scaffolding. *Annual Review of Psychology, 60*(1), 173–196. <https://doi.org/10.1146/annurev.psych.59.103006.093656>

Parkin, A. J., & Lawrence, A. (1994). A dissociation in the relation between memory tasks and frontal lobe tests in the normal elderly. *Neuropsychologia, 32*(12), 1523–1532. [https://doi.org/10.1016/0028-3932\(94\)90124-4](https://doi.org/10.1016/0028-3932(94)90124-4)

Peterzell, D. H. (2016). Discovering Sensory Processes Using Individual Differences : A Review and Factor Analytic Manifesto. *Electronic Imaging, 2016*(16), 1–11. <https://doi.org/10.2352/ISSN.2470-1173.2016.16HVEI-112>

Pichora-Fuller, Kathleen, M. (2003). Cognitive aging and auditory information processing. *International Journal of Audiology, 42*(2), 26–32.

Pilz, K. S., Bennett, P. J., & Sekuler, A. B. (2010). Effects of aging on biological motion discrimination. *Vision Research, 50*(2), 211–219. <https://doi.org/10.1016/j.visres.2009.11.014>

Pilz, K. S., Kunchulia, M., Parkosadze, K., & Herzog, M. H. (2015). Ageing and visual spatiotemporal processing. *Experimental Brain Research, 233*(8), 2441–2448. <https://doi.org/10.1007/s00221-015-4314-9>

Pilz, K. S., Miller, L., & Agnew, H. C. (2017). Motion coherence and direction discrimination in healthy aging. *Journal of Vision, 17*(1), 31–31. <https://doi.org/10.1167/17.1.31>

Raz, D., Perlman, I., Percicot, C. L., Lambrou, G. N., & Ofri, R. (2003). Functional Damage to Inner and Outer Retinal Cells in Experimental Glaucoma. *Investigative Ophthalmology & Visual Science, 44*(8), 3675. <https://doi.org/10.1167/iovs.02-1236>

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., ... Acker, J. D. (2005). Regional Brain Changes in Aging Healthy Adults: General Trends, Individual Differences and Modifiers. *Cerebral Cortex*, *15*(11), 1676–1689. <https://doi.org/10.1093/cercor/bhi044>

Roinishvili, M., Chkonia, E., Stroux, A., Brand, A., & Herzog, M. H. (2011). Combining vernier acuity and visual backward masking as a sensitive test for visual temporal deficits in aging research. *Vision Research*, *51*(4), 417–423. <https://doi.org/10.1016/j.visres.2010.12.011>

Roudaia, E., Bennett, P. J., & Sekuler, A. B. (2013). Contour integration and aging: the effects of element spacing, orientation alignment and stimulus duration. *Frontiers in Psychology*, *4*, 356. <https://doi.org/10.3389/fpsyg.2013.00356>

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*. <https://doi.org/10.1016/j.neurobiolaging.2008.09.023>

Salthouse, T. A., & Czaja, S. J. (2000). Structural constraints on process explanations in cognitive aging. *Psychology and Aging*, *15*(1), 44–55. <https://doi.org/10.1037/0882-7974.15.1.44>

Salthouse, T. A., Hancock, H. E., Meinz, E. J., & Hambrick, D. Z. (1996). Interrelations of Age, Visual Acuity, and Cognitive Functioning. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *51B*(6), P317-330. <https://doi.org/10.1093/geronb/51B.6.P317>

Schulze-Bonsel, K., Feltgen, N., Burau, H., Hansen, L., & Bach, M. (2006). Visual Acuities “Hand Motion” and “Counting Fingers” Can Be Quantified with the Freiburg Visual Acuity Test. *Investigative Ophthalmology & Visual Science*, *47*(3), 1236. <https://doi.org/10.1167/iovs.05-0981>

Spear, P. D. (1993). Neural bases of visual deficits during aging. *Vision Research*, *33*(18), 2589–2609. [https://doi.org/10.1016/0042-6989\(93\)90218-L](https://doi.org/10.1016/0042-6989(93)90218-L)

Spencer, J. M. Y., Sekuler, A. B., Bennett, P. J., Giese, M. A., & Pilz, K. S. (2016). Effects of aging on identifying emotions conveyed by point-light walkers. *Psychology and Aging*, *31*(1), 126–138. <https://doi.org/10.1037/a0040009>

Stratta, P., Daneluzzo, E., Mattei, P., Bustini, M., Casacchia, M., & Rossi, A. (1997). No deficit in Wisconsin Card Sorting Test performance of schizophrenic patients' first-degree relatives. *Schizophrenia Research*, *26*(2–3), 147–151. [https://doi.org/10.1016/S0920-9964\(97\)00047-9](https://doi.org/10.1016/S0920-9964(97)00047-9)

Strauss, E., Sherman, E. M. S., Spreen, O., & Spreen, O. (2006). *A compendium of neuropsychological tests : administration, norms, and commentary*. Oxford University Press.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *The Journal of the Acoustical Society of America*, *41*(4A), 782–787.

Thorpe, S. J., & Fabre-Thorpe, M. (2001). Seeking Categories in the Brain. *Science*, *291*(5502).

Tibber, M. S., Guedes, A., & Shepherd, A. J. (2006). Orientation Discrimination and Contrast Detection Thresholds in Migraine for Cardinal and Oblique Angles. *Investigative Ophthalmology & Visual Science*, *47*(12), 5599. <https://doi.org/10.1167/iovs.06-0640>

Treutwein, B. (1995). Adaptive Psychophysical Procedures. *Vision Research*, *35*(17), 2503–2522.

Tucker-Drob, E. M. (2009). Differentiation of cognitive abilities across the life span. *Developmental Psychology*, *45*(4), 1097–1118. <https://doi.org/10.1037/a0015864>

Ward, J., Rothen, N., Chang, A., & Kanai, R. (2017). The structure of inter-individual differences in visual ability : Evidence from the general population and synaesthesia. *Vision Research*, *141*, 293–302. <https://doi.org/10.1016/j.visres.2016.06.009>

Wesemann, W. (2002). Visual acuity measured via the Freiburg visual acuity test (FVT), Bailey Lovie chart and Landolt Ring chart. *Klinische Monatsblätter Fur Augenheilkunde*, *219*(9), 660–667. <https://doi.org/10.1055/s-2002-35168>

Wood, A. G., Saling, M. M., Abbott, D. F., & Jackson, G. D. (2001). A Neurocognitive Account of Frontal Lobe Involvement in Orthographic Lexical Retrieval: An fMRI Study. *NeuroImage*, *14*(1), 162–169. <https://doi.org/10.1006/nimg.2001.0778>

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, *17*(1), 37–49. [https://doi.org/10.1016/0022-3956\(82\)90033-4](https://doi.org/10.1016/0022-3956(82)90033-4)

6. Supplemental Material

Table S1. Lambda exponents from an optimized power transformation that maximizes normality for each test according to the Shapiro-Wilk test.

	FrACT	VbFlu	DigitF	DigitB	WCST	SimpRT	Simon	VernDur	VernOfs	BMask25	BMask5	Contr	Motion	Bio200i	Bio800i	Bio200	Bio800	Orient	VSrchSL	VSrchRT
λ	0.38	0.55	0.28	0.28	-0.42	-2.00	3.48	0*	-0.77	-0.09*	-0.46	0.09		2.01*				-0.16	-0.10	-1.31

* λ_{VernDur} was manually set to 0 because strong floor effects did not lead to reasonable fits of λ . The λ s for BMask5 and BMask25 were constrained to be the same for both tasks. Likewise, the λ s for the Bio tasks were constrained to be identical. The original Simon test values (x) were converted to positive values by $x'=0.5+x/100$ before fitting λ .

Table S2. Correlation matrices used to run factor analysis for both young (upper triangle, N = 104) and old (lower triangle, N = 92) participants. Similar tests are merged, missing data imputed and all data are z-transformed.

	FrACT	VbFlu	Digit	WCST	SimpRT	Simon	Vern	BMask	Contr	Motion	BioInv	Bio	Orient	VSrch
FrACT	1	0.0669	0.0064	0.0448	0.1913	-0.0965	0.2147	0.2801	0.3944	0.0103	0.0543	0.0604	0.1522	-0.0633
VbFlu	-0.0131	1	0.3750	0.3165	0.2140	0.0171	0.2302	0.1761	0.1036	0.3973	0.1635	0.2562	0.2922	0.2066
Digit	-0.1948	0.2291	1	0.3871	0.1194	0.2410	0.1518	0.2235	0.0845	0.3908	0.3247	0.3129	0.2973	0.1871
WCST	0.1565	0.1807	0.1332	1	0.3488	0.1472	0.2022	0.2051	-0.0897	0.2289	0.2861	0.1610	0.2558	0.1209
SimpRT	0.0534	-0.0226	0.1720	-0.0779	1	-0.1300	0.2614	0.3469	0.0291	0.0696	0.2411	0.1661	0.2486	0.2721
Simon	-0.0183	0.0017	-0.0528	0.1690	-0.1464	1	-0.0070	0.1737	-0.1047	0.1993	0.3484	0.1567	0.0212	-0.0072
Vern	0.2935	-0.0840	-0.0075	0.1407	0.1344	0.0538	1	0.5371	0.1961	0.4345	0.3138	0.2171	0.3790	0.2442
BMask	0.1249	0.0195	0.1470	0.0138	0.1529	-0.0434	0.0712	1	0.2030	0.2238	0.3396	0.0705	0.3126	-0.0608
Contr	0.4971	0.0666	-0.0909	0.0848	0.0395	0.1105	0.2836	0.2818	1	0.1385	-0.0561	-0.2026	0.2005	0.1598
Motion	0.0657	0.0683	0.0697	0.1990	0.1038	0.1375	0.1389	0.1228	0.1667	1	0.4032	0.2767	0.3141	0.2435
BioInv	0.0778	0.0683	-0.0017	0.1183	0.0687	0.0764	0.1919	0.0286	0.0288	0.1017	1	0.3212	0.2241	0.0788
Bio	0.0449	-0.0074	0.0561	0.0693	0.1899	0.0282	0.0651	-0.0587	-0.0679	0.3694	0.3070	1	0.1992	0.1987
Orient	0.3219	0.0419	0.1776	0.2353	0.2695	0.1253	0.4632	0.1829	0.3325	0.2621	0.2006	0.1353	1	0.1789
VSrch	-0.0482	0.1103	0.2094	-0.0401	0.1007	-0.0018	0.1560	-0.0653	0.0895	0.1158	-0.0577	0.0168	0.3243	1

Table S3. Statistics for estimated coefficients of the CFA for the 3-factor model for the young population.

Coeff	Estimate	SE	z value	Pr(> z)	Path
F1_VbFlu	0.55	0.113	4.83	1.39E-06	VbFlu <--- F1
F1_Digit	0.45	0.113	4.00	6.21E-05	Digit <--- F1
F1_WCST	0.40	0.115	3.48	5.00E-04	WCST <--- F1
F1_SimpRT	0.32	0.113	2.82	4.87E-03	SimpRT <--- F1
F1_Vern	0.38	0.101	3.73	1.92E-04	Vern <--- F1
F1_Motion	0.53	0.109	4.90	9.59E-07	Motion <--- F1
F1_BioInv	0.28	0.114	2.46	1.40E-02	BioInv <--- F1
F1_Bio	0.42	0.115	3.65	2.64E-04	Bio <--- F1
F1_Orient	0.43	0.110	3.94	8.04E-05	Orient <--- F1
F1_VSrch	0.50	0.115	4.34	1.45E-05	VSrch <--- F1
F2_FrACT	0.41	0.110	3.73	1.88E-04	FrACT <--- F2
F2_SimpRT	0.37	0.107	3.41	6.39E-04	SimpRT <--- F2
F2_Vern	0.58	0.102	5.65	1.56E-08	Vern <--- F2
F2_BMask	0.80	0.113	7.04	1.90E-12	BMask <--- F2
F2_Contr	0.34	0.111	3.05	2.29E-03	Contr <--- F2
F2_Motion	0.11	0.099	1.10	2.70E-01	Motion <--- F2
F2_BioInv	0.17	0.105	1.58	1.13E-01	BioInv <--- F2
F2_Orient	0.31	0.104	3.03	2.45E-03	Orient <--- F2
F3_VbFlu	0.16	0.125	1.31	1.90E-01	VbFlu <--- F3
F3_Digit	0.46	0.121	3.78	1.54E-04	Digit <--- F3
F3_WCST	0.30	0.123	2.40	1.63E-02	WCST <--- F3
F3_Simon	0.63	0.129	4.90	9.56E-07	Simon <--- F3
F3_Vern	0.16	0.120	1.36	1.73E-01	Vern <--- F3
F3_BMask	0.41	0.116	3.55	3.90E-04	BMask <--- F3
F3_Motion	0.37	0.122	3.01	2.57E-03	Motion <--- F3
F3_BioInv	0.56	0.122	4.62	3.83E-06	BioInv <--- F3
F3_Bio	0.27	0.124	2.19	2.86E-02	Bio <--- F3
F3_Orient	0.14	0.124	1.13	2.57E-01	Orient <--- F3
V[FrACT]	0.83	0.125	6.66	2.72E-11	FrACT <--> FrACT
V[VbFlu]	0.68	0.115	5.94	2.81E-09	VbFlu <--> VbFlu
V[Digit]	0.60	0.104	5.76	8.63E-09	Digit <--> Digit
V[WCST]	0.76	0.116	6.53	6.55E-11	WCST <--> WCST
V[SimpRT]	0.75	0.118	6.34	2.33E-10	SimpRT <--> SimpRT
V[Simon]	0.60	0.145	4.17	3.06E-05	Simon <--> Simon
V[Vern]	0.48	0.096	5.04	4.73E-07	Vern <--> Vern
V[BMask]	0.22	0.137	1.64	1.02E-01	BMask <--> BMask
V[Contr]	0.89	0.129	6.87	6.61E-12	Contr <--> Contr
V[Motion]	0.58	0.099	5.84	5.27E-09	Motion <--> Motion
V[BioInv]	0.59	0.107	5.53	3.15E-08	BioInv <--> BioInv
V[Bio]	0.76	0.116	6.52	7.20E-11	Bio <--> Bio
V[Orient]	0.68	0.110	6.22	5.02E-10	Orient <--> Orient
V[VSrch]	0.75	0.125	6.04	1.54E-09	VSrch <--> VSrch

Table S4. Statistics for estimated coefficients of the CFA for the 4-factor model for the old population.

Coeff	Estimate	SE	z value	Pr(> z)	Path
F1_FrACT	0.67	0.116	5.80	6.56E-09	FrACT <--- F1
F1_SimpRT	0.15	0.118	1.23	2.18E-01	SimpRT <--- F1
F1_Vern	0.47	0.115	4.11	3.95E-05	Vern <--- F1
F1_BMask	0.28	0.122	2.27	2.30E-02	BMask <--- F1
F1_Contr	0.69	0.116	5.93	3.09E-09	Contr <--- F1
F1_Orient	0.55	0.106	5.15	2.62E-07	Orient <--- F1
F2_SimpRT	0.28	0.129	2.21	2.72E-02	SimpRT <--- F2
F2_Vern	0.24	0.123	1.93	5.40E-02	Vern <--- F2
F2_Motion	0.48	0.136	3.53	4.19E-04	Motion <--- F2
F2_BioInv	0.42	0.134	3.14	1.68E-03	BioInv <--- F2
F2_Bio	0.66	0.148	4.45	8.67E-06	Bio <--- F2
F2_Orient	0.32	0.112	2.81	5.03E-03	Orient <--- F2
F3_VbFlu	0.24	0.141	1.68	9.31E-02	VbFlu <--- F3
F3_Digit	0.56	0.152	3.68	2.36E-04	Digit <--- F3
F3_SimpRT	0.31	0.136	2.27	2.29E-02	SimpRT <--- F3
F3_BMask	0.14	0.138	1.03	3.02E-01	BMask <--- F3
F3_Orient	0.45	0.125	3.63	2.87E-04	Orient <--- F3
F3_VSrch	0.44	0.143	3.10	1.93E-03	VSrch <--- F3
F4_VbFlu	0.22	0.165	1.35	1.77E-01	VbFlu <--- F4
F4_WCST	0.47	0.203	2.32	2.01E-02	WCST <--- F4
F4_SimpRT	-0.41	0.185	-2.22	2.65E-02	SimpRT <--- F4
F4_Simon	0.36	0.175	2.05	4.07E-02	Simon <--- F4
V[FrACT]	0.55	0.124	4.44	9.11E-06	FrACT <--> FrACT
V[VbFlu]	0.89	0.154	5.79	6.93E-09	VbFlu <--> VbFlu
V[Digit]	0.69	0.166	4.15	3.35E-05	Digit <--> Digit
V[WCST]	0.78	0.203	3.83	1.29E-04	WCST <--> WCST
V[SimpRT]	0.67	0.178	3.75	1.73E-04	SimpRT <--> SimpRT
V[Simon]	0.87	0.165	5.29	1.25E-07	Simon <--> Simon
V[Vern]	0.71	0.123	5.79	6.86E-09	Vern <--> Vern
V[BMask]	0.91	0.140	6.49	8.31E-11	BMask <--> BMask
V[Contr]	0.53	0.125	4.25	2.14E-05	Contr <--> Contr
V[Motion]	0.77	0.145	5.30	1.19E-07	Motion <--> Motion
V[BioInv]	0.82	0.143	5.73	9.94E-09	BioInv <--> BioInv
V[Bio]	0.57	0.174	3.26	1.12E-03	Bio <--> Bio
V[Orient]	0.40	0.117	3.40	6.70E-04	Orient <--> Orient
V[VSrch]	0.80	0.150	5.36	8.52E-08	VSrch <--> VSrch

Appendix F

Cretenoud, A. F., Barakat, A., Milliet, A., Choung, O.-H., Bertamini, M., Constantin, C., & Herzog, M. H. (submitted). How do visual skills relate to action video game performance?

This manuscript is currently under review at *Journal of Vision*.

Copyright: Reprinted with permission from ARVO

How do visual skills relate to action video game performance?

Aline F. Cretenoud¹, Arthur Barakat^{1,2,3}, Alain Milliet³, Oh-Hyeon Choung¹, Marco Bertamini^{4,5}, Christophe Constantin³, & Michael H. Herzog¹

¹Laboratory of Psychophysics, Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

²Laboratory of Behavioral Genetics, Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

³Logitech Europe S.A., Innovation Park EPFL, Lausanne, Switzerland

⁴Department of Psychological Sciences, University of Liverpool, UK

⁵Department of General Psychology, University of Padova, Italy

Abstract

It has been claimed that video gamers possess increased perceptual and cognitive skills compared to non-video gamers. Here, we examined to which extent gaming performance in CS:GO (Counter-Strike: Global Offensive) correlates with visual performance. We tested 94 players ranging from beginners to experts with a battery of visual paradigms, such as visual acuity and contrast detection. In addition, we assessed performance in specific gaming skills, such as shooting and tracking, and administered personality traits. All measures together explained more than 70% of the variance of the players' ranks. In particular, regression models showed that a few visual abilities, such as visual acuity in the periphery and the susceptibility to the Honeycomb illusion, were strongly associated with the players' ranks. While the causality of the effect remains unknown, our results suggest that certain visual skills are enhanced in high-rank players.

Keywords: action video games, vision, common factor, prediction

1. Introduction

Basic visual skills, such as contrast detection and orientation discrimination, are the building blocks for visual processing. It has been suggested that playing video games is associated with better performance in these basic perceptual abilities (for reviews, see Bavelier et al., 2012; Bediou et al., 2018; Boot et al., 2011; Chopin et al., 2019). For example, Hutchinson and Stocks (2013) observed that action video game (i.e., a subset of video games, which rely on physical challenges such as hand-eye coordination and reaction time) players (AVGPs) performed better in a random-dot kinematograms task compared to non-video game players (NVGPs), suggesting that AVGPs are better at global motion detection. Likewise, AVGPs were observed to have improved perceptual speed (Dye et al., 2009) in the Test of Variables of Attention (T.O.V.A.[®]) compared to NVGPs, while the speed-accuracy tradeoff was similar in both groups. Li, Polat, Scalzo, and Bavelier (2010) trained participants with video games to establish the causal effect of action gaming on temporal dynamics and observed reduced backward masking performance (i.e., reduced threshold elevation in a masked contrast detection task) in video game players (VGPs) compared to NVGPs. In addition, VGPs outperformed NVGPs in other perceptual skills, such as multiple object tracking (Green & Bavelier, 2006), task-switching (Shawn Green et al., 2012), spatial resolution (Green & Bavelier, 2007), and contrast sensitivity (Li et al., 2009). These studies were either intervention studies (e.g., Li et al., 2010), i.e., participants were trained with a specific video game, or cross-sectional (e.g., Hutchinson & Stocks, 2013).

Studies have also examined the benefits of playing video games on cognitive abilities (for reviews, see Bavelier et al., 2012; Bediou et al., 2018; Campbell et al., 2018; Spence & Feng, 2010). For example, VGPs showed enhanced change detection performance compared to NVGPs (Clark et al., 2011). Kowal, Toth, Exton, and Campbell (2018) tested AVGPs and NVGPs with a Stroop test, which tests inhibitive abilities, and a Trail-Making test (TMT), which measures processing speed and task-switching abilities. In both tasks, AVGPs showed significantly faster reaction times compared to NVGPs, suggesting that processing speed and task-switching abilities are enhanced in AVGPs. However, AVGPs made significantly more errors in the Stroop test compared to NVGPs (no significant difference in the TMT test), which indicates that inhibitive abilities are boosted at the expense of the speed-accuracy tradeoff.

Perceptual learning studies have often reported dramatic improvements in perceptual sensitivity, such as in vernier (Fahle & Edelman, 1993; Spang et al., 2010), orientation (Schoups et al., 1995; Shiu & Pashler, 1992), and motion (Ball & Sekuler, 1987) discrimination. For example, participants improved performance when trained with a bisection stimulus, i.e., they were able to discriminate smaller offsets after training (e.g., Aberg & Herzog, 2009; Grzeczowski et al., 2017, 2019). Video gamers may similarly

be exposed to learning effects (see Shawn Green et al., 2010), resulting in substantial positive effects in both perceptual and cognitive skills (but see Ferguson, 2007).

However, most studies focused on only one – or very few – task(s) and thus it is unclear whether gaming performance is related to some specific skills or to a common factor. In the latter case, we expect to find strong correlations between gamers' performance in visual tasks. However, this prediction is in contrast with the weak evidence for a unique common factor for vision (e.g., Cappe, Clarke, Mohr, & Herzog, 2014; but see Bosten et al., 2017; for reviews, see Mollon, Bosten, Peterzell, & Webster, 2017; Tulver, 2019). It seems that visual perception is highly multifactorial. For example, there were only weak correlations between the susceptibility to different illusions, while strong correlations were observed between different variants of the same illusion, suggesting that there are illusion-specific factors (Cretenoud et al., 2019; Cretenoud, Francis, et al., 2020; Cretenoud, Grzeczowski, et al., 2020; Grzeczowski et al., 2017). Similarly, there seems to be no unique common factor in eye movements (Bargary et al., 2017), hue scaling (e.g., Emery, Volbrecht, Peterzell, & Webster, 2017), and contrast perception (Bosten & Mollon, 2010; Peterzell, 2016; Peterzell et al., 2000).

The popularity of electronic sports (esports), which are video games played at a competitive – sometimes professional – level, has exploded in the last decades with a growing interest in athletes' performance (e.g., Wagner, 2006). One of the leading esports is Counter-Strike: Global Offensive (CS:GO; e.g., Nazhif Rizani & Iida, 2019), a first person shooter (FPS) action video game (see Figure S1 in the Supplementary File), in which players are split into two groups, i.e., terrorists and counter-terrorists. Players are usually matched against other players with similar ranks.

Specific motor and cognitive abilities are required to play these video games. For example, flicking, i.e., the motor coordination between the player's move and shooting via the computer mouse, is crucial to eliminate the players of the opposite team in FPS. Because some of these aspects rely on low-level visual skills (e.g., detecting an enemy strongly relies on vision and detection in the periphery), it is of interest to examine the different aspects of the game and their relationship with basic visual tasks.

To the best of our knowledge, all studies measuring visual abilities in VGPs compared performances between groups, i.e., VGPs and NVGPs. However, there is no well-defined criterion to discriminate a VGP from a NVGP. For example, Hutchinson and Stocks (2013) considered participants who played video games for more than 10 hours per week as VGPs, whereas 5 hours per week during the last six months was sufficient in Green and Bavelier (2007). Here, we tested a broad range of CS:GO players, i.e., from low- to high-rank players, with a battery of different visual tasks to examine what aspects are

associated with expertise, and whether there is a unique, common factor underlying visual perception in AVGPs.

2. Materials and Methods

2.1. Participants

Ninety-four participants were recruited (18-35 years; $M = 21.9$; $SD = 3.2$). All participants were AVGPs, played CS:GO at least once in the six months before the experiment, and had a CS:GO rank. Participants signed informed consent prior to the experiment and were paid 20 Swiss Francs per hour. Procedures were conducted in accordance with the Declaration of Helsinki, except for pre-registration (§ 35), and were approved by the local ethics committee.

2.2. Procedure

The experiment consisted of four parts. First, participants answered a survey about their gaming experience. Participants reported their actual CS:GO rank ($M = 9.0$, $SD = 5.2$; Figure 1), best CS:GO rank ever ($M = 11.8$, $SD = 4.9$; if “best” is not specified, we later refer to the actual ranks; Figure 1), and the total ($M = 1232$, $SD = 1879$) and weekly ($M = 13.8$, $SD = 11.7$) number of hours they played CS:GO. Note that ordinal ranks were converted to numerical equivalents from 1 to 18, with 18 being the highest rank (see Table S1 in the Supplementary File).

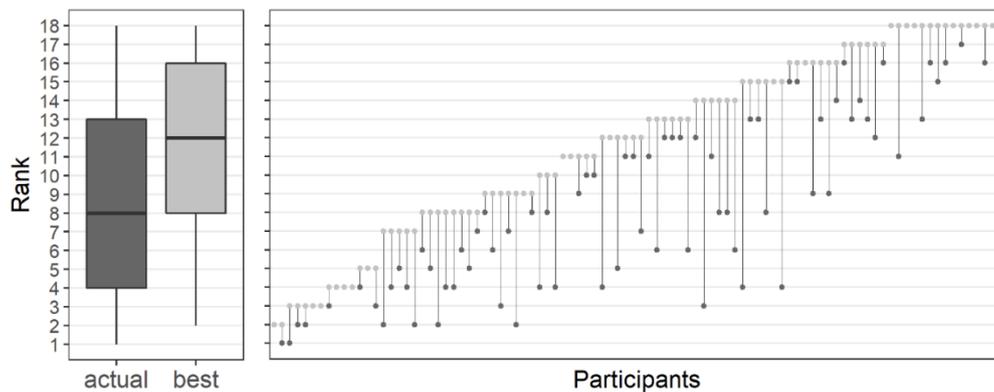


Figure 1. Actual (dark grey) and best (light grey) CS:GO ranks summarized in boxplots (left panel) and shown for each participant (right panel). The higher the rank, the better.

Table 1. Visual paradigms

Paradigm	Variable	Procedure	Lighting condition	Test-retest	Feedback	Distance to screen (cm)	Nb of trials	Nb of training trials
Crw	CrwSize	PEST	dim	No	Negative	60	96	
	CrwPeri						160	32
Ctr	Ctr	PEST	dim	No	Negative	200	80	
Hc/Ex	HcBlack	Adjustment	on	Yes	No	60	2	1
	HcWhite						2	
	ExBlack						2	1
	ExWhite						2	
III	CS left	Adjustment	on	Yes	No	60	2	1
	CS right						2	
	EB small						2	1
	EB large						2	
	ML in						2	1
	ML out						2	
	PD left						2	1
	PD right						2	
	PZ down						2	1
	PZ up						2	
	TT left						2	1
	TT right						2	
	VH hor						2	1
	VH ver						2	
	WH left						2	1
	WH right						2	
	ZN left						2	1
	ZN right						2	
NB	NB		dim	No	Negative	60	40	10
Ori	Ori	PEST	dim	No	Negative	200	80	
RDK	RDKh	QUEST	on	No	Negative	100	80	4
	RDKr						80	4
RT	RT		dim	No	No	200	80	
Sac	proTravel		on	No	Positive and negative	100	16	4
	proSac						16	
	antiTravel						16	4
	antiSac						16	
VA	VA	QUEST	dim	Yes	No	500	24	24
VBM	VBM	PEST	dim	Yes	Negative	200	80	
VS	VS4		dim	No	Negative	200	40	
	VS16						40	

Trials with response times longer than 3 sec after the stimulus onset were replaced in the Crw, Ctr, Ori, RDK, VBM, and VS paradigms. In the Sac paradigm, positive or negative feedback was provided at the end of each trial as a happy or sad smiley, respectively. In contrast, negative auditory feedback was provided in the Crw, Ctr, NB, Ori, RDK, VBM, and VS paradigms.

Second, participants performed a battery of 12 visual paradigms: crowding (Crw), contrast sensitivity (Ctr), the Honeycomb and Extinction illusions (Hc/Ex), a battery of other illusions (Ill), N-back (NB), orientation discrimination (Ori), random dot kinematograms (RDK), simple reaction times (RT), pro- and anti-saccades (Sac), Freiburg visual acuity (VA), visual backward masking (VBM), and visual search (VS). The visual paradigms were presented in random order.

Several variables were extracted for six out of the 12 paradigms. For instance, the visual search paradigm was tested with two conditions, i.e., with either four or 16 distractors. In total, 38 variables were considered for further analysis, and are listed in Table 1. When psychometric functions were used (Crw, Ctr, Ori, RDK, VA, and VBM paradigms), we discarded blocks when the fit was invalid, i.e., when the point of subjective equality (PSE) was outside of the search space, the goodness of fit < 0.05 , or the process did not converge. 1.3% of values were discarded.

Third, participants' gaming skills were assessed through six CS:GO mini-games, which were developed by Logitech (Lausanne, Switzerland) in collaboration with the University of Limerick (Ireland) and are publicly available on playmaster.gg. Playmaster is a training space for CS:GO that tests and compares gaming skills among the community and professionals. We extracted six gaming skills, namely flicking, holding, peeking, shooting, spraying, and tracking (see Figure S1 in the Supplementary File), as weighted sums of different features measured in the mini-games.

Fourth and last, participants answered seven self-report questionnaires, which were presented in random order: the Autism-Spectrum Quotient questionnaire (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), which consists of 50 items; a short version of the Liverpool Inventory of Feelings and Experiences questionnaire (O-LIFE; Mason, Linney, & Claridge, 2005), which investigates positive and negative schizotypy traits with 43 items; the short revised HEXACO personality inventory (HEXACO-60; Ashton & Lee, 2009), which measures 60 items of the six major dimensions of personality (HH: honesty-humility; EM: emotionality; EX: extraversion; AG: agreeableness; CO: conscientiousness; OP: openness to experience); the short version of the Barratt Impulsiveness Scale (BIS; Spinella, 2007), which measures impulsivity with 15 items; the Competitiveness Index (CI; Harris & Houston, 2010; Smither & Houston, 1992), which assesses competitive behavior with 14 items measured on a 5-point Likert scale; the Edinburgh handedness inventory (Oldfield, 1971), which assesses participants' hand dominance; and the Personality Research Form dominance subscale (PRFd; Jackson, 1974), which examines social dominance motivation with a 16-item true or false questionnaire. The AQ, BIS, CI, HEXACO, and O-LIFE questionnaires comprise several subscales. Participants could choose between English and French versions of the questionnaires.

Visual tasks and questionnaires were completed at EPFL individually in a quiet room. Because of technical issues, seven participants had to perform the gaming tasks in a gaming room at Logitech (Innovation Park, Switzerland), while the others ($n = 87$) performed the gaming tasks at EPFL. Participants were informed that they could quit the experiment at any time. The experimenter stayed in the experimental (or gaming) room with the participant and answered questions at any time.

2.3. Apparatus

Stimuli were presented on a BenQ XL2540 LCD monitor (resolution of 1920 x 1080 pixels; screen size: 24.5") with a refresh rate of 240 Hz. Gaming tasks were performed on an ASUS VG248QE monitor (resolution of 1920 x 1080 pixels; screen size: 24") with a refresh rate of 144 Hz.

2.4. Visual paradigms

Table 1 summarizes details for each visual paradigm, such as the distance to the screen and the light conditions. Stimulus luminance varied between 1 cd/m² (black) and 98 cd/m² (white).

2.4.1. Crowding

A paradigm similar to the one used in Green and Bavelier (2007) was performed. First, an E optotype was shown in the periphery, while participants fixated a red dot in the center of the screen (Figure 2a). The red dot was presented for 250 ms. The E optotype was shown for 150 ms with a delay of 100 ms compared to the red dot and at an eccentricity of 10 arcdeg to the right of the red dot. Participants were asked to report the orientation of the optotype within 3 secs by using push buttons, i.e., either standard (right button) or mirrored (left button) orientation. Using an adaptive staircase procedure (parameter estimation by sequential tracking PEST; starting value: 65 arcmin; range value: 10 to 200 arcmin; Taylor & Creelman, 1967), the stimulus size was varied to reach a threshold of 80% of correct responses (CrwSize).

Second, the task was the same as before and two optotype distractors were added above and below the optotype target. The distractors were randomly oriented in one of the four cardinal directions, and the orientations were counterbalanced in a full factorial fashion. The size of the target was fixed according to the first part of the paradigm (i.e., CrwSize) and the distance between the target and the distractors was manipulated using a PEST procedure (starting value: 200 arcmin; range value: 6 x CrwSize to 300 arcmin) to reach 75% of correct responses (CrwPeri).

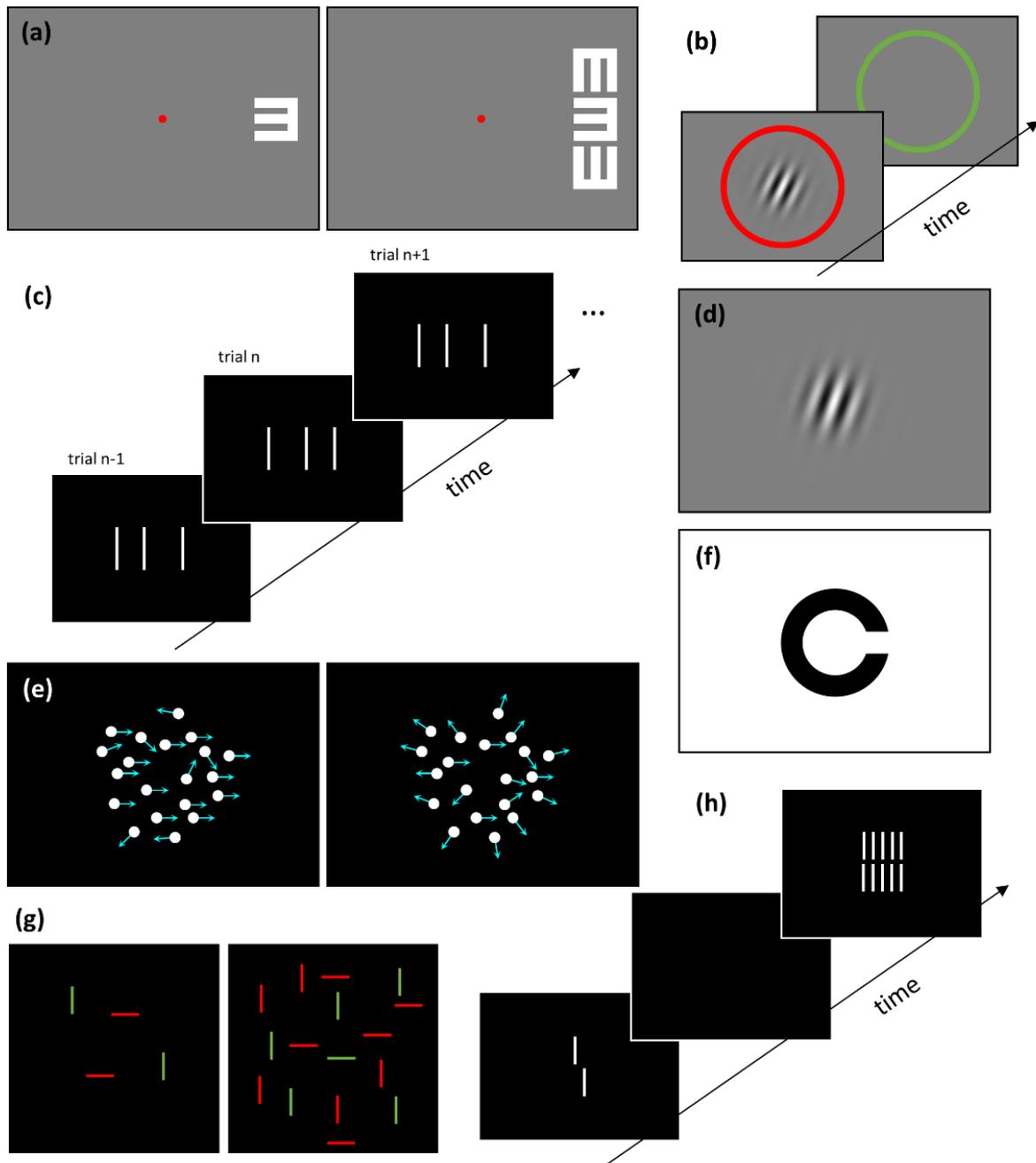


Figure 2. Schematic and exemplary representations of some of the visual paradigms tested. (a) Crowding: the size of the E optotype (left panel) and the distance between the target and distracting optotypes (right panel) varied according to a staircase procedure; (b) contrast sensitivity; (c) N-back with $N = 1$; (d) orientation discrimination; (e) horizontal (left panel) and radial (right panel) random dot kinematograms (cyan arrows indicate motion direction and were not part of the stimulus); (f) Freiburg visual acuity; (g) visual search (left panel: 4-line condition; right panel: 16-line condition); (h) visual backward masking with a 5-element grading.

2.4.2. Contrast sensitivity

Contrast sensitivity was measured with a 2IFC task (see Lahav, Levkovitch-Verbin, Belkin, Glovinsky, & Polat, 2011). A red fixation dot was presented in the middle of the screen. A red and a green circle (2 arcdeg in diameter) were then presented subsequently. Participants indicated in which circle a Gabor patch (spatial frequency: 4.0 cy/arcdeg; duration of presentation: 100 ms; envelope sigma: 0.30 arcdeg; Figure 2b) was presented by pushing a red or green push button, respectively. The mean luminance was 50% and Gabors were rendered using dithering in order to virtually increase gray level resolution. A PEST procedure (starting value: 10%) was used to measure the contrast threshold level at which participants reached 75% of correct responses.

2.4.3. Honeycomb and Extinction illusions

This paradigm was based on a previous study by Bertamini, Herzog, and Bruno (2016; see also Bertamini, Cretenoud, & Herzog, 2019). The Honeycomb and Extinction illusions are characterized by an inability of the participants to see shapes (barbs in the case of the Honeycomb illusion; dots in the case of the Extinction illusion) in the periphery of a uniform texture. The background image (Figure 3a-d) filled the screen. While fixating a red central cross, participants adjusted the size of a red ellipse on the x and y axes using the computer mouse, so that all barbs (Honeycomb) or dots (Extinction) inside the ellipse were perceptible to them.

The red ellipse was displayed with a random size (within the screen size) at the beginning of each trial. Both illusions were tested with two contrast polarity conditions, i.e., either black or white barbs in the Honeycomb illusion and either black or white dots in the Extinction illusion. Hence, there were four conditions (HcBlack, HcWhite, ExBlack, ExWhite) and each condition was tested twice in a random order. There was no time limit for the adjustment. Random light and dark grey checkerboards (40 random masks presented for 0.5 sec each and made of squares of 0.52 arcdeg in side with 0.35 and 0.65 of the maximum luminance) were shown after each trial to reduce the aftereffect. The extracted value was the area of the adjusted ellipse.

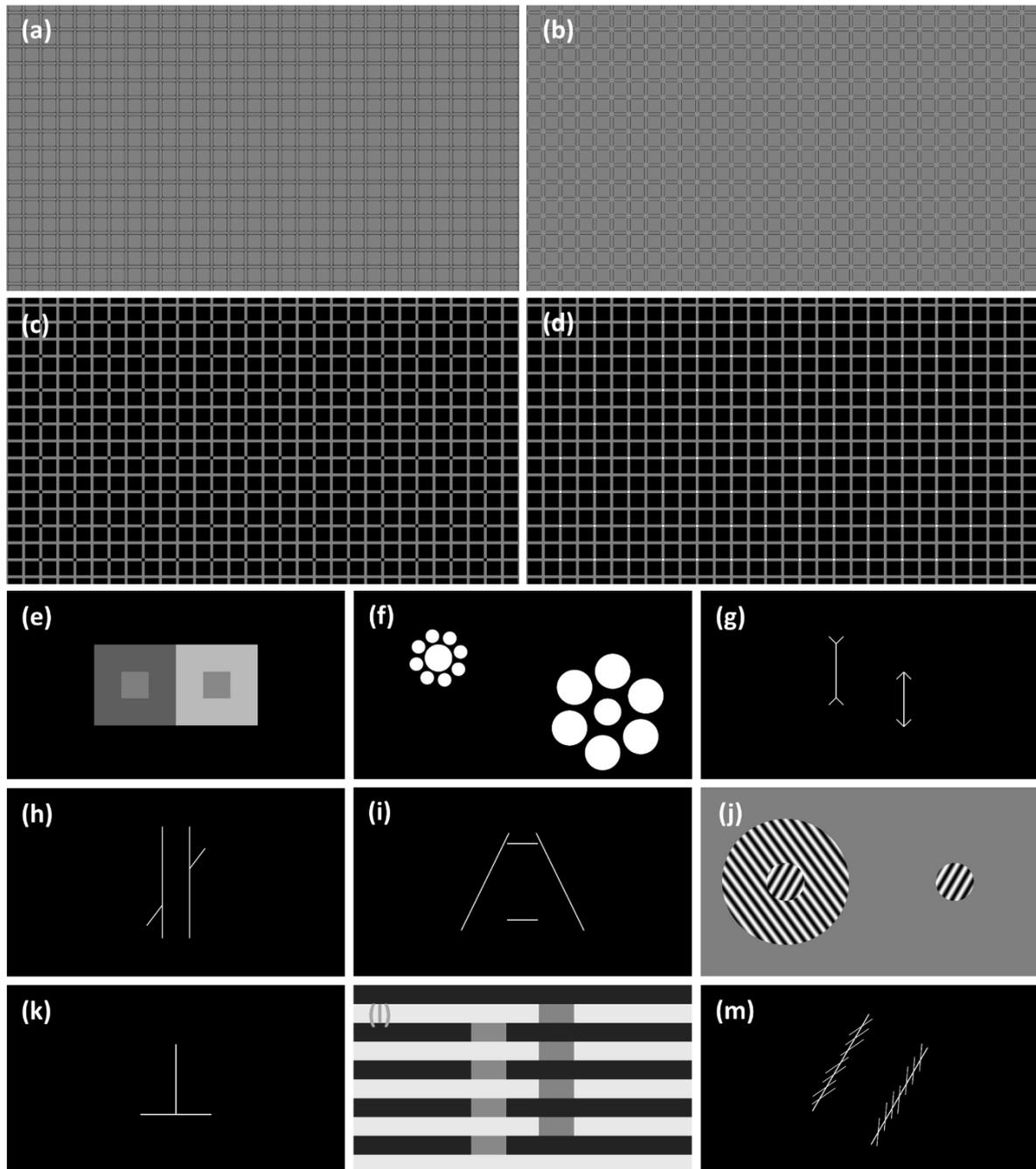


Figure 3. The Honeycomb illusion with (a) black and (b) white barbs and the Extinction illusion with (c) black and (d) white dots. The red adjustable ellipse is not depicted here. The image (a) to (d) need to be enlarged so as to fill a large proportion of the visual field; for details, see Bertamini et al. (2016). The battery of other illusions: (e) contrast, (f) Ebbinghaus, (g) Müller-Lyer, (h) Poggendorff, (i) Ponzo, (j) Tilt, (k) vertical-horizontal, (l) White, and (m) Zöllner. Illusions (e) to (m) were all tested with two configurations. For example, the upper horizontal line of the Ponzo illusion was adjusted to match the length of the lower horizontal line, or inversely.

2.4.4. Illusions

A battery of nine other illusions was tested (Figure 3e-m): contrast (CS), Ebbinghaus (EB), Müller-Lyer (ML), Poggendorff (PD), Ponzo (PZ), Tilt (TT), vertical-horizontal (VH), White (WH) and Zöllner (ZN). A method of adjustment was used to measure illusion susceptibility, i.e., participants were asked to adjust the size (EB, ML, PZ, VH), shade of grey (CS, WH), orientation (TT, ZN), or position (PD) of an element to match the size, shade of grey, orientation, or position, respectively, of a reference on the screen by moving the computer mouse. The reference and adjustable elements were the inside squares in the CS illusion, the central disks in the EB illusion, the vertical segments with inward- and outward-pointing arrows in the ML illusion, the left and right parts of the interrupted diagonal in the PD illusion, the upper and lower horizontal segments in the PZ illusion, the small left and right Gabor patches in the TT illusion, the horizontal and vertical segments in the VH illusion, the two columns of rectangles in the WH illusion, and the two main streams in the ZN illusion. Each illusion was tested in two conditions: one element (or series of elements, in the case of the White illusion) was in turn the reference or the adjustable element. For example, in the Ebbinghaus illusion, the task was either to adjust the size of the left central disk so that it appeared to be the same size as the right central disk or to adjust the size of the right central disk so that it appeared to be the same size as the left central disk. The order of presentation of the different illusions and conditions was randomized across participants and there was no time constraint. For a detailed description of the illusions, refer to Cretenoud et al. (2019) and Grzeczowski et al. (2017). The extracted values were the illusion magnitudes expressed as a difference compared to the reference. Positive and negative illusion magnitudes indicate over- and under-adjustments, respectively.

2.4.5. N-back

We tested a 1-back paradigm based on a bisection stimulus, which consists in three vertical lines with the central line being either offset to the left or right compared to the veridical center. The vertical lines were 1200 arcsec in length and the offset was fixed at 100 arcsec. Each trial consisted in a bisection stimulus, which was shown for 150 ms. Participants were asked to report whether the offset of the current stimulus was on the same or opposite side compared to the offset of the previous stimulus (1-back; Figure 2c) using two push buttons. Forty-one bisection stimuli were shown. We extracted the percentage of correct responses.

2.4.6. Orientation discrimination

Participants performed an adapted version of the orientation discrimination paradigm used in Tibber, Guedes, and Shepherd (2006). Each trial consisted in a red central dot followed by a Gabor patch

(spatial frequency: 3.3 cy/arcdeg; duration of presentation: 100 ms; envelope sigma along orientation: 0.57 arcdeg; envelope sigma perpendicular to orientation: 0.19 arcdeg), which was centrally displayed (Figure 2d). Gabors were rendered using dithering in order to virtually increase gray level resolution. The mean luminance was 50% and the target contrast was 80%. Participants were asked to discriminate between clockwise and counterclockwise stimuli by using two push buttons. The Gabor orientation at which participants gave 75% of correct responses was estimated using a staircase PEST procedure (starting value: 5°).

2.4.7. Random dot kinematograms

The random dot kinematograms paradigm measures global motion perception (Edwards & Badcock, 1995; Hutchinson & Stocks, 2013; Newsome & Park, 1988). Two thousand dots were moving at 5 arcsec/s in a circular aperture (inner diameter: 1 arcdeg; outer diameter: 12 arcdeg) for 500 ms. Each trial consisted of a certain ratio of dots moving coherently while the rest of the dots moved independently from each other (i.e., distractors; Figure 2e). Participants had to discriminate between leftward and rightward (horizontal, RDKh) or inward and outward (radial, RDKr) global motion. The ratio of dots moving coherently was adapted using a staircase procedure (QUEST with the prior for coherence centered at 60% with SD 50%; Watson & Pelli, 1979; Watson & Pelli, 1983) to reach 75% of correct responses. The two conditions were tested sequentially, and the order was randomized across participants.

2.4.8. Simple reaction times

We used a modified version of the classic Hick-paradigm (Hick, 1952). Participants were instructed to press a mouse button as quickly as possible after a white square (3 arcdeg in side) appeared on a black background. To prevent participants from predicting when the white square appeared, the inter-trial interval (ITI) varied randomly (minimum: 1500 ms; maximum: 3500 ms). The extracted value was the median reaction time (outlier trials were removed using a modified z-score; Iglewicz & Hoaglin, 1993).

2.4.9. Pro- and anti-saccades

Participants gazed at a fixation dot in the center of the screen and were asked to make a pro- or an anti-saccade towards or away from a target, respectively. The color of the fixation dot, i.e., green or red, indicated whether a pro- or anti-saccade was required, respectively. The target was randomly displayed to the left or to the right of the fixation dot. A positive or negative feedback was provided at the end of each trial as a happy or sad smiley, respectively. Participants were positioned in the head rest of an SMI iViewXHi-Speed 1250 eye tracker (Sensomotoric Instruments, Teltow, Germany) and eye

movements were recorded binocularly at 500 Hz. For both pro- and anti-saccades, we extracted the median travel time (i.e., saccade duration; proTravel and antiTravel) and median saccade time (i.e., delay between the target onset and the saccade onset; proSac and antiSac). As in the simple reaction times paradigm, a modified z-score was used to detect and remove outlier trials.

2.4.10. Freiburg visual acuity

Visual acuity was measured following the procedure of the Freiburg visual acuity test (Bach, 1996). Participants were presented with Landolt-C optotypes (Figure 2f) with randomized gap orientations and were asked to indicate the direction of the gap (“up”, “up-right”, “right”, “down-right”, “down”, “down-left”, “left”, or “up-left”) using an 8-button controller. The size of the optotype was varied according to a staircase QUEST procedure and we extracted the size corresponding to 75% of correct responses.

2.4.11. Visual backward masking

In a visual backward masking paradigm (Herzog et al., 2004; Herzog & Koch, 2001; Roinishvili et al., 2011), a vernier stimulus, which consists of two vertical bars offset in the horizontal direction, was presented for 10 ms. The offset between the two horizontal bars was fixed at 75 arcsec. The vernier stimulus was followed by a variable inter-stimulus interval (ISI), i.e., a blank screen, and by a grating for 300 ms (Figure 2h). The grating consisted of 5 aligned elements of the same length as the vernier stimulus. Participants were asked to report the offset direction of the lower bar in the vernier stimulus. The ISI was varied using a PEST procedure (starting value: 190 ms) so that participants reached 75% of correct responses.

Visual search

In the visual search paradigm, four (VS4) or 16 (VS16) lines were presented randomly within a black square. Using two push buttons, participants had to report as quickly as possible whether a green horizontal line was present within an array of distractors (green vertical, red vertical and horizontal lines; Figure 2g). The green horizontal line, i.e., the target, was present in 50% of the trials. The median reaction time was extracted for correct trials in both conditions (after outlier trials were excluded according to modified z-scores).

2.5. **Pre-processing and data analysis**

Data were extracted in Matlab (The Mathworks Inc.) and analyses were performed in R (R Core Team, 2018), except when mentioned.

2.5.1. Test-retest reliability

We computed test-retest reliability for the variables extracted from visual paradigms, which were tested twice, i.e., the Honeycomb and Extinction illusions, the battery of other illusions, visual acuity, and visual backward masking. As suggested by Koo and Li (2016) and by Shrout and Fleiss (1979), two-way mixed effects models (intraclass correlations of type (3,1) or $ICC_{3,1}$) were computed. The test-retest reliabilities were significant even after Bonferroni correction was applied for multiple comparisons for all but one variable (WH left: ICC coef. = 0.110; 95% CI [-0.093, 0.305], $p = 0.144$; see Table S2 in the Supplementary File). Hence, we averaged the extracted values across repeated trials.

2.5.2. Illusions

Illusion magnitudes are shown in Figure S2 (Supplementary File). Correlations were computed between the magnitudes of the two conditions of an illusion (CS: $r = -0.500$, $p < 0.001$; EB: $r = -0.479$, $p < 0.001$; ML: $r = -0.191$, $p = 0.065$; PD: $r = 0.721$, $p < 0.001$; PZ: $r = -0.154$, $p = 0.139$; TT: $r = -0.122$, $p = 0.243$; VH: $r = -0.421$, $p < 0.001$; WH: $r = -0.217$, $p = 0.034$; ZN: $r = -0.712$, $p < 0.001$). Even though between-condition correlations were not always strong according to Gignac and Szodorai (2016), the two conditions of each illusion from the Ill paradigm were combined into a global illusion magnitude, which was expressed as the sum of the absolute effects in the two conditions.

Bertamini, Cretenoud, and Herzog (2019) recently observed a dissociation between the Honeycomb and Extinction illusions depending on contrast polarity, suggesting that different mechanisms are operating in the black and white conditions of both illusions, respectively. Here, we computed a repeated measures ANOVA and similarly observed a significant interaction ($F[1,93] = 118.7$, $p < 0.001$; see Figure S3 in the Supplementary File) between the illusion type (Honeycomb or Extinction illusion) and contrast polarity (black or white). Therefore, the two conditions of the Honeycomb and Extinction illusions were not combined into a global illusion magnitude.

2.5.3. Data transformation and outlier removal

The normality assumption was tested by computing a Shapiro-Wilk test for each variable. Some distributions violated the normality assumption (see Table S3 in the Supplementary File). Hence, each distribution was rescaled to approximate a normal distribution. First, we shifted the data distribution to positive values only. Second, we computed modified z-scores, which are based on the median and median absolute deviation rather than the mean and standard deviation, respectively, and removed outliers according to a 3.5 criterion (Iglewicz & Hoaglin, 1993). Third, we optimized the λ exponent of a Tukey power transformation (see Table S3 in the Supplementary File) to maximize normality

according to the Shapiro-Wilk test. Fourth, including the previously removed outliers, data were transformed using the Tukey transformation with the optimized λ parameter. Fifth, we standardized the data, as above. Outliers were removed only in the visual variables. Last, we flipped the sign of visual variables when lower values indicated better performance (Crw, Ctr, Ill, Ori, RDK, RT, Sac, VBM, and VS paradigms). Higher values indicate better performance in all gaming variables.

2.5.4. Questionnaires

To reduce the complexity of our dataset, the three subscales of the BIS (NI: nonplanning impulsivity, MI: motor impulsivity; AI: attentional impulsivity), which showed strong correlations with each other (NI-MI: $r = 0.441$, $p < 0.001$; NI-AI: $r = 0.406$, $p < 0.001$; MI-AI: $r = 0.532$, $p < 0.001$), were summed in a total score, which was considered for further analysis. Similarly, we summed the two subscales of the CI (EC: enjoyment of competition; CO: contentiousness) for further analysis, since they significantly correlated ($r = 0.314$, $p = 0.002$). Similarly, we later only considered the total score (i.e., we summed the subscales) of the AQ (SS: social skills; AS: attention switch; AD: attention to detail; CO: communication; IM: imagination) and O-LIFE (UE: unusual experiences; IA: introverted anhedonia; CD: cognitive disorganization; IN: impulsive nonconformity) questionnaires. However, the HEXACO subscales were kept as separate variables because they showed weak inter-correlations (see Table 2).

3. Results

3.1. Correlations

Pairwise correlations were computed for all variables extracted (Table 2). Correlations were in general weak between visual variables, except between pairs of variables that were extracted from the same paradigm, e.g., between the two conditions of the visual search paradigm (VS4-VS16: $r = 0.790$, $p < 0.001$). However, in line with previous studies (e.g., Cretenoud et al., 2019; Grzechkowski et al., 2017, 2018), between-illusion correlations were rather weak ($M_r = 0.108$; $SD_r = 0.102$), except between the Honeycomb and Extinction illusions (all $ps < 0.001$; HcBlack-HcWhite: $r = 0.819$; HcBlack-ExBlack: $r = 0.702$; HcBlack-ExWhite: $r = 0.501$; HcWhite-ExBlack: $r = 0.767$; HcWhite-ExWhite: $r = 0.527$; ExBlack-ExWhite: $r = 0.575$). Gaming variables showed stronger inter-correlations ($M_r = 0.291$; $SD_r = 0.121$). In addition, correlations between questionnaire variables, which are related to CS:GO (NbHourPerWeek, NbTotalHours, actual CS:GO rank, best CS:GO rank), and gaming variables were high too ($M_r = 0.314$; $SD_r = 0.182$). The correlations between CS:GO related questionnaire variables were even higher ($M_r = 0.586$; $SD_r = 0.256$).

3.2. Exploratory Factor Analysis

We imputed outlying and missing values using the 'mice' function from the mice R package with method 'norm' (Bayesian linear regression with 20 imputation samples). We then computed the Kaiser-Meyer-Olkin (KMO) test for sampling adequacy, which quantifies the degree of inter-variable correlations. Variables that showed an unacceptable measure of sampling adequacy (i.e., MSA < 0.5) were removed sequentially until all variables showed an acceptable MSA. PD, TT, CrwPeri, antiTravel, RDKh, CS, CrwSize, WH, VH, ML, Ori, HEXACO HH, RDKr, EB, BIS, and O-LIFE were therefore removed for the exploratory factor analysis (EFA). The global MSA index after variable removal was 0.695.

Factors were extracted with a common factor analysis to reflect the variance shared between variables (i.e., the common variance). We used an oblique rotation (promax; see Costello & Osborne, 2005) because we had no reason to preclude factors to correlate.

A parallel analysis suggested a seven-factor model, whereas only three factors were suggested by scree plot inspection (see Figure S4 in the Supplementary File). Because the eigenvalues for factors 4 to 7 were very close to those of a resampled dataset (RF1: 4.802; RF2: 2.762; RF3: 2.093; RF4: 1.230; RF5: 1.110; RF6: 0.871; RF7: 0.815), we retained the three-factor model (TLI = 0.704; RMSEA = 0.064 with 90% CI [0.053, 0.076]). The three factors together explained 29.5% of the variance (RF1: 12.3%; RF2: 9.5%; RF3: 7.7%). Loadings are reported in Table 3. According to a simulation published in Hair et al. (2018), loadings larger than 0.55 are considered as significant with a sample size of 100.

The first factor was mainly composed of the CS:GO related and gaming variables. The second factor was mainly related to the Honeycomb and Extinction variables. The third factor mainly loaded on the visual variables related to reaction times, such as RT, VS4, and VS16, and to the pro- and anti-saccade paradigm. Inter-factor correlations were weak (RF1-RF2: $r = 0.093$; RF1-RF3: $r = 0.011$; RF2-RF3: 0.006).

Table 3. Rotated factor loadings from an EFA after promax rotation for visual (green), gaming (orange), and questionnaire (purple) variables. A color scale from blue (negative loadings) to red (positive loadings) is shown. Factor loadings larger than 0.55 or smaller than -0.55 are highlighted (bold).

	RF1	RF2	RF3
Ctr	-0.110	-0.226	0.275
HcBlack	0.011	0.798	0.176
HcWhite	0.070	0.836	0.159
ExBlack	-0.082	0.751	0.251
ExWhite	0.092	0.580	0.182
PZ	-0.301	0.115	0.094
ZN	-0.243	-0.024	0.276
NB	-0.161	-0.093	0.365
RT	0.175	0.049	0.538
proTravel	0.111	-0.038	0.352
proSac	0.174	0.060	0.543
antiSac	0.111	-0.008	0.569
VA	-0.089	-0.200	0.237
VBM	-0.055	-0.267	0.525
VS4	0.060	0.168	0.537
VS16	-0.022	0.272	0.489
Shoot	0.701	0.110	0.108
Spray	0.500	0.043	-0.011
Track	0.596	-0.076	0.133
Hold	0.212	0.020	0.289
Flick	0.407	-0.109	0.022
Peek	0.220	0.069	0.146
AQ	-0.003	-0.263	-0.253
CI	-0.183	0.454	-0.060
HEXACO EM	-0.073	-0.405	0.056
HEXACO EX	-0.095	0.129	0.216
HEXACO AG	0.166	-0.143	0.057
HEXACO CO	-0.087	0.277	-0.128
HEXACO OP	-0.270	0.018	-0.023
Handedness	0.142	0.271	-0.148
PRFd	-0.333	0.186	0.095
NbHoursPerWeek	0.405	0.151	-0.072
NbTotalHours	0.806	0.140	0.054
Actual CS:GO rank	0.848	0.131	0.122
Best CS:GO rank	0.876	0.105	0.083

3.3. Regression

3.3.1. Path model

We aimed at determining to which extent the visual, gaming, and questionnaire scores predict the actual CS:GO ranks of the players. In addition, we wondered how gaming variables are related to visual variables. Hence, we designed a complex, multiple regression model (i.e., a path model), which is schematically represented in Figure 4. The actual CS:GO rank is an outcome variable (i.e., endogenous variable), while the visual and questionnaire variables are predictors (i.e., exogenous variables). The gaming variables are both outcomes and predictors. Note that the CS:GO related questionnaire variables (i.e., best CS:GO rank, NbHoursPerWeek, and NbTotalHours) were not included for further analysis.

Standardized path coefficients are reported in Table 4. The visual, gaming, and questionnaire variables explained 72.3% of the variance of the CS:GO ranks. Between 28.2% and 40.4% of the variance of each gaming variable was accounted for by the visual variables. Not only the gaming variables showed significant standardized path coefficients on the CS:GO ranks but also variables related to visual paradigms, such as crowding, the Honeycomb and Extinction illusions, and visual search. Similarly, visual variables showed some strong standardized path coefficients on the gaming variables. For example, the CrwPeri variable significantly loaded on the Shoot, Spray, Hold and Flick gaming variables (standardized path coefficients of 0.500, 0.354, 0.334 and 0.503, respectively).

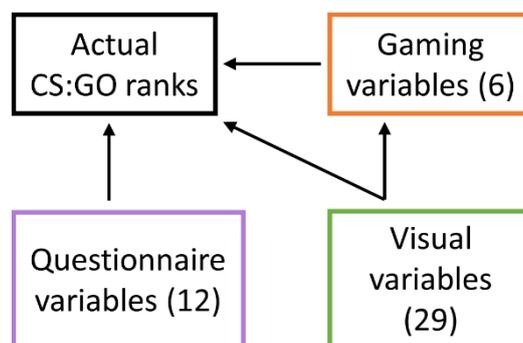


Figure 4. Schematic representation of the path model computed to determine to what extent the players' rank can be predicted by the visual, gaming, and questionnaire variables. The numbers in brackets indicate the number of variables considered. The visual variables not only regressed on the actual CS:GO ranks but also on the gaming variables.

Table 4. Standardized path coefficients (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$) from the path model (see Figure 4) and variance explained (r^2) of each endogenous variable. The strength of the standardized path coefficients is indicated with a color scale from blue (negative loadings) to red (positive loadings).

	Actual CS:GO rank	Shoot	Spray	Track	Hold	Flick	Peek
CrwSize	0.231**	-0.030	-0.058	-0.197	0.152	-0.101	0.065
CrwPeri	-0.420**	0.500**	0.354*	0.112	0.334*	0.503**	0.249
Ctr	0.108	-0.069	-0.163	-0.139	-0.063	-0.082	0.010
HcBlack	-0.205	0.350*	0.228	0.013	0.109	0.257	0.160
HcWhite	0.513***	-0.056	0.123	0.093	-0.156	-0.270	-0.207
ExBlack	-0.489**	0.102	-0.102	-0.263	0.431*	-0.110	0.190
ExWhite	0.128	-0.267	-0.027	0.279	-0.364*	0.181	-0.025
CS	0.069	0.005	-0.060	-0.113	0.157	-0.055	-0.217*
EB	-0.082	0.317**	0.129	0.148	0.254*	0.201	-0.015
ML	0.085	-0.273*	-0.119	-0.131	-0.022	-0.054	0.241*
PD	0.274**	-0.439***	-0.114	-0.028	-0.237	-0.139	-0.102
PZ	-0.018	-0.047	0.074	-0.008	0.020	0.096	-0.283**
TT	0.022	-0.014	-0.200	0.022	0.035	-0.088	0.067
VH	-0.021	0.006	0.275**	-0.024	0.003	0.180	-0.160
WH	-0.095	-0.103	-0.103	-0.111	-0.046	-0.151	-0.034
ZN	-0.130	-0.136	0.005	0.109	-0.240*	0.105	-0.057
NB	-0.09	-0.055	-0.047	-0.072	0.042	-0.038	0.142
Ori	-0.145	0.155	0.430***	0.088	0.110	0.187	-0.033
RDKh	0.081	0.115	0.326**	0.364**	0.076	0.279**	0.063
RDKr	-0.173*	0.101	0.052	-0.146	0.152	0.085	0.158
RT	-0.166	0.226	0.023	0.147	0.355**	-0.010	-0.152
proTravel	-0.003	-0.055	0.148	-0.081	0.067	-0.127	0.401*
proSac	0.124	0.378*	0.327*	0.088	0.134	0.186	-0.031
antiTravel	-0.028	0.117	0.029	0.190	-0.026	0.254	-0.415*
antiSac	-0.109	-0.217	-0.201	-0.101	-0.068	0.030	0.118
VA	-0.103	0.166	-0.152	0.163	0.034	-0.388**	-0.038
VBM	0.256*	-0.232	0.030	0.019	-0.039	0.040	0.000
VS4	0.433**	-0.181	-0.305	0.220	-0.141	-0.137	-0.410*
VS16	-0.049	0.121	-0.021	-0.196	-0.038	-0.075	0.298
Shoot	0.497***						
Spray	0.179**						
Track	0.195**						
Hold	-0.039						
Flick	0.056						
Peek	-0.128*						
AQ	0.081						
BIS	0.197						
CI	0.221*						
HEXACO HH	0.270**						
HEXACO EM	-0.069						
HEXACO EX	0.018						
HEXACO AG	0.178*						
HEXACO CO	-0.101						
HEXACO OP	-0.021						
Handedness	0.148						
O-LIFE	-0.093						
PRFd	-0.004						
r^2	0.723	0.369	0.350	0.369	0.317	0.404	0.282

3.3.2. Elastic net model

Using the scikit-learn package in Python (Pedregosa et al., 2011), we aimed at predicting the actual CS:GO players' ranks by fitting an elastic net model (Zou & Hastie, 2005), i.e., a regressor, which both uses L1 and L2 regularizations, therefore reducing the risks of overfitting and the dimensionality of the model.

The dataset was split into a training (80%) and test (20%) set. Using a search grid with a 5-fold cross-validation, we optimized the model's generalization performance on the training set by tuning two hyperparameters, namely alpha and the L1 ratio (i.e., $L1/(L1+L2)$). The lower alpha, the more complex the model (i.e., less strict regularization).

Performance on the training set was optimized for alpha = 0.15 and with an L1 ratio of 0.45. With these values for the hyperparameters, the training and test set accuracies were $r^2 = 0.640$ and $r^2 = 0.263$, respectively. The *MSEs* for the training and test sets were 0.21 and 0.25, respectively. In contrast, a dummy regressor resulted in an *MSE* of 0.60 and 0.37 in the training and test sets, respectively. The variables, which showed non-zero coefficients, were CrwSize (0.031), HcWhite (0.085), ZN (-0.110), proSac (0.087), VA (-0.038), VS4 (0.007), Shoot (0.225), Spray (0.059), Track (0.169), Flick (0.002), BIS (0.045), HEXACO CO (-0.011), and HEXACO OP (-0.099). Gaming variables were expected to show non-zero coefficients, because they are obviously related to the players' ranks (see Table 2).

Our results suggest that the Honeycomb illusion and crowding variables are predictors of the players' ranks, i.e., participants who perceived barbs in larger areas (HcWhite) and who needed a smaller optotype to achieve 75% of performance (CrwSize), tend to have higher ranks. Note that both paradigms are related to visual perception in the periphery. However, participants with higher ranks tend to have worse visual acuity in the fovea (VA) and to be more susceptible to the Zöllner illusion (ZN). In addition, our results suggest that faster reaction times (proSac and VS4) are associated with higher ranks. Lastly, participants with weaker conscientiousness (HEXACO CO), weaker openness to experience (HEXACO OP), and with higher score on the Barratt Impulsiveness Scale (BIS), tend to have higher ranks.

4. Discussion

4.1. Summary

We tested 94 CS:GO players ranging from beginners to experts with 12 visual paradigms, specific gaming skills, and personality traits to examine what aspects are associated with expertise, and whether there is a unique, common factor underlying visual perception in AVGPs.

First, we observed only weak correlations between visual variables, except between variables that belong to the same paradigm, which can be taken as a measure of good test-retest reliability (paradigms that were tested more than once mostly showed significant intraclass correlations, also suggesting reliable measurements). In addition, gaming variables showed strong inter-correlations. A factor analysis revealed three factors explaining about 30% of the variance, which suggests a poor factor structure. Second, a path model showed that more than 70% of the variance of the actual players' rank is predicted by visual, gaming, and questionnaire scores. Not only gaming variables but also some visual and questionnaire scores showed strong loadings on the actual players' rank. Lastly, we computed an elastic net model to select the features with stronger predictive power on the actual ranks (i.e., to reduce the dimensionality of the dataset). The model retained 13 variables (among which six were visual variables), which altogether led to better predictions of the ranks compared to a dummy model. The visual variables, which were retained in the elastic net model and showed significant standardized loadings in the path model, were *CrwSize* (crowding size), *HcWhite* (Honeycomb illusion with white barbs), and *VS4* (visual search with 4 elements). Note that the best CS:GO rank, *NbHourPerWeek*, and *NbTotalHours* were not included in the path and elastic net models.

Importantly, the path model accounted for a larger proportion of the variance in the data compared to the elastic net model ($r^2 = 0.723$ versus $r^2_{training} = 0.640$, respectively), because the former used more variables than the latter (47 versus 13 variables, respectively). While the dimensionality of the dataset was much reduced in the elastic net model, the decrease in performance compared to the path model was rather small, which suggests that most variables do not significantly predict the players' rank. However, the test set accuracy was much lower than the training set accuracy in the elastic net model ($r^2_{test} = 0.263$ versus $r^2_{training} = 0.640$), which suggests overfitting even though the elastic net model showed a better test *MSE* compared to a dummy regressor. The small test sample size (20%, i.e., 19 participants only) may partially explain the rather low test set accuracy.

We expected many aspects of gaming to rely on (low-level) visual skills. However, our results suggest that there is no strong common factor for visual perception in CS:GO players. Similarly, there is only

weak evidence for a common factor for visual perception in general (Mollon et al., 2017; Tulver, 2019). For example, many specific factors were reported in oculomotor tasks (Bargary et al., 2017), in the perception of faces (Verhallen et al., 2017), and visual illusions (e.g., Cretenoud et al., 2019; Grzeczowski et al., 2017). More generally, basic visual paradigms only weakly correlated with each other (e.g., Cappe, Clarke, Mohr, & Herzog, 2014).

4.2. **Positive association between peripheral vision and the players' rank**

Rather than a strong common factor for visual perception in CS:GO players, specific visual paradigms and personality traits seem to be strongly predictive of the players' rank. For example, players who perceived more barbs in the Honeycomb white illusion, tended to have higher ranks. Since the four variables extracted from the Honeycomb and Extinction illusion paradigm (HcBlack, HcWhite, ExBlack, and ExWhite) strongly correlated with each other (Table 2), we expected that either all four variables or none would be significantly associated with the players' rank. However, only one variable (i.e., HcWhite) showed a non-zero coefficient in the elastic net model and two out of four only showed significant standardized path loadings (HcWhite and ExBlack), suggesting that not all variants of the illusions are associated with the players' rank.

Interestingly, similar illusion magnitudes were observed in the Honeycomb illusion with black and white barbs (HcBlack and HcWhite; see Figure S3 in the Supplementary File), while the mean extent of visible region was previously shown to be larger in the white compared to the black variant (Bertamini et al., 2019). A difference in the experimental design may explain the discrepancies in the results. To estimate the mean extent of the region in which barbs were visible, participants adjusted the size of an ellipse (i.e., on both x and y axes) in the present investigation, while a disk (i.e., a single dimension) was adjusted in Bertamini et al. (2019). The background images were the same in both studies. Note that barbs (or disks in the Extinction illusion) were removed during the adjustment in Bertamini et al. (2019), unlike in the present investigation. Despite these differences, an interaction between the illusion type (Extinction, Honeycomb) and contrast polarity (black, white) was observed in both studies. It may be worth considering that the magnitude of these illusions conflate a perceptual and a response bias aspect. Participants may differ in the tendency to report what they may "know" rather than what they see, or, even without awareness of this, to "cheat" by not maintaining fixation.

The crowding paradigm similarly seems to be a strong predictor of the players' rank. Higher ranks were associated with a better visual acuity in the periphery, as reflected by the CrwSize variable. While Green and Bavelier (2007b) previously reported an increased spatial resolution in AVGPs compared to NVGPs, we observed a negative association between spatial resolution, as measured with a crowding

paradigm (CrwPeri), and the players' rank. However, further investigation is needed to verify this association, since the CrwPeri did not show up as a non-zero coefficient in the elastic net model, which may indicate that the association is only reliable in combination with other specific skills or unreliable. In addition, a radial-tangential anisotropy was reported in crowding (Chung, 2013; Greenwood et al., 2017), suggesting that the association may be different along the horizontal axis.

Both the Hc/Ex and Crw paradigms are related to peripheral vision and were strongly associated with the players' rank. However, the Hc/Ex and Crw variables only weakly correlated ($M = 0.095$, $SD = 0.124$; see Table 2). As in foveal vision, it is likely that vision in the periphery is multifactorial, i.e., there is no strong common factor for peripheral vision. For example, Yashar, Wu, Chen, and Carrasco (2019) reported no common mechanism for crowding across different visual features. The authors tested different visual features under crowding to determine at which processing stage crowding occurs. They observed that orientation and spatial frequency errors were interdependent, while orientation and color errors were independent, suggesting that peripheral vision is feature-dependent.

While different features are processed differently in the periphery, our results suggest that peripheral vision in general plays an important role in CS:GO. Specifically, it seems that peripheral vision is enhanced in high-rank compared to low-rank players, which adds to previous results reporting evidence for enhanced performance in peripheral vision in AVGPs compared to NVGPs (for a review and meta-analysis, see Chopin et al., 2019). Similarly, enhanced peripheral visual skills are beneficial to team sports players, such as basketball or soccer players (Faubert & Sidebottom, 2012; Knudson & Kluka, 1997). However, note that the CrwSize variable only weakly correlated with the players' rank (Table 2), suggesting that its role in CS:GO is important when interacting with other specific skills only.

4.3. Negative association between central vision and the players' rank

Surprisingly, visual acuity in the fovea (VA) was negatively associated with the players' rank. Patino and colleagues (2010) reported that central and peripheral visual acuities were negatively correlated in a large sample of subjects. Here, we did not observe any significant correlation between central (VA) and peripheral (CrwSize) visual acuities ($r = 0.159$, $p = 0.129$; Table 2). Our results therefore suggest that central and peripheral vision are independently engaged while playing video games, as was shown for reaching (Prado et al., 2005).

4.4. Other associations between visual paradigms and the players' rank

Faster reaction times (when making saccades or searching for a target, i.e., proSac and VS4, respectively) were also associated with higher ranks. Similarly, Bosten and colleagues (2017) reported

that the time spent playing computer games significantly correlated with a factor for oculomotor speed.

In addition, we observed other associations, which however explained only a small proportion of the variance of the ranks and were not always consistent across analyses (path model vs. elastic net model), suggesting that they may not be reliable. Hence, reduced backward masking was in general observed in players with higher ranks (Table 4). Li and colleagues (2010) previously reported that playing action game reduces the effects of backward masking.

Likewise, VGPs were reported to perform significantly better than NVGPs at discriminating contracting, but not expanding, elements in a radial random dot kinematograms paradigm (Hutchinson & Stocks, 2013). The authors suggested that VGPs are more sensitive than NVGPs to visual characteristics, which are enhanced in gaming (e.g., contracting patterns) relative to those encountered in the real world (e.g., expanding patterns). Here, however, players with higher ranks showed in general worse performance in the radial random dot kinematograms paradigm (RDKr; Table 4). Note that both contracting and expanding conditions were considered together in the RDKr variable.

Previous studies suggested that action video games enhance certain type of perceptual features. Some of these associations were reproduced here (e.g., VBM), while others were not. For example, Li and colleagues (2009) reported that contrast sensitivity at intermediate and higher spatial frequencies was enhanced after action video game training, which suggests that high-rank players may have better sensitivity to contrast than low-rank players. However, we did not observe any significant association between contrast sensitivity (Ctr) and the players' rank. It seems unlikely that the spatial frequency used here (4.0 cy/arcdeg) was too low to find an effect, since a small but significant effect was previously reported with a spatial frequency of 3.0 cycles per degree (Li et al., 2009). Similarly, we did not observe any significant association between the players' rank and perceptual speed (i.e., RT variable), contrary to Dye and colleagues (2009). Importantly, we tested only gamers ranging from beginners to experts but not non-gamers, as used in most studies before, which may explain the discrepancies in the results. For example, it may be that training video game improves contrast detection and perceptual speed in NVGs but does not further improve with additional training.

4.5. Gaming variables

Among the six gaming variables that were extracted, four were retained in the elastic net model (Shoot, Spray, Track, and Flick) and two significantly loaded on the first factor in the EFA (Shoot and Track). To estimate to which extent the players' rank can be predicted from the performances in the six gaming variables, we computed another path model, in which the gaming variables only loaded on

the actual CS:GO ranks. The six gaming variables accounted for 48% of the variance of the actual CS:GO ranks. Note that extracting more gaming variables could have resulted in a larger proportion of the variance explained. However, our results highlight that the Shoot, Spray, and Track variables (which showed up in both the path and elastic net models) are building blocks for the game.

4.6. Questionnaire variables

The associations between CS:GO unrelated questionnaire variables and the gaming variables or the players' rank were weak and inconsistent across analyses. Indeed, three questionnaire variables showed weak but significant loadings on the players' rank in the path model, namely the CI, HEXACO HH, and HEXACO AG, indicating that high scores in competitiveness, honesty-humility, and agreeableness were in general associated with higher ranks. In contrast, the elastic net model suggested that players with high impulsivity scores (BIS) and low scores in conscientiousness (HEXACO CO) and openness to experience (HEXACO OP) tended to have higher ranks.

4.7. Limitations

While we only tested CS:GO players, our results may hold true for other action video games. Importantly, the present investigation does not allow us to claim that high-rank CS:GO players develop specific visual skills, such as better visual acuity in the periphery while playing. Neither can we infer from the data that specific visual skills or personality traits are required to become an excellent player. However, a reliable dose-response effect in intervention studies was suggested as evidence for a causal effect of action video gaming on perception (Chopin et al., 2019). While we are not able to infer causality, our experimental design avoids the methodological shortcomings usually related to intervention studies (Boot et al., 2011), such as differential placebo effects driven by the treatment versus control interventions (e.g., Tetris-trained participants may predict that they will have a better post-training performance in a mental rotation task). Likewise, all participants were active gamers, reducing the risks of strategy changes impacting our results. We cannot exclude gender-specific effects since only male participants took part in the present study (e.g., gender disparity in mental rotation ability decreases following video game training; see Feng et al., 2007). Last, we considered the rank as a continuous variable, even though it is ordinal. As the distance between two ranks may not be constant, we consider this as a limitation of the present investigation.

4.8. Conclusions

To summarize, our results suggest that there is no unique, strong common factor for visual perception in CS:GO players. However, some visual tasks strongly predict the players' rank. In particular, visual

perception in the periphery, as measured with a crowding paradigm and the Honeycomb illusion, seems to be enhanced in high-rank players. Even though causative relationships cannot be derived from these results, the present investigation gives clues about visual paradigms, which may be part of future training programs for esports.

5. Acknowledgments

We thank Marc Reknow for technical support and Ophélie Favrod for helping with the selection of the visual paradigms. This work was supported by the project “Basics of visual processing: from elements to figures” (project no. 320030_176153/1) of the Swiss National Science Foundation (SNSF) and by a National Centre of Competence in Research (NCCR Synapsy) grant from the SNSF (51NF40-185897).

6. Declaration of interests

AB, AM, and CC were employed at the company Logitech Europe S.A., a manufacturer of computer peripherals and software, in the R&D department.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7. References

- Aberg, K. C., & Herzog, M. H. (2009). Interleaving bisection stimuli - randomly or in sequence - does not disrupt perceptual learning, it just makes it more difficult. *Vision Research*, *49*(21), 2591–2598. <https://doi.org/10.1016/j.visres.2009.07.006>
- Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, *91*, 340–345.
- Bach, M. (1996). The Freiburg Visual Acuity Test-Automatic Measurement of Visual Acuity. *Optometry and Vision Science*, *73*(1), 49–53.
- Ball, K., & Sekuler, R. (1987). Direction-Specific Improvement In Motion Discrimination. *Vision Research*, *27*(6), 953–965.
- Bargary, G., Bosten, J. M., Goodbourn, P. T., Lawrance-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). Individual differences in human eye movements: An oculomotor signature? *Vision Research*, *141*, 157–169. <https://doi.org/10.1016/j.visres.2017.03.001>
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17. <https://doi.org/10.1023/A:1005653411471>
- Bavelier, D., Shawn Green, C., Pouget, A., & Schrater, P. (2012). Brain plasticity through the life span: Learning to learn and action video games. *Annual Review of Neuroscience*, *35*, 391–416. <https://doi.org/10.1146/annurev-neuro-060909-152832>
- Bediou, B., Adams, D. M., Mayer, R. E., Tipton, E., Shawn Green, C., & Bavelier, D. (2018). Meta-analysis of action video game impact on perceptual, attentional, and cognitive skills. *Psychological Bulletin*, *144*(1), 77–110. <https://doi.org/10.1037/bul0000168>
- Bertamini, M., Cretenoud, A. F., & Herzog, M. H. (2019). Exploring the Extent in the Visual Field of the Honeycomb and Extinction Illusions. *I-Perception*, *10*(4), 2041669519854784. <https://doi.org/10.1177/2041669519854784>
- Bertamini, M., Herzog, M. H., & Bruno, N. (2016). The honeycomb illusion: Uniform textures not perceived as such. *I-Perception*, *7*(4), 1–15. <https://doi.org/10.1177/2041669516660727>

Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do Action Video Games Improve Perception and Cognition? *Frontiers in Psychology, 2*(226), 1–6. <https://doi.org/10.3389/fpsyg.2011.00226>

Bosten, J. M., Goodbourn, P. T., Bargary, G., Verhallen, R., Lawrance-Owen, A. J., Hogg, R. E., & Mollon, J. D. (2017). An exploratory factor analysis of visual performance in a large population. *Vision Research, 141*, 303–316. <https://doi.org/10.1016/j.visres.2017.02.005>

Bosten, J. M., & Mollon, J. D. (2010). Is there a general trait of susceptibility to simultaneous contrast? *Vision Research, 50*(17), 1656–1664. <https://doi.org/10.1016/j.visres.2010.05.012>

Campbell, M. J., Toth, A. J., Moran, A. P., Kowal, M., & Exton, C. (2018). eSports: A new window on neurocognitive expertise? *Progress in Brain Research, 240*, 161–174. <https://doi.org/10.1016/bs.pbr.2018.09.006>

Cappe, C., Clarke, A., Mohr, C., & Herzog, M. H. (2014). Is there a common factor for vision? *Journal of Vision, 14*(8), 1–11. <https://doi.org/10.1167/14.8.4>

Chopin, A., Bediou, B., & Bavelier, D. (2019). Altering perception: the case of action video gaming. *Current Opinion in Psychology, 29*, 168–173. <https://doi.org/10.1016/j.copsyc.2019.03.004>

Chung, S. T. L. (2013). Cortical reorganization after long-term adaptation to retinal lesions in humans. *Journal of Neuroscience, 33*(46), 18080–18086. <https://doi.org/10.1523/JNEUROSCI.2764-13.2013>

Clark, K., Fleck, M. S., & Mitroff, S. R. (2011). Enhanced change detection performance reveals improved strategy use in avid action video game players. *Acta Psychologica, 136*(1), 67–72. <https://doi.org/10.1016/j.actpsy.2010.10.003>

Costello, A. B., & Osborne, J. W. (2005). Best Practices in Exploratory Factor Analysis : Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research & Education, 10*(7), 86–99. <https://doi.org/10.1.1.110.9154>

Cretenoud, A. F., Francis, G., & Herzog, M. H. (2020). When illusions merge. *Journal of Vision, 20*(8), 1–15. <https://doi.org/10.1167/JOV.20.8.12>

Cretenoud, A. F., Grzeczowski, L., Bertamini, M., & Herzog, M. H. (2020). Individual differences in the Muller-Lyer and Ponzo illusions are stable across different contexts. *Journal of Vision, 20*(6), 1–14. <https://doi.org/10.1167/JOV.20.6.4>

Cretenoud, A. F., Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, *19*(14), 1–21. <https://doi.org/10.1167/19.14.12>

Dye, M. W. G., Green, C. S., & Bavelier, D. (2009). Increasing Speed of Processing With Action Video Games. *Current Directions in Psychological Science*, *18*(6), 321–326. <https://doi.org/10.1111/j.1467-8721.2009.01660.x>

Edwards, M., & Badcock, D. R. (1995). Global Motion Perception: No Interaction Between the First-and Second-order Motion Pathways. *Vision Research*, *35*(18), 2589–2602.

Emery, K. J., Volbrecht, V. J., Peterzell, D. H., & Webster, M. A. (2017). Variations in normal color vision. VII. Relationships between color naming and hue scaling. *Vision Research*, *141*, 66–75. <https://doi.org/10.1016/j.visres.2016.12.007>

Fahle, M., & Edelman, S. (1993). Long-term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision Research*, *33*(3), 397–412. [https://doi.org/10.1016/0042-6989\(93\)90094-D](https://doi.org/10.1016/0042-6989(93)90094-D)

Faubert, J., & Sidebottom, L. (2012). Perceptual-cognitive training of athletes. *Journal of Clinical Sport Psychology*, *6*(1), 85–102. <https://doi.org/10.1123/jcsp.6.1.85>

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science*, *18*(10), 850–855. <https://doi.org/10.1111/j.1467-9280.2007.01990.x>

Ferguson, C. J. (2007). The good, the bad and the ugly: A meta-analytic review of positive and negative effects of violent video games. *Psychiatric Quarterly*, *78*(4), 309–316. <https://doi.org/10.1007/s11126-007-9056-9>

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78. <https://doi.org/10.1016/j.paid.2016.06.069>

Green, C. S., & Bavelier, D. (2006). Enumeration versus multiple object tracking: the case of action video game players. *Cognition*, *101*(1), 217–245. <https://doi.org/10.1016/j.cognition.2005.10.004>

Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science*, *18*(1), 88–94. <https://doi.org/10.1111/j.1467-9280.2007.01853.x>

Greenwood, J. A., Szinte, M., Sayim, B., & Cavanagh, P. (2017). Variations in crowding, saccadic precision, and spatial localization reveal the shared topology of spatial vision. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(17), E3573–E3582. <https://doi.org/10.1073/pnas.1615504114>

Grzeczowski, L., Clarke, A. M., Francis, G., Mast, F. W., & Herzog, M. H. (2017). About individual differences in vision. *Vision Research*, *141*, 282–292. <https://doi.org/10.1016/j.visres.2016.10.006>

Grzeczowski, L., Cretenoud, A. F., Herzog, M. H., & Mast, F. W. (2017). Perceptual learning is specific beyond vision and decision making. *Journal of Vision*, *17*(6), 1–11. <https://doi.org/10.1167/17.6.6>

Grzeczowski, L., Cretenoud, A. F., Mast, F. W., & Herzog, M. H. (2019). Motor response specificity in perceptual learning and its release by double training. *Journal of Vision*, *19*(6), 1–14. <https://doi.org/10.1167/19.6.4>

Grzeczowski, L., Roinishvili, M., Chkonia, E., Brand, A., Mast, F. W., Herzog, M. H., & Shaqiri, A. (2018). Is the perception of illusions abnormal in schizophrenia? *Psychiatry Research*, *270*, 929–939. <https://doi.org/10.1016/j.psychres.2018.10.063>

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2018). *Multivariate data analysis* (8th ed.). Cengage Learning EMEA.

Harris, P. B., & Houston, J. M. (2010). A Reliability Analysis of the Revised Competitiveness Index. *Psychological Reports*, *106*(3), 870–874. <https://doi.org/10.2466/pr0.106.3.870-874>

Herzog, M. H., & Koch, C. (2001). Seeing properties of an invisible object: Feature inheritance and shine-through. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(7), 4271–4275. <https://doi.org/10.1073/pnas.071047498>

Herzog, M. H., Kopmann, S., & Brand, A. (2004). Intact figure-ground segmentation in schizophrenia. *Psychiatry Research*, *129*(1), 55–63. <https://doi.org/10.1016/j.psychres.2004.06.008>

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, *4*(1), 11–26. <https://doi.org/10.1080/17470215208416600>

Hutchinson, C. V., & Stocks, R. (2013). Selectively Enhanced Motion Perception in Core Video Gamers. *Perception*, *42*(6), 675–677. <https://doi.org/10.1068/p7411>

Iglewicz, B., & Hoaglin, D. (1993). Volume 16: how to detect and handle outliers. In *The ASQC basic references in quality control: statistical techniques*. Asq Press.

Jackson, D. N. (1974). *Personality Research Form Manual*. Research Psychologists Press.

Knudson, D., & Kluka, D. A. (1997). The Impact of Vision and Vision Training on Sport Performance. *Journal of Physical Education*, 68(4), 17–24. <https://doi.org/10.1080/07303084.1997.10604922>

Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>

Kowal, M., Toth, A. J., Exton, C., & Campbell, M. J. (2018). Different cognitive abilities displayed by action video gamers and non-gamers. *Computers in Human Behavior*, 88, 255–262. <https://doi.org/10.1016/j.chb.2018.07.010>

Lahav, K., Levkovitch-Verbin, H., Belkin, M., Glovinsky, Y., & Polat, U. (2011). Reduced mesopic and photopic foveal contrast sensitivity in glaucoma. *Archives of Ophthalmology*, 129(1), 16–22. <https://doi.org/10.1001/archophthalmol.2010.332>

Li, R., Polat, U., Makous, W., & Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nature Neuroscience*, 12(5), 549–551. <https://doi.org/10.1038/nn.2296>

Li, R., Polat, U., Scalzo, F., & Bavelier, D. (2010). Reducing backward masking through action game training. *Journal of Vision*, 10(14), 1–13. <https://doi.org/10.1167/10.14.33>

Mason, O., Linney, Y., & Claridge, G. (2005). Short scales for measuring schizotypy. *Schizophrenia Research*, 78(2–3), 293–296. <https://doi.org/10.1016/j.schres.2005.06.020>

Mollon, J. D., Bosten, J. M., Peterzell, D. H., & Webster, M. A. (2017). Individual differences in visual science: What can be learned and what is good experimental practice? *Vision Research*, 141, 4–15. <https://doi.org/10.1016/j.visres.2017.11.001>

Nazhif Rizani, M., & Iida, H. (2018). Analysis of Counter-Strike: Global Offensive. *Proceedings of 2018 International Conference on Electrical Engineering and Computer Science (ICECOS)*, 373–378. <https://doi.org/10.1109/ICECOS.2018.8605213>

Newsome, W. T., & Park, E. B. (1988). A Selective Impairment of Motion Perception Following Lesions of the Middle Temporal Visual Area (MT). *The Journal of Neuroscience*, 8(6), 2201–2211.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.

Patino, C. M., McKean-Cowdin, R., Azen, S. P., Allison, J. C., Choudhury, F., & Varma, R. (2010). Central and Peripheral Visual Impairment and the Risk of Falls and Falls with Injury. *Ophthalmology*, 117(2), 206. <https://doi.org/10.1016/j.ophtha.2009.06.063>

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.

Peterzell, D. H. (2016). Discovering Sensory Processes Using Individual Differences : A Review and Factor Analytic Manifesto. *Electronic Imaging*, 2016(16), 1–11. <https://doi.org/10.2352/ISSN.2470-1173.2016.16HVEI-112>

Peterzell, D. H., Scheffrin, B. E., Tregear, S. J., & Werner, J. S. (2000). Spatial frequency tuned covariance channels underlying scotopic contrast sensitivity. In *Vision Science and its Applications: OSA Technical Digest*. Optical Society of America.

Prado, J., Clavagnier, S., Otzenberger, H., Scheiber, C., Kennedy, H., & Perenin, M. T. (2005). Two cortical systems for reaching in central and peripheral vision. *Neuron*, 48(5), 849–858. <https://doi.org/10.1016/j.neuron.2005.10.010>

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/>

Roinishvili, M., Chkonia, E., Stroux, A., Brand, A., & Herzog, M. H. (2011). Combining vernier acuity and visual backward masking as a sensitive test for visual temporal deficits in aging research. *Vision Research*, 51(4), 417–423. <https://doi.org/10.1016/j.visres.2010.12.011>

Schoups, A. A., Vogels, R., & Orban, G. A. (1995). Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularly. *The Journal of Physiology*, 483(3), 797–810. <https://doi.org/10.1113/jphysiol.1995.sp020623>

Shawn Green, C., Li, R., & Bavelier, D. (2010). Perceptual Learning During Action Video Game Playing. *Topics in Cognitive Science*, 2(2), 202–216. <https://doi.org/10.1111/j.1756-8765.2009.01054.x>

Shawn Green, C., Sugarman, M. A., Medford, K., Klobusicky, E., & Bavelier, D. (2012). The effect of action video game experience on task-switching. *Computers in Human Behavior*, 28(3), 984–994. <https://doi.org/10.1016/j.chb.2011.12.020>

Shiu, L. P., & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception & Psychophysics*, 52(5), 582–588. <https://doi.org/10.3758/BF03206720>

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428.

Smither, R. D., & Houston, J. M. (1992). The Nature of Competitiveness: The Development and Validation of the Competitiveness Index. *Educational and Psychological Measurement*, 52(2), 407–418. <https://doi.org/10.1177/0013164492052002016>

Spang, K., Grimsen, C., Herzog, M. H., & Fahle, M. (2010). Orientation specificity of learning vernier discriminations. *Vision Research*, 50(4), 479–485. <https://doi.org/10.1016/j.visres.2009.12.008>

Spence, I., & Feng, J. (2010). Video Games and Spatial Cognition. *Review of General Psychology*, 14(2), 92–104. <https://doi.org/10.1037/a0019491>

Spinella, M. (2007). Normative data and a short form of the Barratt Impulsiveness Scale. *International Journal of Neuroscience*, 117(3), 359–368. <https://doi.org/10.1080/00207450600588881>

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *The Journal of the Acoustical Society of America*, 41(4), 782–787. <https://doi.org/https://doi.org/10.1121/1.1910407>

Tibber, M. S., Guedes, A., & Shepherd, A. J. (2006). Orientation Discrimination and Contrast Detection Thresholds in Migraine for Cardinal and Oblique Angles. *Investigative Ophthalmology & Visual Science*, 47(12), 5599–5604. <https://doi.org/10.1167/iovs.06-0640>

Tulver, K. (2019). The factorial structure of individual differences in visual perception. *Consciousness and Cognition*, 73:102762. <https://doi.org/https://doi.org/10.1016/j.concog.2019.102762>

Verhallen, R. J., Bosten, J. M., Goodbourn, P. T., Lawrance-owen, A. J., Bargary, G., & Mollon, J. D. (2017). General and specific factors in the processing of faces. *Vision Research*, *141*, 217–227. <https://doi.org/10.1016/j.visres.2016.12.014>

Wagner, M. G. (2006). On the Scientific Relevance of eSports. *International Conference on Internet Computing*, 437–442.

Watson, A. B., & Pelli, D. G. (1979). The QUEST staircase procedure. *Applied Vision Association Newsletter*, *14*, 6–7.

Watson, A. B., & Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*(2), 113–120. <https://doi.org/10.3758/BF03202828>

Yashar, A., Wu, X., Chen, J., & Carrasco, M. (2019). Crowding and Binding: Not All Feature Dimensions Behave in the Same Way. *Psychological Science*, *30*(10), 1533–1546. <https://doi.org/10.1177/0956797619870779>

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301–320. <https://doi.org/10.1111/J.1467-9868.2005.00503.X>

8. Supplementary File

8.1. Supplementary figures

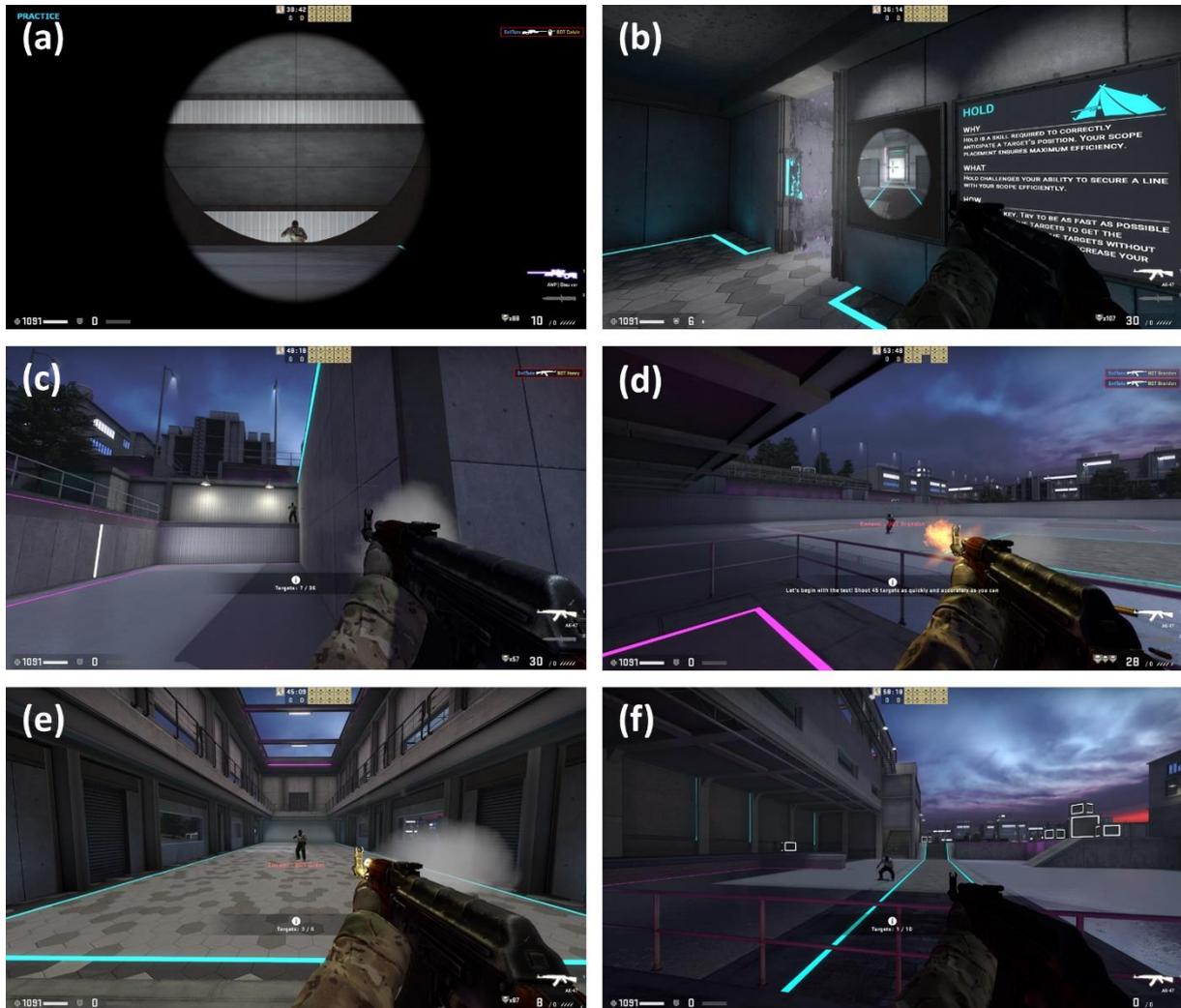


Figure S1. Screenshots of the (a) flick, (b) hold, (c) peek, (d) shoot, (e) spray, and (f) track CS:GO mini-games, which are publicly available on playmaster.gg.

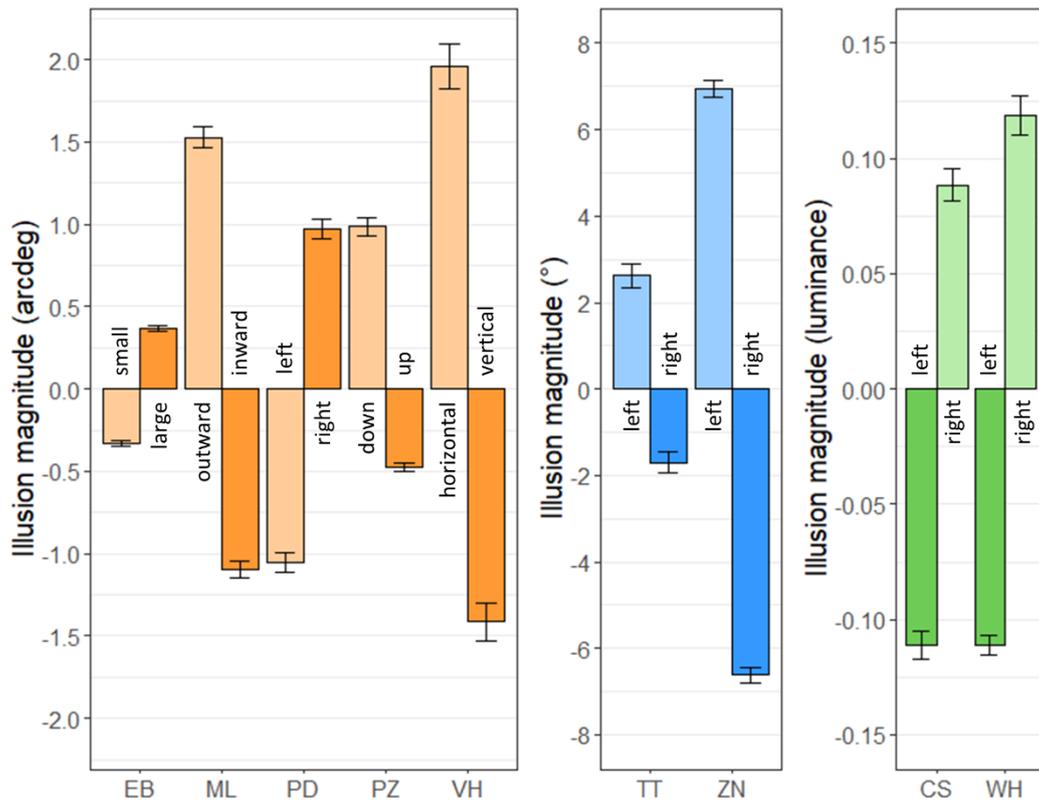


Figure S2. Illusion magnitudes \pm SE (standard error) for both conditions of each illusion from the Ill paradigm. Positive and negative magnitudes indicate over- and under-adjustments compared to the reference, respectively. Left panel (orange): illusions with adjustment of size, length, or position measured in units of arcdeg (EB: Ebbinghaus; ML: Müller-Lyer; PD: Poggendorff; PZ: Ponzo; VH: vertical-horizontal); middle panel (blue): illusions with orientation adjustment measured in degrees (TT: Tilt; ZN: Zöllner); right panel (green): illusions with contrast adjustment measured in normalized luminance (CS: contrast; WH: White).

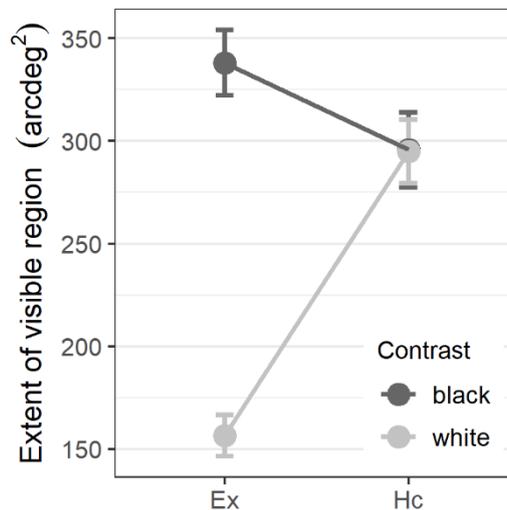


Figure S3. Mean extent of the region (in units of arcdeg²) in which bars (Honeycomb) and disks (Extinction) were perceptible as a function of the contrast polarity (black or white) in the Hc/Ex paradigm. Error bars show standard errors (SE).

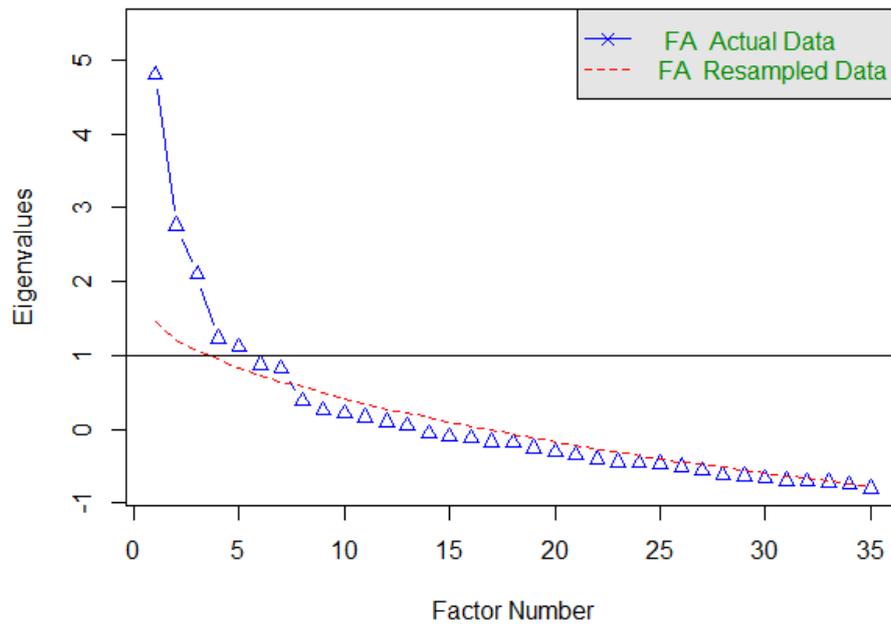


Figure S4. Scree plot from an exploratory factor analysis, in which a common factor analysis was computed to extract the factors. Eigenvalues are shown for the actual data (in blue) and for resampled data (in red). A three-factor model is suggested by scree plot inspection, while a parallel analysis suggested a seven-factor model.

8.2. Supplementary tables

Table S1. Conversion from ordinal CS:GO ranks to numerical equivalents

Ordinal	Numerical
Silver I	1
Silver II	2
Silver III	3
Silver IV	4
Silver Elite	5
Silver Elite Master	6
Gold Nova I	7
Gold Nova II	8
Gold Nova III	9
Gold Nova Master	10
Master Guardian I	11
Master Guardian II	12
Master Guardian Elite	13
Distinguished Master Guardian	14
Legendary Eagle	15
Legendary Eagle Master	16
Supreme Master First Class	17
The Global Elite	18

Table S2. Test-retest reliabilities expressed as intraclass coefficients (ICCs) for all variables extracted from the Hc/Ex and III paradigms, visual acuity (VA) and visual backward masking (VBM).

Paradigm	Variable	ICC	95% CI	df	<i>p</i>
Hc/Ex	HcBlack	0.872	[0.814, 0.913]	[93, 93]	<0.001
	HcWhite	0.822	[0.743, 0.878]	[93, 93]	<0.001
	ExBlack	0.787	[0.696, 0.853]	[93, 93]	<0.001
	ExWhite	0.848	[0.780, 0.897]	[93, 93]	<0.001
III	CS left	0.561	[0.405, 0.685]	[93, 93]	<0.001
	CS right	0.569	[0.415, 0.692]	[93, 93]	<0.001
	EB small	0.723	[0.610, 0.807]	[93, 93]	<0.001
	EB large	0.605	[0.459, 0.719]	[93, 93]	<0.001
	ML in	0.638	[0.500, 0.744]	[93, 93]	<0.001
	ML out	0.528	[0.365, 0.659]	[93, 93]	<0.001
	PD left	0.722	[0.609, 0.806]	[93, 93]	<0.001
	PD right	0.666	[0.536, 0.765]	[93, 93]	<0.001
	PZ down	0.639	[0.501, 0.744]	[93, 93]	<0.001
	PZ up	0.371	[0.183, 0.533]	[93, 93]	<0.001
	TT left	0.317	[0.123, 0.488]	[93, 93]	0.001
	TT right	0.399	[0.215, 0.556]	[93, 93]	<0.001
	VH hor	0.530	[0.367, 0.661]	[93, 93]	<0.001
	VH ver	0.756	[0.653, 0.831]	[93, 93]	<0.001
	WH left	0.110	[-0.093, 0.305]	[93, 93]	0.144
	WH right	0.508	[0.341, 0.644]	[93, 93]	<0.001
	ZN left	0.714	[0.599, 0.801]	[93, 93]	<0.001
	ZN right	0.715	[0.599, 0.801]	[93, 93]	<0.001
VA		0.847	[0.778, 0.895]	[93, 93]	<0.001
VBM		0.847	[0.778, 0.896]	[93, 93]	<0.001

Table S3. Statistics from Shapiro-Wilk tests and lambda (λ) exponent from an optimized Tukey transformation that maximizes normality for each visual (green), gaming (orange), and questionnaire (purple) variable according to the Shapiro-Wilk test

	CrwSize	CrwPeri	Ctr	HcBlack	HcWhite	ExBlack	ExWhite	CS
SW statistic	0.925**	0.877**	0.732**	0.904**	0.947*	0.960*	0.820**	0.980
SW statistic Z	0.987	0.989	0.982	0.987	0.992	0.993	0.987	0.986
lambda (λ)	-0.70	0.45	0.10	0.40	0.40	0.55	0.45	1.10
	EB	ML	PD	PZ	TT	VH	WH	ZN
SW statistic	0.990	0.984	0.986	0.903**	0.955*	0.984	0.982	0.977
SW statistic Z	0.991	0.992	0.987	0.989	0.982	0.991	0.992	0.990
lambda (λ)	0.95	1.30	0.90	0.35	0.55	0.75	0.65	1.15
	NB	Ori	RDKh	RDKr	RT	proTravel	proSac	antiTravel
SW statistic	0.960*	0.623**	0.765**	0.796**	0.923**	0.554**	0.808**	0.486**
SW statistic Z	0.967*	0.988	0.971*	0.976	0.991	0.989	0.985	0.983
lambda (λ)	1.95	-0.35	0.05	0.05	-1.90	1.55	-0.75	0.70
	antiSac	VA	VBM	VS4	VS16	Shoot	Spray	Track
SW statistic	0.967*	0.982	0.820**	0.947**	0.812**	0.975	0.983	0.955*
SW statistic Z	0.993	0.986	0.968*	0.989	0.991	0.988	0.983	0.955*
lambda (λ)	-0.80	1.40	0.20	0.00	-1.10	2.05	0.85	1.00
	Hold	Flick	Peek	AQ SS	AQ AS	AQ AD	AQ CO	AQ IM
SW statistic	0.989	0.983	0.986	0.943**	0.939**	0.956*	0.959*	0.950*
SW statistic Z	0.991	0.990	0.987	0.954*	0.940**	0.958*	0.959*	0.950*
lambda (λ)	2.00	1.85	1.10	0.80	1.10	1.20	1.00	1.05
	BIS NI	BIS MI	BIS AI	CI EC	CI CO	HEXACO HH	HEXACO EM	HEXACO EX
SW statistic	0.976	0.976	0.965*	0.932**	0.977	0.986	0.988	0.978
SW statistic Z	0.977	0.980	0.981	0.979	0.978	0.990	0.988	0.985
lambda (λ)	0.75	0.60	0.25	2.65	0.85	1.45	1.10	1.65
	HEXACO AG	HEXACO CO	HEXACO OP	Handedness	O-LIFE UE	O-LIFE CD	O-LIFE IA	O-LIFE IN
SW statistic	0.987	0.960*	0.975	0.725**	0.929**	0.958*	0.932**	0.950*
SW statistic Z	0.990	0.983	0.990	0.933**	0.954*	0.968*	0.945**	0.959*
lambda (λ)	1.40	2.35	1.95	3.55	0.75	0.75	0.80	0.80
	PRFd	NbHoursPerWeek	NbTotalHours	Actual CSGO rank	Best CSGO rank	NbHoursSleep		
SW statistic	0.974	0.828**	0.555**	0.929**	0.928**	0.946**		
SW statistic Z	0.976	0.975	0.973*	0.946**	0.930**	0.968*		
lambda (λ)	0.80	0.00	0.25	0.45	1.20	2.45		

The Shapiro-Wilk test was run both before (SW statistic) and after (SW statistic Z) data were transformed using a Tukey transformation with the optimized λ exponent (outliers were removed only in the visual variables). Significant statistics indicate a violation of the normality assumption. * $p < 0.05$, ** $p < 0.05/62$ (Bonferroni correction for multiple comparisons).

Appendix G

Rashal, E., Cretenoud, A. F., & Herzog, M. H. (2020). Perceptual grouping leads to objecthood effects in the Ebbinghaus illusion. *Journal of Vision*, 20(8):11, 1–15, <https://doi.org/10.1167/jov.20.8.11>.

Copyright holder: © ARVO

Published under CC-BY-NC-ND 4.0

Perceptual grouping leads to objecthood effects in the Ebbinghaus illusion

Einat Rashal

Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland



Aline F. Cretenoud

Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland



Michael H. Herzog

Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland



The Ebbinghaus illusion is argued to be a product of low-level contour interactions or a higher cognitive comparison process. We examined the effect of grouping on the illusion by manipulating objecthood, i.e., the degree to which an object is a cohesive perceptual entity. We hypothesized that reduced objecthood would decrease the illusion magnitude, because the objects become less efficient in the comparison process. To test this hypothesis, we used a version of the illusion where the target and flanking objects were squares that were composed from their corners or sides. Degree of objecthood was manipulated by changing the gap size or rotation angle of the elements constructing the objects, so that larger gaps and angles produced less cohesive objects than smaller. Participants performed an adjustment procedure on the test target to match a control target in size. In addition, subjective reports of the objects' shape were collected as a measure of perceived shape. Our results show decreased illusion magnitude with increasing gap size and rotation angle. Surprisingly, the perceived shape of the objects did not correlate with illusion magnitude. These results provide novel evidence of the role of mid-level processes in the Ebbinghaus illusion and point to a dissociation between subjective and objective measures of objecthood.

Coren & Enns, 1993; Coren & Miller, 1974; Massaro & Anderson, 1971; Vuk & Podlesek, 2005), whereas the latter poses that the target is perceived as stretched or contracted because the target's contours are attracted or repulsed by the contours of the flankers (e.g., Jaeger, 1978; Jaeger & Klahs, 2015; Sherman & Chouinard, 2016; Schwarzkopf & Rees, 2013; Todorović and Jovanović, 2018; Weintraub & Schneck, 1986).

An important distinction can be made between the two accounts: the first would have to consider some qualities of the objects available for comparison, whereas the second is more concerned with the position of contours around the target. To provide good standards for the comparison, the flankers should give ample evidence that they are similar to the target. Such evidence comes, for example, from their number and their distance, because a larger number of flankers provides more information, and the closer the flankers are to the target the more relevant they are for the comparison (Massaro & Anderson, 1971). Shape was suggested to play a role as well, because flankers and targets with similar shapes can be better compared than objects with different shapes (Coren & Miller, 1974). On the other hand, for contour interactions to occur, contours should be strategically located in close and far positions relative to the target to affect its size perception (e.g., Jaeger, 1978).

A recent study by Jaeger and Klahs (2015) tested the relative strengths of the two accounts by studying the effect of the number of flankers on the illusion. In that study, a few small flankers surrounded the target and as the number of flankers increased, they led to circular configurations (i.e., with the maximum number of flankers, the target was flanked by large circular groups of small circles). The authors reasoned that increasing the number of small flankers would result in an increase of illusion magnitude (i.e., the target will appear

Introduction

In the Ebbinghaus illusion, a target looks smaller when it is surrounded by larger flanking objects or looks bigger when it is surrounded by smaller ones. This phenomenon has been offered two accounts: a cognitive size contrast mechanism or low-level contour interactions. The former proposes that the size of the target is judged in comparison to the size of the surrounding stimuli (e.g., Choplin & Medin, 1999;

Citation: Rashal, E., Cretenoud, A. F., & Herzog, M. H. (2020). Perceptual grouping leads to objecthood effects in the Ebbinghaus illusion. *Journal of Vision*, 20(8):11, 1–15, <https://doi.org/10.1167/jov.20.8.11>.



even larger) according to the size contrast account, because of the increasing salience of the standard in the comparison. On the other hand, according to the contour interaction account, this manipulation would cause the illusion to change direction (i.e., the target will appear to be larger and then decreasing in size with increasing number of small flankers), as repulsion will be generated from the flankers on the outer perimeter of that configuration as small flankers are added. Their results supported the second hypothesis, replicating a similar switching effect reported by Weintraub and Schneck (1986), who manipulated the length of flanking contours, starting from small arcs and progressing onto full large flanking circles. In both studies, this effect was interpreted as evidence in favor of the contour interaction account, however, as mentioned by Jaeger and Klahs (2015), the switch in direction could have been due to the larger circular configurations effectively functioning as large flankers, in which case, the effect can be explained by the size contrast account as well.

Influence of mid-level organizational process such as the grouping of small flankers reported in Jaeger and Klahs (2015) would support the cognitive size contrast account and challenge the low-level contour interaction account. However, the role of perceptual organization in the Ebbinghaus illusion has not been tested directly yet. In the current study, we examined the effect of mid-level processes on the Ebbinghaus illusion by measuring the effect of *objecthood* on the illusion magnitude, objecthood being the degree to which an object is a cohesive perceptual entity. We theorized that a degraded object would be less efficient in the comparison. Objecthood was defined in terms of grouping strength between object parts—stronger grouping leads to more cohesive objects (Kimchi, Yeshurun, Spehar, & Pirkner, 2015). In our experiments, we presented observers a variant of the classic Ebbinghaus illusion using square target and flankers. Objecthood was manipulated by varying the percentage of the visible contour of the squares (Experiments 1–3), or the rotation angle of the elements constructing the squares (Experiment 4). Importantly, the visible contour of the flankers was presented in a way that the ratio between close and far contours relative to the target did not change (relative to the center of each flanker). According to the contour interaction account, varying objecthood in this way should not change the illusion magnitude, as the ratio between attraction and repulsion from the flankers' contours on the target does not change¹. However, according to the size contrast account, a decrease in visible contour should lead to a weaker illusion because the objects would become less “object-like.” This would make the comparison less effective since the objects would become less similar. Thus our hypothesis was that the illusion would be weaker with decreasing objecthood of either the target or the flankers. Specifically, the illusory

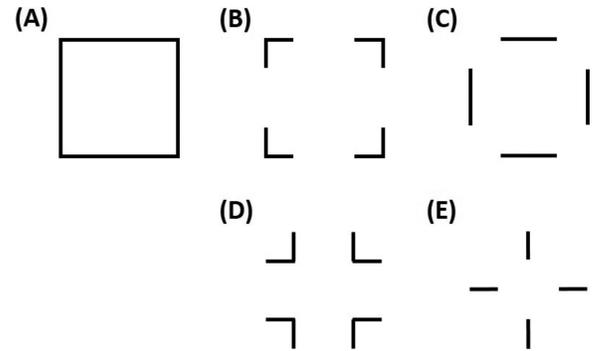


Figure 1. Example stimuli employed in Experiments 1–3: Complete square (A), square constructed by corners (B), and square constructed by sides (C). In Experiment 3 flankers had a cross-like shape constructed by corners (D), or lines (E). Note that illustrations B–E depict the objects with a gap size of 0.5 (see text for details).

effect on the estimation of target size should decrease with a decreasing percentage of visible contour or with increasing rotation angle.

Experiment 1

Our first experiment was designed to establish whether an objecthood effect can be found on the Ebbinghaus illusion using squares with a gap-size manipulation. To that end, we manipulated the percentage of visible contour of either the target or the flankers. Using squares allowed us to manipulate objecthood also by means of different grouping cues (Figure 1). Collinearity between object segments, like the corners of a square, contributes to a stable representation of the shape (Kimchi et al., 2015; Hadad & Kimchi, 2008). Thus, if the illusion is sensitive to objecthood, there may be a difference between conditions displaying squares made of corners or squares made out of sides. That is, the illusion magnitude may decrease to a different degree when the target is composed of corners or surrounded by flankers composed of corners compared with when it is composed of sides or surrounded by flankers composed of sides.

Corners are also powerful in carrying shape information (e.g., Persike & Meinhardt, 2017; Poirier & Wilson, 2007). Because it is difficult to disentangle shape from objecthood—low objecthood will almost certainly lead to a degraded perception of the object's shape—there may be a correlation between the perceived shape of the object and the magnitude of the size illusion. To test this, we also collected subjective reports of shape and objecthood of the different objects in the displays. We hypothesized that

the percentage of visible contour and element type (i.e., sides/corners) would affect the perceived shape of the object. Specifically, we expected a decreasing percentage of “square” reports with decreasing percentage of visible contour and a difference in the reports between objects made of corners or made of sides.

Method

Participants

Ten naïve students from the École Polytechnique Fédérale de Lausanne (EPFL) participated in this experiment (three females; mean age: 21.40 years; range 18–26 years). Participants signed informed consent forms and were paid for their participation. Procedures were conducted in accordance with the Declaration of Helsinki (except for preregistration) and were approved by the local ethics committee.

Apparatus

Participants sat at a distance of ≈ 57 cm from the screen. Stimuli were presented on a BenQ XL2420T monitor driven by a PC computer using Matlab (R2014b, 64 bits) and the Psychophysics toolbox (Brainard, 1997; Pelli, 1997; version 3.1, 64 bits) at a 1920- × 1080-pixel resolution with a 60-Hz refresh rate. Participants used a Logitech LS1 computer mouse to adjust stimuli on the screen. The screen luminance was measured and controlled before the experiment with a Minolta LS-100 luminance meter.

Stimuli

The stimuli were based on a square version of the Ebbinghaus illusion (Figure 2, top panel). The squares were presented in white (≈ 176 cd/m²) with a stroke width of two pixels on a black background (≈ 1 cd/m²). Following previous studies (e.g., Cretenoud, Karimpur, Grzeczowski, Francis, Hamburger, & Herzog, 2019), targets were positioned 8.6° from the center of the screen on both sides. A configuration on the left depicted the *reference target* (i.e., a square surrounded by smaller squares), and it was centered at 2.45° above the medial axis of the screen. The configuration on the right depicted the *test target* surrounded by larger flankers, centered at 2.45° below the medial axis of the screen. The misalignment on the horizontal meridian was designed to prevent participants strategically drawing imaginary lines between the outlines of the target and reference. The reference target subtended 2.45° on each side and was surrounded by eight smaller squares subtending 1.22° on each side. The center-to-center distance between target and flankers in the reference configuration was 2.6°. The test target was

surrounded by six flankers subtending 3.18° on each side, with a center-to-center distance of 4.5°.

The test target or the flankers in both configurations were composed of complete or incomplete squares. The incomplete squares were constructed from separate L-shaped elements (i.e., corners) or lines (i.e., sides), forming a square by means of grouping operations (see Figure 1). In each display, when the test target was an incomplete square, the flankers (in both test and reference configurations) were complete squares, and when the test target was a complete square the flankers were incomplete squares. In conditions where the squares were incomplete, the percentage of visible contour was between 10–90%. The gap size between the segments of visible contour varied in steps of 10% on each side of the square. Example stimuli in Experiment 1 with minimum (0.1) and maximum (0.9) gap size can be found in the Supplementary File (Supplementary Figure S1). Where corners were presented, gaps were formed from the centers of the square’s sides, increasing in size symmetrically towards the corners of the square. Where sides were presented, gaps were formed in the same manner from the corners of the square toward its center. The reference target was always a complete square, to prevent participants from adopting a strategy in which they matched the gap size on one side of the targets instead of the overall size of the object.

Procedure

Adjustment task

Before displaying the stimuli on the computer screen, the experimenter explained the task to the participant using a paper drawing of the illusion. Participants were instructed to base their adjustments only on their subjective perception.

Participants had to match the size of the test target to the size of the reference target. The factor of object role (i.e., target/flankers as incomplete squares) was manipulated between blocks. The factor of element type (i.e., corners/sides) was manipulated independently. Hence, in one block, the test target was a “*corners-target*” or a “*sides-target*” (i.e., a square composed of corners or sides, respectively; Figures 2A and 2B). In another block, the flankers were incomplete squares; “*corners-flankers*” consisting of corners, or “*sides-flankers*” consisting of sides (Figures 2C and 2D). The latter also included a condition where the targets were presented without flankers, and a baseline condition with complete squares only. The order of the blocks was randomized between participants. Gap size and element type were chosen randomly within a block, with a restriction that two consecutive adjustments were made on the exact same condition (e.g., flankers

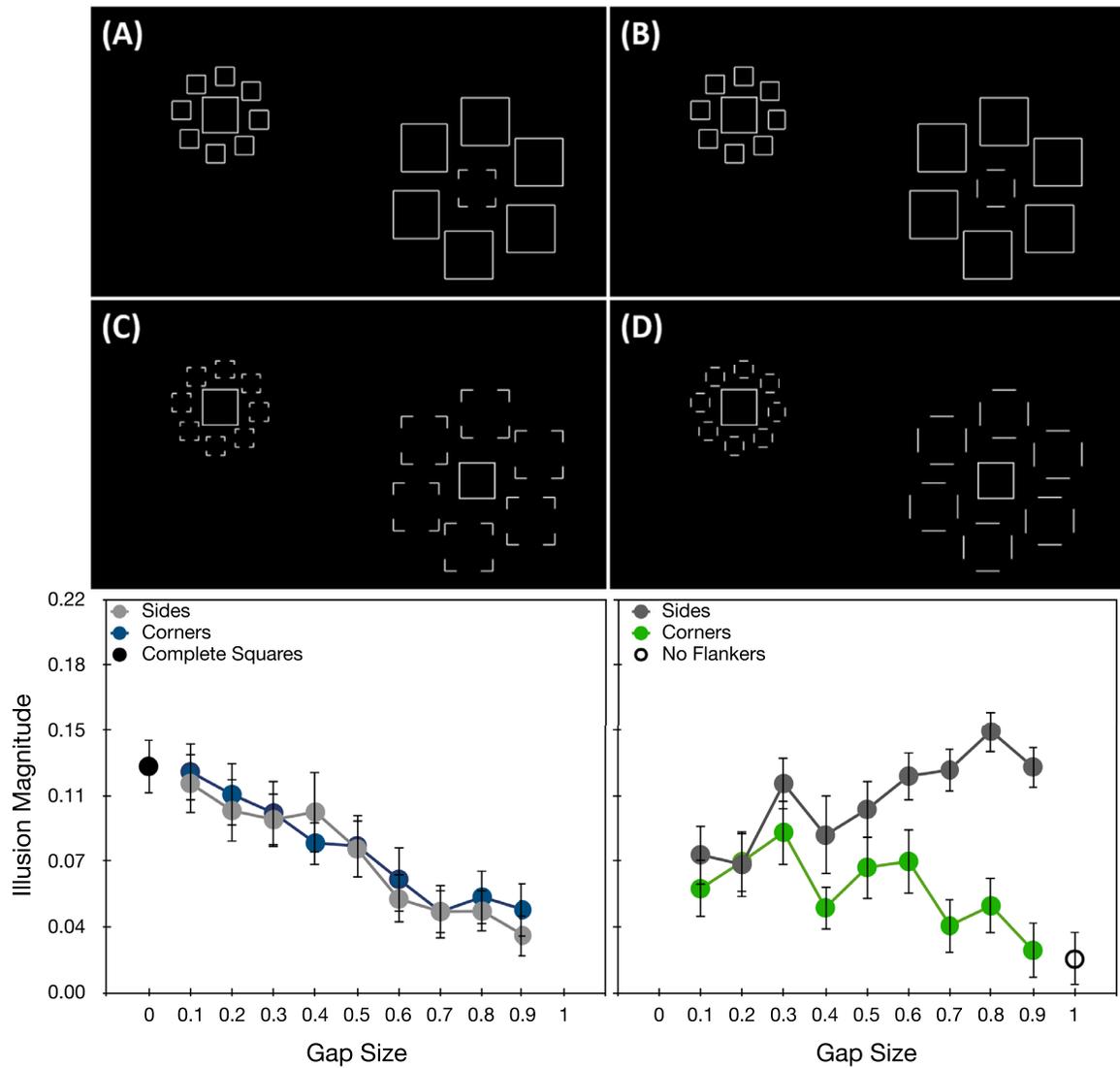


Figure 2. Top: examples of stimuli used in [Experiment 1](#). The reference configurations on the left present a reference target surrounded by smaller flankers. The test configuration on the right presents the adjustable test target surrounded by larger flankers. (A) Corners-target (i.e., a target composed of corners). (B) Sides-target (i.e., a target composed of sides). (C) Corners-flankers (i.e., flankers composed of corners). (D) Sides-flankers (i.e., flankers composed of sides). Gap size in these examples is 0.5. The reference target was always a complete square. Bottom: mean illusion magnitude for different element type conditions as a function of gap size for flankers (left) and targets (right) in [Experiment 1](#). Error bars indicate standard errors of the means.

formed by corners with 0.7 gap size). The size of the test target was adjusted by moving the computer mouse on the horizontal axis: a rightward movement increased the size and a leftward movement reduced it. The initial size of the test target was determined randomly at the beginning of each trial. The maximal size for adjusting the test target was 4° (on each side), to avoid any overlap with the surrounding flankers. To validate a trial, participants pressed the left button of the computer mouse. There was no time restriction, and no feedback was provided. Overall, there were 76 trials.

Subjective reports

After completing both blocks of the adjustment task, participants were asked to report their perception of the grouped target or flankers for each combination of part type and gap size. The displays were the ones containing gaps from the adjustment task (i.e., not including the ones without flankers or with complete squares only). In addition, the size of the targets in the two configurations of the display was identical. Participants indicated whether they perceived the elements of the grouped target or flankers as: (a) a

square, (b) another shape, or (c) unrelated elements, by pressing designated keys. Each display was evaluated once, resulting in a total of 36 trials.

Results and discussion

Adjustment task

Consecutive adjustments for each condition were averaged for each participant. Illusion magnitude was computed as a percentage of error in size adjustment compared to the reference target (i.e., adjusted target size minus reference size divided by reference size). The p -values were corrected with Greenhouse-Geisser epsilon in cases of sphericity violation. Illusion magnitude was subjected to a three-way repeated measures analysis of variance (ANOVA) with gap size, object role, and element type as within-subject factors, excluding the conditions with complete squares or no-flankers (Figure 2). The analysis revealed a main effect of gap size [$F(8, 72) = 5.06, p < 0.001, \eta_p^2 = 0.36$], showing decreased illusion magnitude with increasing gap size, and a main effect of element type [$F(1, 9) = 84.93, p < 0.001, \eta_p^2 = 0.90$], showing a higher magnitude for sides ($M = 0.09, SD = 0.6$) compared with corners ($M = 0.07, SD = 0.6$). The interaction between the two factors was significant [$F(8, 72) = 2.4, p < 0.02, \eta_p^2 = 0.21$], as were the interactions between object role and element type [$F(1, 9) = 40.82, p < 0.001, \eta_p^2 = 0.82$], object role and gap size [$F(8, 72) = 9.05, p < 0.001, \eta_p^2 = 0.50$], and among all three factors [$F(8, 72) = 2.8, p = 0.01, \eta_p^2 = 0.24$]. The main effect of object role did not reach significance [$F < 1$].

The data from the different object roles were further subjected to separate two-way repeated measures ANOVAs with gap size and element type as within-subject factors. For flankers as grouped squares, the analysis revealed a significant main effect of gap size, showing a decrease in magnitude with an increase in gap size [$F(8, 72) = 18.67, p < 0.001, \eta_p^2 = 0.68$]. The effect of element type and its interaction with gap size were not significant [$F(8, 72) = 2.58, p = 0.14, \eta_p^2 = 0.22; F < 1$, respectively]. For targets as grouped squares, the analysis revealed a significant main effect of element type [$F(1, 9) = 59.76, p < 0.001, \eta_p^2 = 0.87$], showing higher magnitude for sides-targets ($M = 0.11, SD = 0.6$) compared with corners-targets ($M = 0.6, SD = 0.6$), and an interaction between element type and gap size [$F(8, 72) = 3.53, p = 0.02, \eta_p^2 = 0.28$], showing a decrease in magnitude with an increase in gap size for corners-targets [$F(8, 72) = 2.75, p = 0.05, \eta_p^2 = 0.24$], and an increase in magnitude with an increase in gap size for sides-targets [$F(8, 72) = 2.53, p = 0.02, \eta_p^2 = 0.22$]. The effect of gap size did not reach significance [$F(8, 72) = 1.76, p = 0.1, \eta_p^2 = 0.16$].

Clearly, our objecthood manipulation affected the Ebbinghaus illusion, because for most conditions the illusion magnitude decreased with a decrease in visible contour. The odd case of sides-targets showing larger adjustment errors with decreasing amount of visible contour poses a challenge for the objecthood hypothesis at first glance, as it would be expected that the contextual effect would be in the same direction whether the grouped object is the target or the flankers. However, we believe it reflects a difficulty in comparing the degraded target to the reference rather than the influence of contextual objects. That is, integration of the target's contours to an imaginary square is impaired once the corners are removed, hence, the contours are integrated into an alternative shape, which is smaller in surface than the original square. Thus trying to match this smaller shape to the larger square will result in an adjustment bias regardless of the surrounding objects. Consequently, as the contour segments become smaller, the larger the difference between the alternative shape and the reference square will be. We find support for this idea in the results of the subjective reports portion of this experiment, described in the following section. In addition, an adjustment bias for the sides-target is confirmed in Experiment 3.

Subjective reports

Pearson's chi-squared tests were conducted on the distributions in the four different flankers and targets conditions (Figure 3). Reports were correlated with element type for flankers and targets [$\chi^2(2) = 29.9, p < 0.0001; \chi^2(2) = 63.48, p < 0.0001$, respectively]. That is, corners-targets were perceived as a square to a greater degree than sides-target (Figures 3A and 3B), and the same was true for corners-flankers versus sides-flankers (Figures 3C and 3D). In addition, reports were correlated with object role for sides but not for corners [$\chi^2(2) = 10.02, p < 0.01; \chi^2(2) = 0.71, p > 0.1$, respectively]. That is, corners-target and corners-flankers were perceived as squares to the same degree (Figures 3A and 3C), but sides-targets were perceived as a square less often than as another shape (38% and 54%, respectively; Figure 3B), whereas sides-flankers were perceived as squares more often than as another shape (57% and 31%, respectively; Figure 3D). Importantly, as clearly evident in the plots, the percentage of "square" reports decreased with increasing gap size, while the percentage of "other shape" reports increased, and the percentage of "unrelated elements" was higher for flankers than for targets formed by their sides (12% and 8%, respectively). Thus the results from the subjective reports are in line with previous findings on grouping and shape perception, showing that grouped corners lead to a better perception of a square compared with side elements (e.g., Hadad & Kimchi, 2008). However, they do not correspond to the results of the adjustment

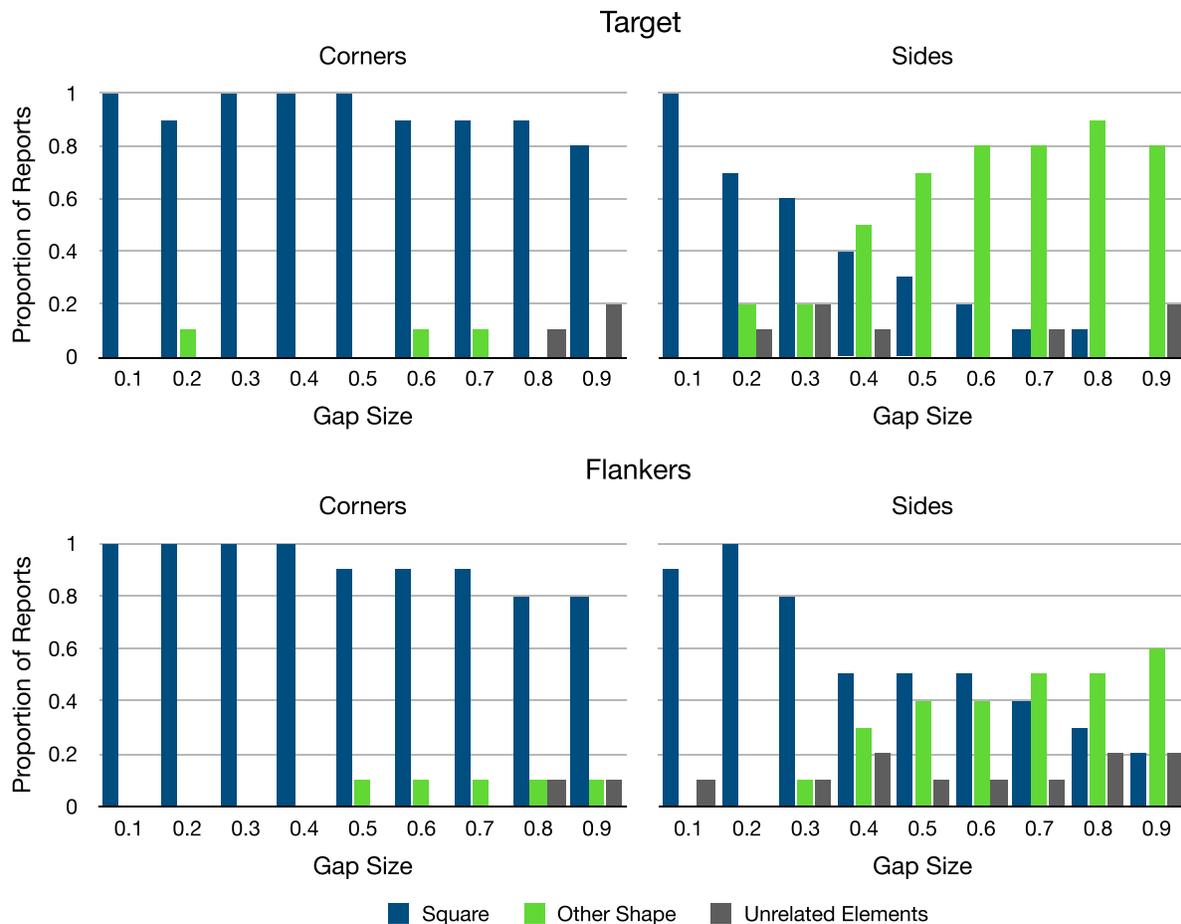


Figure 3. Proportion of reports at each gap size in the different object role and element type conditions in [Experiment 1](#).

task. This discrepancy suggests the Ebbinghaus illusion is not affected directly by the perceived shape of the objects in the configuration. [Experiment 2](#) was designed to explore this question further.

Experiment 2

The lack of compatibility between subjective reports and the contextual effect found in [Experiment 1](#) raises an important question, because it was previously suggested that the magnitude of the illusion depends on similarity between the target and its surrounding flankers (e.g., [Choplin & Medin, 1999](#); [Coren & Enns, 1993](#); [Coren & Miller, 1974](#); [Vuk & Podlesek, 2005](#)). [Coren and Miller \(1974\)](#), for example, demonstrated a correlation between illusion magnitude and similarity ratings, showing a weaker illusion with decreasing similarity. Later, [Vuk and Podlesek \(2005\)](#) suggested that the flankers' effects depend on their configuration as a whole rather than their global shape. That is, the combination of elements constructing the target and flankers could be more important than their outlines. This issue is pertinent to our study because the subjective reports

in [Experiment 1](#) indicate that sides-flankers resemble squares less than corners-flankers, thus creating a difference in their global shape with the gap-size manipulation. Still, it is possible that the two types of incomplete square flankers were equally dissimilar to the target, which was a complete square, resulting in a similar effect on illusion magnitude. To examine this possibility, in [Experiment 2](#) we varied the degree of similarity between target and flankers by presenting incomplete targets with incomplete flankers in every possible combination of element type. For example, a corners-target surrounded by corners-flankers should have greater similarity than a corners-target surrounded by sides-flankers. If configural similarity is a factor in the objecthood effect, we expected to find a higher illusion magnitude when the target and flankers were composed of the same elements compared with when they were composed from different elements. The gap size of the target was fixed at 0.5 in this experiment. Due to the difference in illusion magnitudes observed in [Experiment 1](#) for the two target types with this gap size we expected an effect in [Experiment 2](#) as well, showing higher adjustment error for the sides-target.

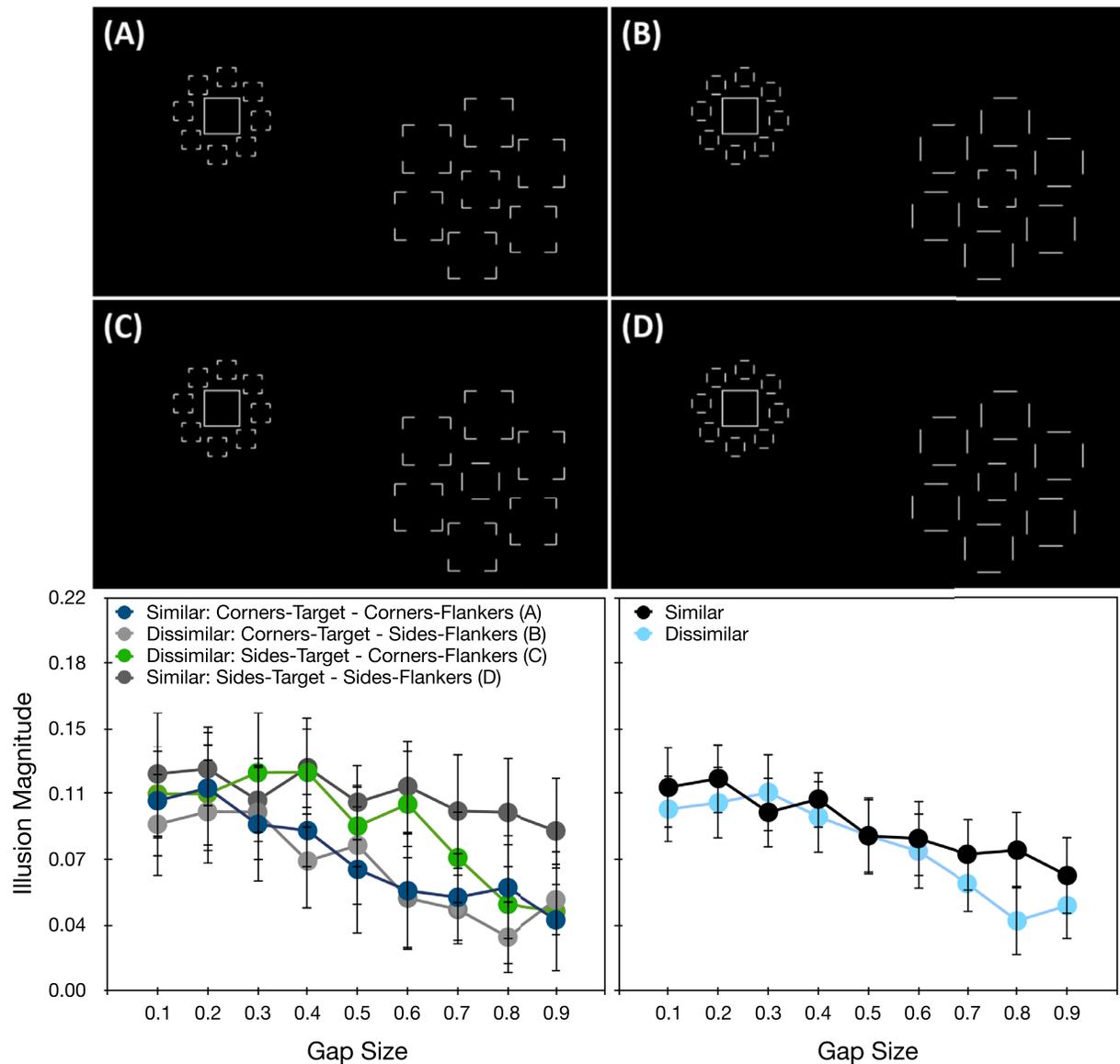


Figure 4. Top: examples of stimuli used in [Experiment 2](#). (A) Similar: corners-target and corners-flankers. (B) Dissimilar: corners-target and sides-flankers. (C) Dissimilar: sides-target and corners-flankers. (D) Similar: sides-target and sides-flankers. Gap size is 0.5 in these examples. The reference target was always a complete square. Bottom: mean illusion magnitude for different target and similarity conditions as a function of flankers' gap size (left), and main effect of similarity (right) in [Experiment 2](#). Error bars indicate standard errors of the means.

Method

Participants

Ten naïve students, new to the study, from the École Polytechnique Fédérale de Lausanne (EPFL) participated in this experiment (three females; mean age: 21.40 years; range: 19–24 years). Participants signed informed consent forms and were paid for their participation. Procedures were conducted in accordance with the Declaration of Helsinki (except for preregistration) and were approved by the local ethics committee.

Apparatus, stimuli, and procedure

[Experiment 2](#) was identical to [Experiment 1](#) except for the following: the test target and flankers were all incomplete objects. Target and flankers' element types were manipulated independently in separate blocks, creating two similarity conditions. That is, a corners-target was surrounded by corners-flankers (i.e., similar, [Figure 4A](#)), or sides-flankers (i.e., dissimilar, [Figure 4B](#)), and a sides-target was surrounded by corners-flankers (i.e., dissimilar, [Figure 4C](#)) or sides-flankers (i.e., similar, [Figure 4D](#)). Only the flankers were subjected to the gap size manipulation.

The gap size of the test target was fixed at 0.5. The four conditions were presented randomly, each composed of 18 adjustments (9 gap sizes \times 2 adjustments trials). Example stimuli in [Experiment 2](#) with minimum (0.1) and maximum (0.9) gap size can be found in the Supplementary File (Supplementary Figure S2).

Results and discussion

Illusion magnitude was subjected to a three-way repeated measures ANOVA with gap size, target element type, and similarity as within-subject factor. The analysis revealed a significant main effect of gap size [$F(8, 72) = 11.98, p < 0.001, \eta_p^2 = 0.57$], showing a decrease in illusion magnitude with increasing gap size, a significant main effect of target element type [$F(1, 9) = 7.34, p = 0.02, \eta_p^2 = 0.45$], showing stronger illusion for sides-target ($M = 0.1, SD = 0.09$) compared with corners-target ($M = 0.07, SD = 0.09$). The effect of similarity was only marginally significant [$F(1, 9) = 3.82, p = 0.08, \eta_p^2 = 0.30$], showing stronger illusion for similar ($M = 0.09, SD = 0.09$) compared with dissimilar ($M = 0.08, SD = 0.09$) target and flankers' element type. This trend seems to come from some differences between similar and dissimilar condition at the largest gap sizes. None of the other effects reached significance ($p > 0.23$). The plots at the bottom of [Figure 4](#) depicts the illusion magnitude in each of the four conditions (left), and for the similarity conditions, for convenience (right).

The results of [Experiment 2](#) give further support to the hypothesis that objecthood affects the Ebbinghaus illusion, as a decrease in illusion magnitude was found with increasing gap size. The higher magnitude shown with sides-target compared with corners-target is in accordance with the results of [Experiment 1](#), suggesting that this type of target is more fragile. The effect of similarity was only marginally significant, showing no difference between conditions for most of the gap sizes. This result suggests that similarity in configuration between target and flankers has little to do with the illusion magnitude, at least in the displays used in this experiment.

Experiment 3

In contrast to previous studies, our results so far point to shape having no role in the Ebbinghaus illusion. To examine this hypothesis further, in [Experiment 3](#) we tested the effect of similarity in global shape between target and flankers by presenting a square target surrounded by cross-like flankers. Similar to the square flankers in the previous experiments, the cross flankers were constructed from straight lines or corners elements. Flankers' element type was manipulated to keep the experimental conditions as similar as possible

to those of [Experiments 1](#) and [2](#). We did not expect this factor to have an effect on the illusion, as the different types of elements did not convey different shape information when constructing the crosses in our displays. On the other hand, a cross shape is considerably different from that of a square. Hence, if similarity in shape plays a role, illusion magnitude should be weaker for dissimilar flankers. That is, illusion magnitude should be lower in this experiment compared with the previous experiment.

As in [Experiment 1](#), we collected subjective reports to get a more comprehensive account of objecthood and shape perception for the stimuli employed in our displays. Since reports of the square targets were already collected in [Experiment 1](#), in this experiment we collected reports only regarding the cross-like flankers.

Method

Participants

Ten naïve students, new to the study, from the École Polytechnique Fédérale de Lausanne (EPFL) participated in this experiment (three females; mean age: 23.40 years; range: 20–26 years). Participants signed informed consent forms and were paid for their participation. Procedures were conducted in accordance with the Declaration of Helsinki (except for preregistration) and were approved by the local ethics committee.

Apparatus, stimuli, and procedure

Adjustment task

[Experiment 3](#) was identical to [Experiment 2](#) except for the following: the test target was an incomplete square and the flankers were crosses. Crosses were constructed from corners segments oriented outwards—“*corners-flankers*” ([Figures 5A](#) and [5C](#)) or line segments positioned to form the crosses' arms—“*arms-flankers*” ([Figures 5B](#) and [5D](#)). Target and flankers' element types were manipulated independently. Only the flankers were subjected to the gap size manipulation, whereas the gap size of the test target was fixed at 0.5. A baseline condition with complete crosses (i.e., gap size 0) was included as well. There were 76 trials in total. Example stimuli in [Experiment 3](#) with minimum (0.1) and maximum (0.9) gap size can be found in the Supplementary File (Supplementary Figure S3).

Subjective reports

After completing the adjustment task, participants were asked to indicate whether they perceived the elements of the grouped flankers as: (a) a square, (b) a cross, (c) another shape, or (d) unrelated elements, for each gap size.

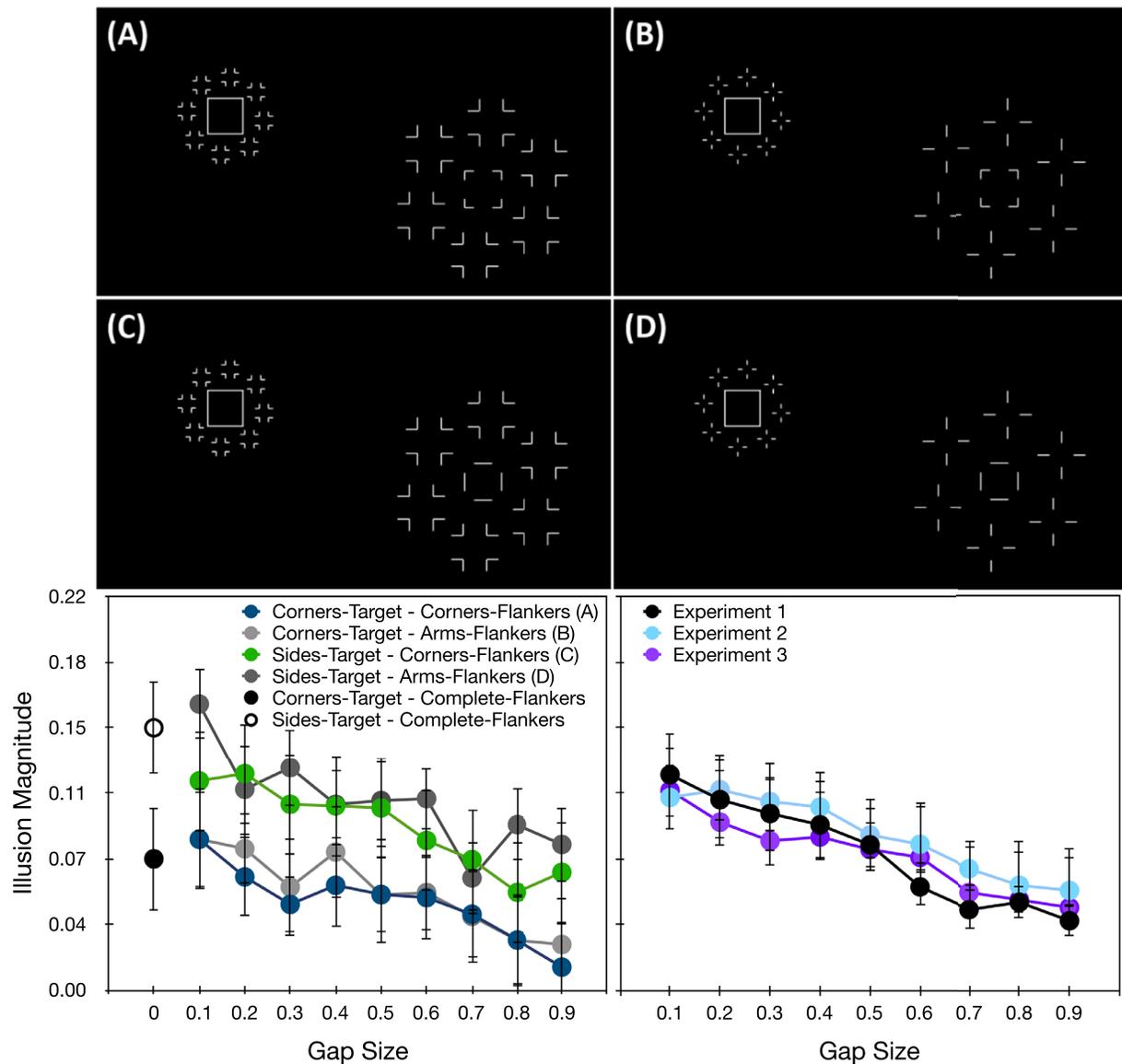


Figure 5. Top: examples of stimuli used in [Experiment 3](#). Targets were squares flanked by crosses. (A) Corners-target and corners-flankers. (B) Corners-target and arms-flankers. (C) Sides-target and corners-flankers. (D) Sides-target and arms-flankers. Gap size is 0.5 in these examples. The reference target was always a complete square. Bottom: mean illusion magnitude for different target and flankers part type conditions as a function of gap size in [Experiment 3](#) (left), and gap-size effect in all three experiments (right). Error bars indicate standard errors of the means.

Results and discussion

Adjustment task

Illusion magnitude was subjected to a three-way repeated measures ANOVA with gap size, target element type, and flankers' element type as within-subject factor. The analysis revealed a significant main effect of gap size [$F(8, 72) = 10.12, p < 0.001, \eta_p^2 = 0.53$], showing a decrease in illusion magnitude with increasing gap size, and a main effect of target element type [$F(1, 9) = 12.85, p = 0.006, \eta_p^2 = 0.59$], showing a higher magnitude for sides-target ($M = 0.1, SD =$

0.08) compared with corners-target ($M = 0.05, SD = 0.07$). The effect of flanker element type did not reach significance [$F(8, 72) = 2.04, p = 0.19, \eta_p^2 = 0.19$], nor did any of the interactions [$F_s < 1$]. Thus, similar to the results of the previous experiments, [Experiment 3](#) showed a reduction in illusion magnitude with decreased objecthood of the flankers ([Figure 5](#)).

To test whether the illusion is affected by shape similarity between target and flankers, we compared illusion magnitudes in [Experiments 2](#) and [3](#). We conducted a mixed-design repeated measures ANOVA with gap size as within-subject factor and experiment as between-subjects factor. The analysis revealed a

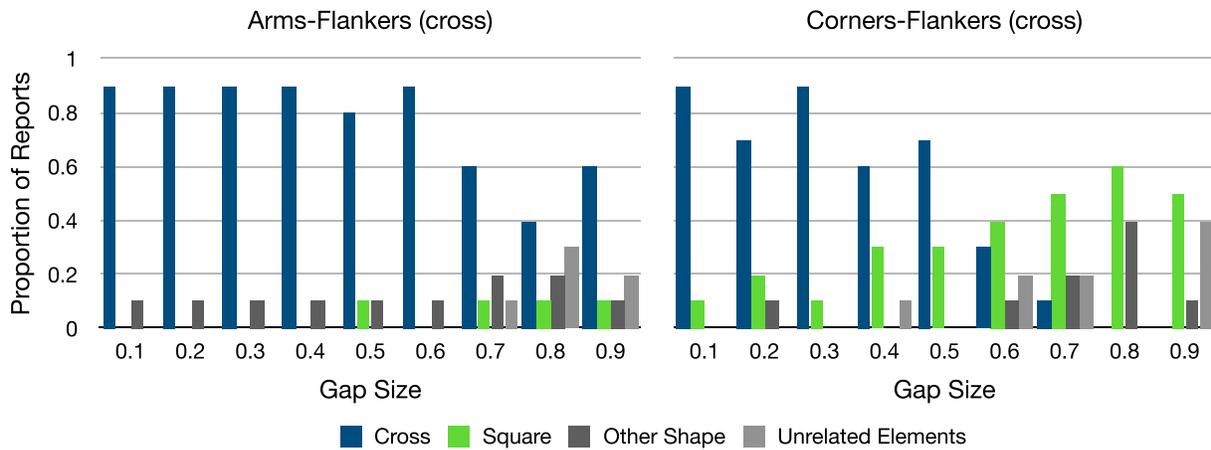


Figure 6. Proportion of reports at each gap size in the different element type conditions in Experiment 3.

significant main effect of gap size [$F(8, 144) = 21.28$, $p < 0.001$, $\eta_p^2 = 0.54$], showing decreasing magnitude with increasing gap size. The main effect of experiment was not significant, as was the interaction between experiment and gap-size [$F_s < 1$]. Thus the comparison between Experiments 2 and 3 shows that the objecthood effect was similar across the different shape similarity manipulations (Figure 5). Thus similarity of shape between target and flankers, global or configurational, does not seem to contribute to the objecthood effect on the illusion.

Subjective reports

Pearson's chi-squared tests were conducted on the distribution of reports from the two flankers' conditions (Figure 6). It was found that reports were correlated with flankers' element type [$\chi^2(3) = 27.25$, $p < 0.0001$]. Specifically, arms-flankers were perceived as crosses in 77% of the trials, as squares in 4%, as another shape in 12%, and as unrelated elements in 7%. On the other hand, corners-flankers were perceived as crosses in 47% of the trials, as squares in 33%, as another shape in 10%, and as unrelated elements in 10%. As clearly evident in the plots, the percentage of "cross" reports decreased with increasing gap size, whereas the percentage of the alternative reports increased, and this trend was more pronounced for corners-flankers compared with arms-flankers. As in Experiment 1, the subjective reports did not match the results of the adjustment task, suggesting that perceived shape does not contribute to the illusion.

Adjustment bias

Similar to Experiments 1 and 2, a main effect of target element type was found. Also, in the baseline conditions, where flankers were complete crosses, a significantly higher illusion magnitude was found for

sides-target ($M = 0.15$, $SD = 0.09$) compared with corners-target ($M = 0.07$, $SD = 0.08$), ($t = 2.91$, $p = 0.02$, two-tailed). This pattern suggests an adjustment bias. To test this possibility, we conducted a control experiment in which 10 new participants (three females; mean age: 21.30 years; range: 19–25 years) were presented with displays containing test and reference targets but no flankers. They were asked to perform the adjustment task, comparing a test target that was either a full square, a square made of corners, a square made of sides, or a cross made of corners to a complete square as a reference target. These grouped targets were presented for two trials each in a random order. The results of this control experiment revealed that illusion magnitude was significantly different from zero for target-sides ($M = 0.07$, $SD = 0.09$, $t = 2.41$, $p = 0.04$, two-tailed), but not for the others ($t \leq 1$). This result matches the difference between the two target types in our experiments, supporting also the idea that a bias in adjustment was underlying the odd results in the sides-target condition in Experiment 1.

Experiment 4

The gap-size effect found in Experiments 1 to 3 supports the hypothesis that the Ebbinghaus illusion is sensitive to the quality of the objects in the comparison. However, it is possible that the gap-size manipulation introduced another factor that contributed to the effect. Specifically, reducing the same proportion of contour from small and large flankers left different amounts of visible contours in each configuration, because the amount of contour for the targets was fixed. As in our paradigm it is not possible to disentangle the effect of small flankers from that of the large flankers on illusion magnitude, it is possible that the varying ratio in amount of visible contour led to the gap-size

effect. In this case a low-level contour interaction could explain the effect. To rule out this possibility, in [Experiment 4](#) we manipulated objecthood in a way that kept both ratio and amount of visible contour constant in both configurations—instead of disrupting the objects by discarding parts of their contour, we introduced a rotation manipulation to the elements forming the objects. We hypothesized that the illusion magnitude would decrease with increasing rotation angle, because of reduced objecthood of the flankers. Subjective reports were collected as well to account for the perceived shape of the flankers with rotated parts.

Methods

Participants

Ten naïve students, new to the study, from the École Polytechnique Fédérale de Lausanne (EPFL) participated in this experiment (four females; mean age: 22 years; range: 18–25 years). Participants signed informed consent forms and were paid for their participation. Procedures were conducted in accordance with the Declaration of Helsinki (except for preregistration) and were approved by the local ethics committee.

Apparatus, stimuli, and procedure

Adjustment task

[Experiment 4](#) was identical to the previous experiments except for the following: the test target was always a complete square and the flankers were four-element objects constructed of corners with a 0.5 gap size. The corners were rotated in two opposite directions (i.e., outward rotation; [Figure 7A](#)), or in the same direction (i.e., clockwise rotation; [Figure 7B](#)). Rotation angle was between 5° and 40° in steps of 5°. This range was chosen to make sure that no incidental collinearity grouping occurred between pairs of rotated elements while allowing a similar number of trials per condition as in the previous experiments. A baseline condition, where the corners were not rotated (i.e., similar to the stimuli in the previous experiments) was included as well. There were 34 trials in total. Example stimuli in [Experiment 4](#) with minimum (5°) and maximum (40°) rotation angle can be found in the Supplementary File (Supplementary Figure S4).

Subjective reports

After completing the adjustment task, participants were asked to indicate whether they perceived the elements of the grouped flankers as: (a) a square, (b)

another shape, or (c) unrelated elements, for each rotation angle.

Results and discussion

Adjustment task

Illusion magnitude was subjected to a two-way repeated measures ANOVA with degree of rotation and flankers' part rotation direction as within-subject factor. The analysis revealed a significant main effect of degree of rotation [$F(7, 63) = 6.25, p < 0.001, \eta_p^2 = 0.41$], showing a decrease in illusion magnitude with increasing rotation angle. There was no significant effect of rotation direction, nor an interaction between the two factors [$F_s < 1$]. Thus, similar to the results of the previous experiments, [Experiment 4](#) showed a reduction in illusion magnitude with decreased objecthood of the flankers ([Figure 7](#)).

Subjective reports

Pearson's chi-squared tests were conducted on the distribution of reports from the two flankers' conditions ([Figure 8](#)). It was found that reports were correlated with flankers' part rotation direction [$\chi^2(2) = 6.01, p < 0.04$]. Specifically, when the parts were rotated clockwise, the objects were perceived as squares in 36% of the trials, as another shape in 36%, and as unrelated elements in 18%. On the other hand, when the parts were rotated outward, the objects were perceived as squares in 21% of the trials, as another shape in 21%, and as unrelated elements in 29%. As clearly evident in the plots, the percentage of "square" reports decreased with increasing rotation angle, while the percentage of the alternative reports increased, and this trend was more pronounced with outward rotation compared with clockwise rotation. As in [Experiments 1](#) and [3](#), the subjective reports do not seem to reflect a factor that affects the adjustment task, suggesting that perceived shape does not contribute to the illusion.

General discussion

The main goal of this study was to examine whether objecthood has a role in the Ebbinghaus illusion. The results of our experiments demonstrate that degrading the quality of an object by means of grouping strength between the elements constructing the object affects size estimation. In line with our hypothesis, we found that reduced grouping strength led to a weaker illusion. Specifically, increasing the gap size ([Experiments 1–3](#)), or rotation angle ([Experiment 4](#)) between the objects' elements resulted in a decreased illusion

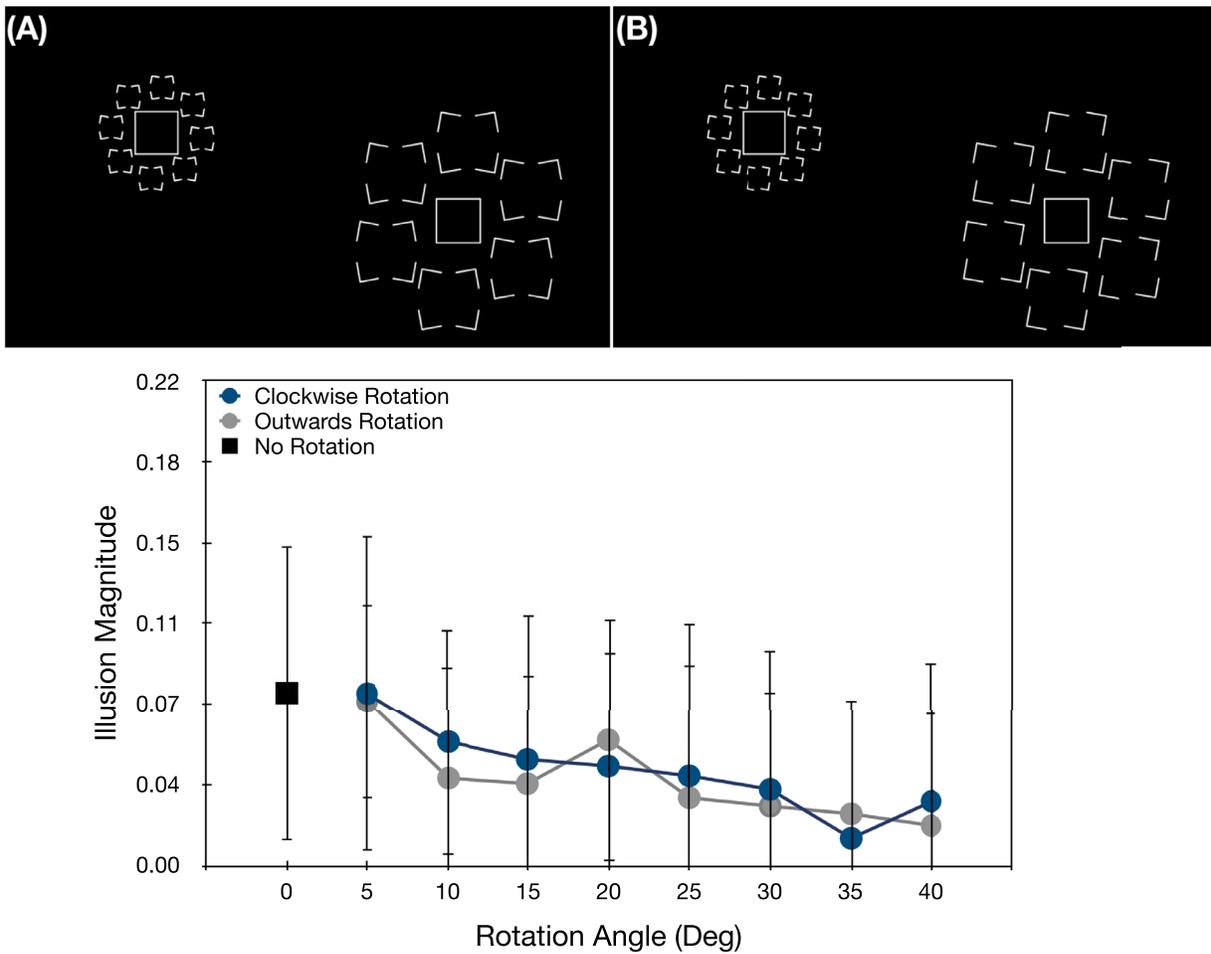


Figure 7. Top: examples of stimuli employed in Experiment 4. Test and reference targets were complete squares. Flankers were made of (A) corners rotated outwards (i.e., opposite directions), or (B) clockwise (i.e., same direction). Rotation is 10° in these examples. Bottom: mean illusion magnitude for different flankers' part rotation direction conditions as a function of degree of rotation in Experiment 4. Error bars indicate standard errors of the means.

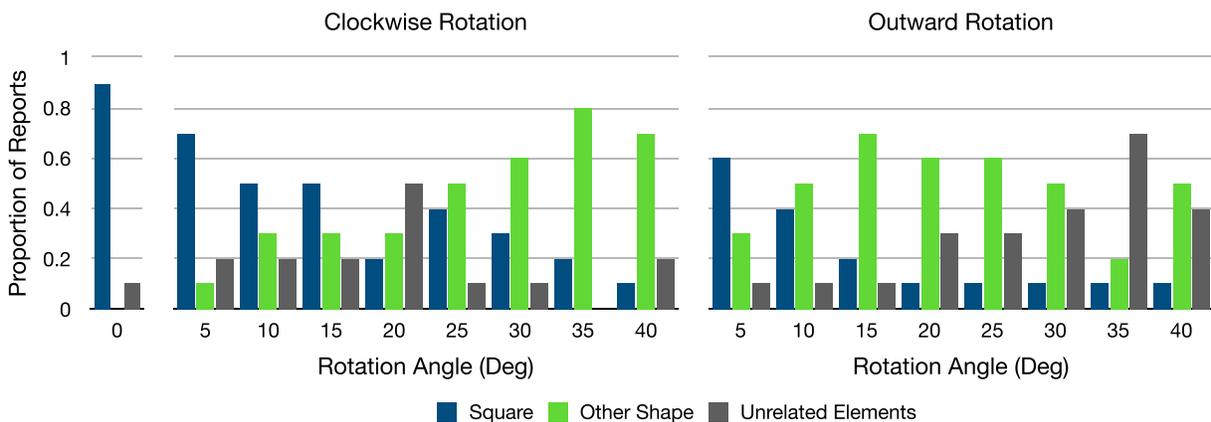


Figure 8. Proportion of reports at each rotation angle in the different rotation direction conditions in Experiment 4.

magnitude. This result is consistent with the size contrast account, which posits that the illusion is a product of a comparative mechanism rather than low-level contour interactions. This is because as the objecthood of the flankers or target was reduced, the less effective they were in the comparison. The contour interaction account is less likely to explain our results because the ratio between close and far contours of the flankers relative to the target was fixed during the objecthood manipulation. If the target was subjected to attraction and repulsion from the flankers, illusion magnitude should have been the same for all gap sizes and rotation angles. [Experiment 4](#) challenged also an alternative explanation that the objecthood effect found in [Experiments 1–3](#) was due to a difference in contour ratio between target and flankers in the test and reference, as this ratio was kept constant in this experiment. It is important to note, however, that the contour interaction account is not developed enough to provide clear predictions ([Todorović and Jovanović, 2018](#)). It is possible that close and far contours are weighted differently depending on some attribute of the flankers, or that the point of reference for far-close contour ratio is other than the center of the object (e.g., the distance between nearest points of flankers and targets). However, our finding of a similar effect with two different manipulations of the visible contours surrounding the target cannot be parsimoniously accommodated by the alternative account. Thus we suggest that our results provide evidence of the role of mid-level processes in the illusion, supporting the size contrast account.

Another objective of this study was to address the question of the role of shape similarity in the illusion, as previous studies showed conflicting results ([Choplin & Medin, 1999](#); [Coren & Miller, 1974](#); [Vuk & Podlessek, 2005](#)). Importantly, we found no evidence for configural similarity, because there was no difference in illusion magnitude, nor in the objecthood effect (i.e., the effect of gap size on illusion magnitude), when target and flankers were composed of similar or different element configurations ([Experiment 2](#)), or when they had the same or different global shapes ([Experiments 2 and 3](#), respectively). Furthermore, measuring the perceived shapes of the objects in our displays show no correspondence with the size error measure of the illusion. That is, the gap size manipulation reduced the perceived “squareness” (or “cross-ness”) of the flankers, but this effect differed for the two element types used to construct them, which was a factor that did not affect the illusion magnitude. The latter may indicate a dissociation between mechanisms involved in phenomenology (e.g., subjective reports) and in visuo-motor integration (e.g., adjustment task). For example, [Aglioti, DeSouza, and Goodale \(1995\)](#) demonstrated a dissociation between perception and action towards an Ebbinghaus display; as the

illusion was obtained with perceptual judgements, however, grasping the test target was not affected by the surrounding stimuli. A dissociation between direct and indirect measures of perceptual organization has been reported by [Schmidt and Schmidt \(2013\)](#), who demonstrated that grouping strength of different grouping rules (e.g., similarity by shape and similarity by brightness) did not correlate with priming effects of those rules. Hence, we propose that the objecthood effects reported here and shape similarity effects in previous studies of the Ebbinghaus illusion reflect an effect of figural “goodness” ([van der Helm, 2014](#); for a comprehensive review on Gestalt factors in visual perception see [Peterson & Kimchi, 2013](#); [Wagemans et al., 2012](#)). Support to this hypothesis may be found in a study by [Rose and Bressan \(2002\)](#), who showed that even when the target and flankers’ shapes were identical, the magnitude of the illusion varied for different shapes. Because some shapes are better than others, their “goodness” could have caused the variations in illusion magnitudes. This is a speculation at this point, since “goodness” was not measured in that study, but it poses an interesting question that merits further investigation.

Interestingly, the adjustment bias we found for sides-target suggests that shape information is crucial for the adjustment procedure. When the test target contains less shape related information, there is difficulty completing it into an imaginary square that is comparable with the reference square. As suggested by the subjective reports, the visible segments can be completed into another shape, for example an octagon, which has a smaller surface. Thus, the size of this shape would result in an adjustment bias to equal that of the square regardless of the surrounding objects. The effect found in [Experiment 1](#) for the sides-target is compatible with this idea: the surface of the target became smaller with a decrease in visible contour, leading to a larger adjustment error with each increment in gap size. This pattern is not evident with the corners-target, because corners convey more shape information (e.g., [Hadad & Kimchi, 2008](#); [Persike & Meinhardt, 2017](#); [Poirier & Wilson, 2007](#)), thus, providing a target that is a comparable square to the reference.

Our results accord with a recent study by [Lavrenteva and Murakami \(2018\)](#). In their study, the target and flankers were defined by first- and second-order attributes (e.g., luminance and local contrast, respectively). Interestingly, an asymmetry in the size estimation error was found, showing that first-order flankers had the same effect on size estimation with all target types, whereas second-order flankers affected first-order targets less than second-order targets. The authors proposed that this asymmetry resulted from different weights given to first- and second-order attributes when target and flankers differed in the sensory evidence they provided. The objecthood effect

found in the current study can be another example of changing weights. That is, a degraded object weighs less than a good object since it contains less coherent information, and thus, it has less of an effect in the size contrast process.

Lastly, it is important to acknowledge that the size contrast account has its own shortcoming, specifically, it being more of a description of the phenomenon rather than an explanation (Todorović and Jovanović, 2018). However, the size contrast account poses that object-level representations are compared. Thus the degree of objecthood would be crucially influencing the process, as the object as a whole is compared, whether it has a complete or grouped outline. A few results from previous studies (e.g., Weintraub & Schneck, 1986) showing the illusion with only dots or lines as flankers are not easily explained with just objecthood. Hence, an interesting line of future research would examine other processes of perceptual organization, such as grouping and segmentation of the target and flankers, which may be involved in such complex displays of the illusion in addition to size contrast, contour interactions and their potential interactions.

Conclusion

Objecthood plays a role in the size adjustment error demonstrated in the Ebbinghaus illusion. The current study provides support for a higher-level size contrast mechanism underlying the illusion and proposes that better objects carry more weight during the comparison process. This opens a new avenue of studying objecthood through visual illusions.

Keywords: objecthood, grouping, size illusion

Acknowledgments

This project has received funding from a grant to MHH, “Basics of visual processing: from elements to figures” (project no. 320030_176153/1) of the Swiss National Science Foundation (SNSF). ER was supported by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 708007.

Commercial relationships: none.
Corresponding author: Einat Rashal.
Email: einatrashal@gmail.com.
Address: École Polytechnique Fédérale de Lausanne (EPFL), Brain Mind Institute, Laboratory of Psychophysics, Station 19, 1015 Lausanne, Switzerland.

Footnote

¹We consider the center of the flankers to be the relative point from which attraction and repulsion are operating by close and far contours, respectively. Although this has not been made explicit in previous studies, this seems to be the underlying assumption. Theoretically, attraction and repulsion may operate within another function, however, it is not clear what it might be (e.g., Todorovic and Jovanovic, 2018).

References

- Aglioti, S., DeSouza, J. F., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology*, 5, 679–685.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- Choplin, J. M., & Medin, D. L. (1999). Similarity of the perimeters in the Ebbinghaus illusion. *Perception & Psychophysics*, 61, 3–12.
- Coren, S., & Miller, J. (1974). Size contrast as a function of figural similarity. *Attention, Perception, & Psychophysics*, 16(2), 355–357.
- Coren, S., & Enns, J. T. (1993). Size contrast as a function of conceptual similarity between test and inducers. *Perception & Psychophysics*, 54, 579–588.
- Cretenoud, A. F., Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, 19(14), 12–12.
- Hadad, B., & Kimchi, R. (2008). Time course of grouping of shape by perceptual closure: Effects of spatial proximity and collinearity. *Perception and Psychophysics*, 70(5), 818–827.
- Jaeger, T. (1978) Ebbinghaus illusions: size contrast or contour interaction phenomena? *Perception and Psychophysics*, 24, 337–342.
- Jaeger, T., & Klahs, K. (2015). The Ebbinghaus illusion: New contextual effects and theoretical considerations. *Perceptual and Motor Skills*, 120(1), 177–182.
- Kimchi, R., Yeshurun, Y., Spehar, B., & Pirkner, Y. (2015). Perceptual organization, visual attention, and objecthood. *Vision Research*, 126, 34–51.
- Lavrenteva, S., & Murakami, I. (2018). The Ebbinghaus illusion in contrast-defined and orientation-defined stimuli. *Vision Research*, 148, 26–36.
- Massaro, D. W., & Anderson, N. H. (1971). Judgmental model of the Ebbinghaus illusion. *Journal of Experimental Psychology*, 89(1), 147.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10, 437–442.

- Persike, M., & Meinhardt, G. (2017). A new angle on contour integration: The role of corners. *Journal of Vision*, 17(12):7, 1–13.
- Peterson, M. A., & Kimchi, R. (2013). Perceptual organization in vision. In D. Reisberg (Ed.), *Oxford Handbook of Cognitive Psychology* (pp. 9–31). New York: Oxford University Press.
- Poirier, F. J. A. M., & Wilson, H. R. (2007). Object perception and masking: Contributions of sides and convexities. *Vision Research*, 47, 3001–3011.
- Rose, D., & Bressan, P. (2002). Going round in circles: Shape effects in the Ebbinghaus illusion. *Spatial Vision*, 15(2), 191–203.
- Schmidt, F., & Schmidt, T. (2013). Grouping principles in direct competition. *Vision Research*, 88, 9–21.
- Schwarzkopf, D. S., & Rees, G. (2013). Subjective size perception depends on central visual cortical magnification in human V1. *PLoS One*, 8(3), e60550.
- Sherman, J. A., & Chouinard, P. A. (2016). Attractive contours of the Ebbinghaus illusion. *Perceptual and Motor Skills*, 122(1), 88–95.
- Todorović, D., & Jovanović, L. (2018). Is the Ebbinghaus illusion a size contrast illusion? *Acta Psychologica*, 185, 180–187, doi:10.1016/j.actpsy.2018.02.011.
- van der Helm, P. A. (2014). Simplicity in perceptual organization. In J. Wagemans (Ed.), *Oxford handbook of perceptual organization*. Oxford, U.K.: Oxford University Press.
- Vuk, M., & Podlesek, A. (2005). The effect of inner elements of the context figures on the Ebbinghaus illusion. *Horizons of Psychology*, 14, 9–22.
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., . . . von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin*, 138, 1172–1217.
- Weintraub, D. J., & Schneck, M. K. (1986). Fragments of Delboeuf and Ebbinghaus illusions: Contour/context explorations of misjudged circle size. *Attention, Perception, & Psychophysics*, 40(3), 147–158.

**Perceptual grouping leads to objecthood effects in the Ebbinghaus illusion –
Supplementary File**

Einat Rashal, Aline F. Cretenoud, & Michael H. Herzog

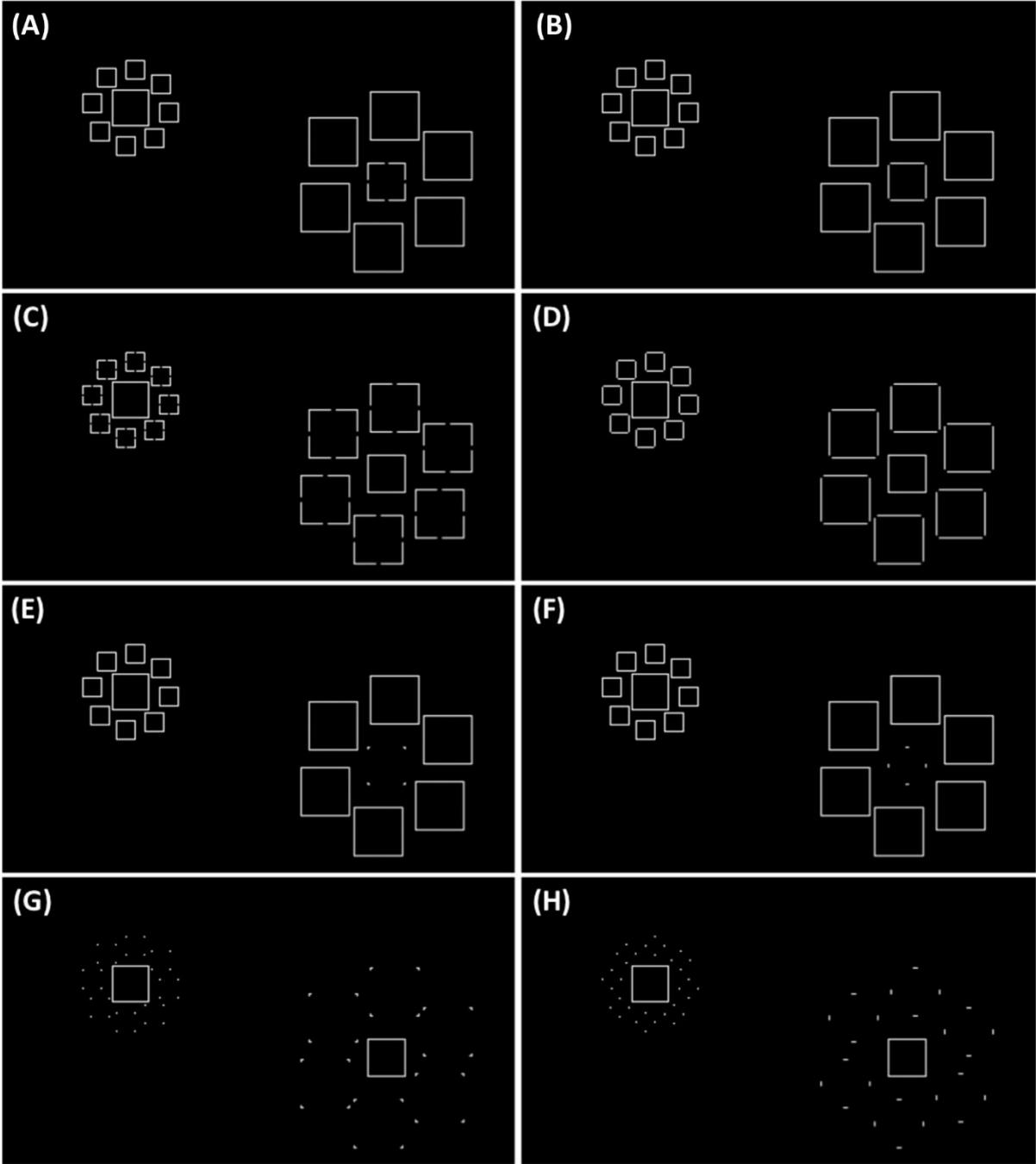


Figure S1. Examples of stimuli used in Experiment 1 with gap sizes of 0.1 (A-D) and 0.9 (E-H): (A/E) Corners-target, (B/F) Sides-target, (C/G) Corners-flankers, and (D/H) Sides-flankers.

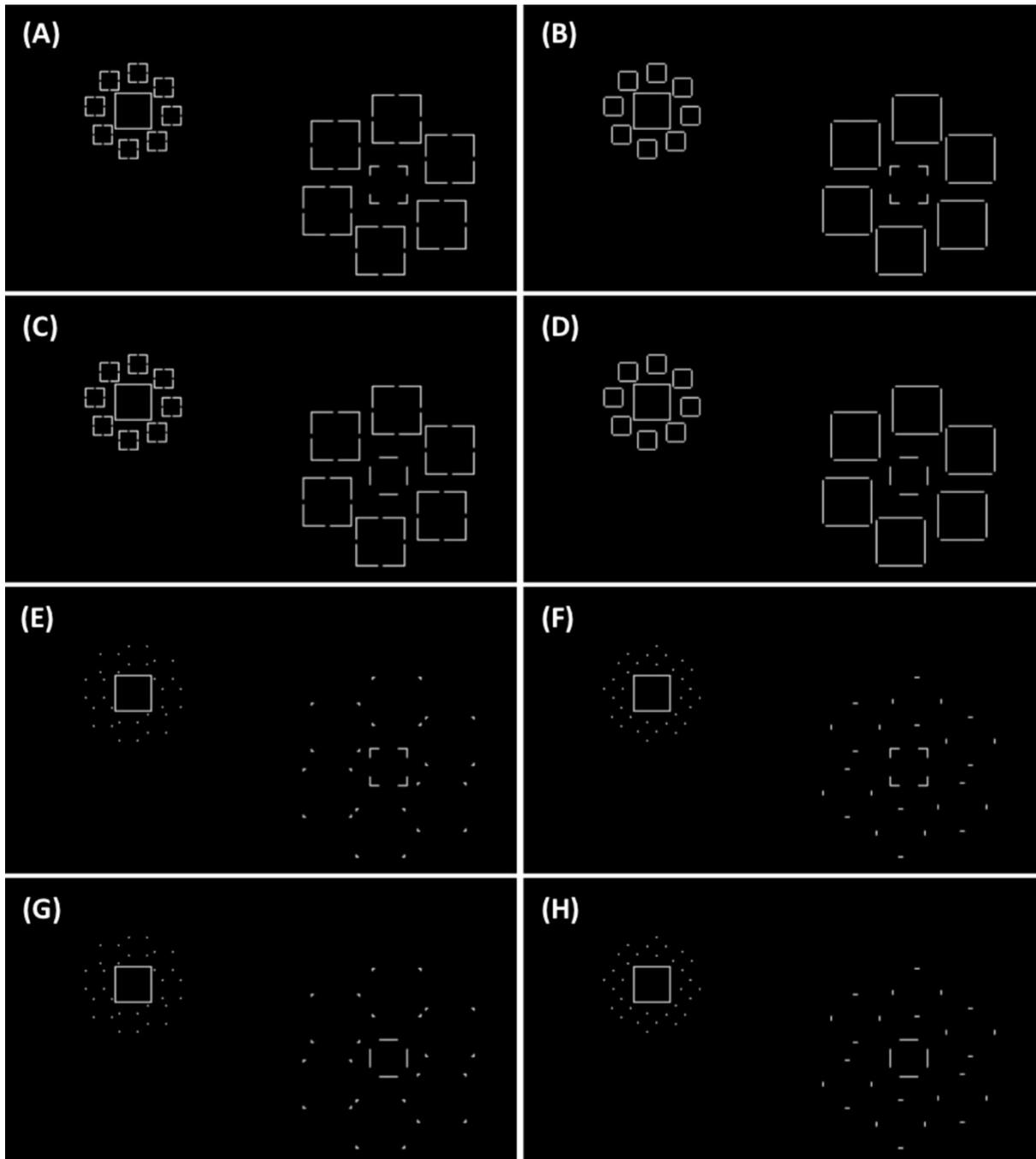


Figure S2. Examples of stimuli used in Experiment 2 with gap sizes of 0.1 (A-D) and 0.9 (E-H): (A/E) Similar: corners-target and corners-flankers. (B/F) Dissimilar: corners-target and sides-flankers. (C/G) Dissimilar: sides-target and corners-flankers. (D/H) Similar: sides-target and sides-flankers. The gap size of the test target was fixed at 0.5.

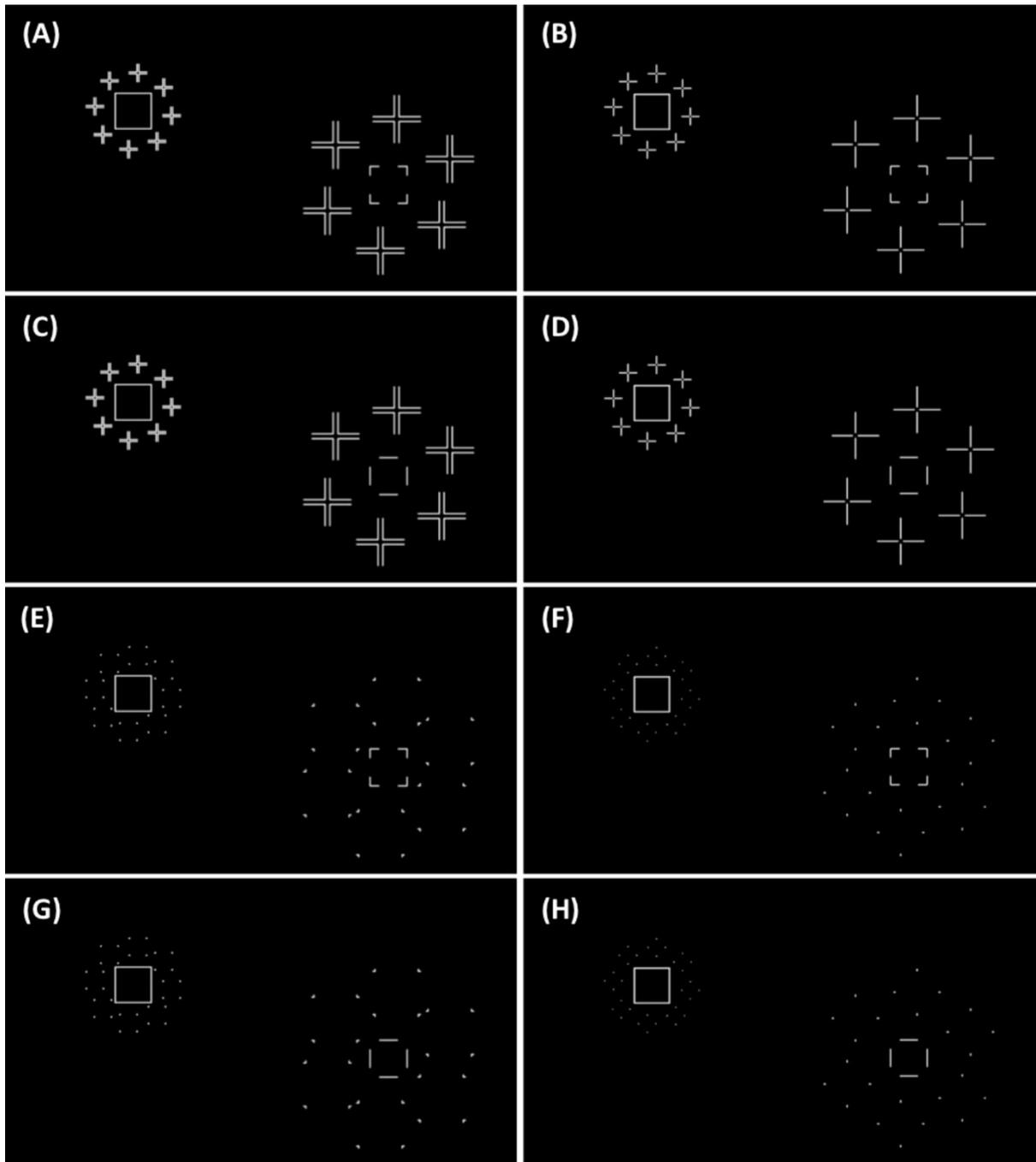


Figure S3. Examples of stimuli used in Experiment 3 with gap sizes of 0.1 (A-D) and 0.9 (E-H): (A/E) Corners-target and corners-flankers, (B/F) Corners-target and arms-flankers, (C/G) Sides-target and corners-flankers, and (D/H) Sides-target and arms-flankers. The gap size of the test target was fixed at 0.5.

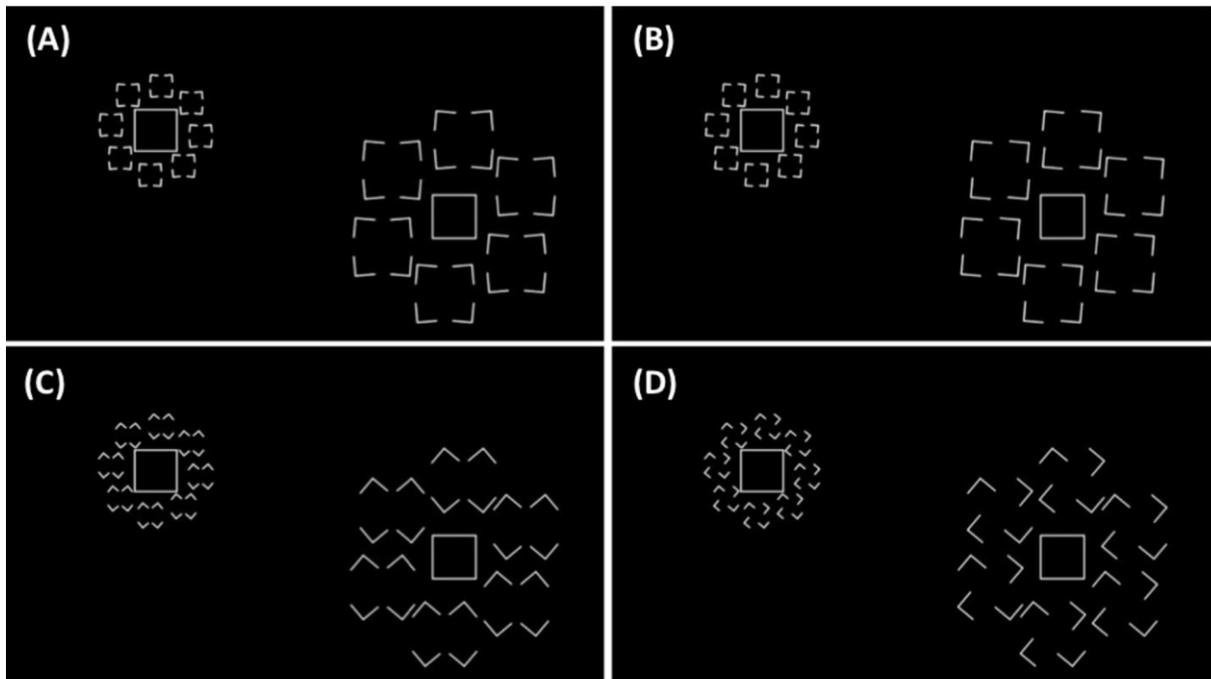


Figure S4. Examples of stimuli used in Experiment 4 with angle rotation of 5 (A-B) and 40 (C-D) degrees: Flankers were made of (A/C) corners rotated outwards (i.e., opposite directions), or (B/D) clockwise (i.e., same direction). Test and reference targets were complete squares.

Appendix H¹¹

Jastrzębowska, M. A., Cretenoud, A. F., Ozkirli, A., Draganski, B., Herzog, M. H. (ongoing). Is there a neural common factor in illusion magnitude?

¹¹ This project has previously been described in Jastrzębowska, M. A. (2020). *Bayesian modeling of brain function and behavior*. PhD thesis, EPFL, Lausanne, Switzerland. Available online: <http://dx.doi.org/10.5075/epfl-thesis-7388>

Introduction

Idiosyncrasies in the visual cortex were recently suggested to be associated with individual variability in the perception of visual illusions (Chen et al., 2020; Moutsiana et al., 2016; Pooresmaeili et al., 2013; Schwarzkopf et al., 2011; Schwarzkopf & Rees, 2013). For example, Schwarzkopf and colleagues (2011, 2013) reported negative associations between the primary visual cortex (V1) surface area and the susceptibility to the Ebbinghaus and corridor Ponzo illusions. Even though V1 surface area was previously reported to strongly correlate with the secondary visual cortex (V2) surface area (Dougherty et al., 2003), there was no significant correlation between the illusion magnitudes and V2 surface area (or higher areas, e.g., V3 and V4). In addition, a significant correlation was observed between the susceptibility to the Delboeuf illusion and population receptive field (pRF) size in V1 (Moutsiana et al., 2016).

Assuming that cortical properties are – at least partially – driving the individual differences in vision, a common factor for vision should emerge. However, as reported in the present thesis and elsewhere (for reviews, see Mollon et al., 2017; Tulver, 2019), visual abilities seem largely uncorrelated. The aim of the present study was to replicate previous results and examine whether they generalize to other illusions. In addition to testing the Ebbinghaus and corridor Ponzo illusions, which were reported to negatively correlate with V1 surface area in Schwarzkopf and colleagues (2011, 2013), other size illusions as well as a bunch of illusions of perceived orientation, uniform textures, and contrast, were tested. While it is not unlikely that perception of orientation and uniform textures is also mediated by cortical idiosyncrasies, there is no *a priori* reason to expect a significant correlation between early visual cortical area size and the susceptibility to a contrast illusion.

Methods

Participants

Thirty participants were tested (14 females, mean age: 25 years, range: 18-35 years). All participants had normal uncorrected visual acuity, as measured with the Freiburg visual acuity test (i.e., monocular and binocular acuity values > 1.0; Bach, 1996). Following the psychophysical experiment, participants underwent a magnetic resonance imaging (MRI) session. They were naïve to the purpose of the experiment and aware that they could quit the experiment at any time. Participants were paid 20 Swiss Francs per hour in the psychophysical session and 30 Swiss Francs per hour in the MRI session. Procedures were conducted in accordance with the Declaration of Helsinki, except for preregistration (§ 35), and were approved by the local ethics committee.

Stimuli and apparatus¹²

For the psychophysical experiment, visual stimuli were displayed on an ASUS VG278HE LCD monitor (refresh rate: 60 Hz; spatial resolution: 1920×1080) at a viewing distance of 76 cm. The mean luminance of the screen was 75 cd/m². Stimulus programs were implemented in MATLAB (The MathWorks, Inc., Natick, MA) using the Psychtoolbox (Brainard, 1997).

The MRI experiment was conducted at the MRI-scanning facilities of the Centre Hospitalier Universitaire Vaudois (CHUV) in Lausanne, Switzerland. Participants laid in the MRI scanner and looked at a screen, placed at the end of the 60 cm scanner bore, through a mirror (viewing distance: 70 cm). A Sony VPL-FH31 projector (size of projected image: 57×32.1 cm; resolution: 1920×1080 pixels; refresh rate: 60 Hz) was used to back-project the stimuli onto the screen, with a mean luminance of 1000 cd/m². As in the psychophysical part, stimulus generation and response collection were done in MATLAB using the Psychtoolbox.

Visual illusions

An adjustment procedure with the computer mouse was used to measure the susceptibility to 13 illusions (Figure 22): the bisection (BS), contrast (CS), Delboeuf (DB), two variants of the Ebbinghaus (EB1 and EB2), Extinction (EX), Honeycomb (HC), Müller-Lyer (ML), Poggendorff (PD), two variants of the Ponzo (PZ and PZc), Tilt (TT), and Zoellner (ZN) illusions. In the BS, CS, DB, EB1, EB2, ML, PD, PZ, PZc, TT, and ZN, one element was in turn the adjustable element or the reference. For example, in the ML illusion, the segment with inward-pointing arrows was adjusted to match the length of the segment with outward-pointing arrows, and inversely. The EX and HC illusions were tested with two contrast polarity conditions, i.e., either black or white barbs in the HC illusion and either black or white dots in the EX illusion. To validate a trial, participants clicked on the left button of the computer mouse. Each condition was tested twice in a random order (13 illusions × 2 conditions × 2 trials = 52). There was no time constraint and no feedback. Lines were shown with a 4-pixel width, except in the DB illusion, where the circles were 2 pixels wide. Black and white had a luminance of ≈ 1 cd/m² and ≈ 150 cd/m², respectively. Illusions are further described below¹².

¹² This was copied (with only minor changes) from Jastrzębowska, M. A. (2020). *Bayesian modeling of brain function and behavior*. PhD thesis, EPFL, Lausanne, Switzerland. Available online: <http://dx.doi.org/10.5075/epfl-thesis-7388>

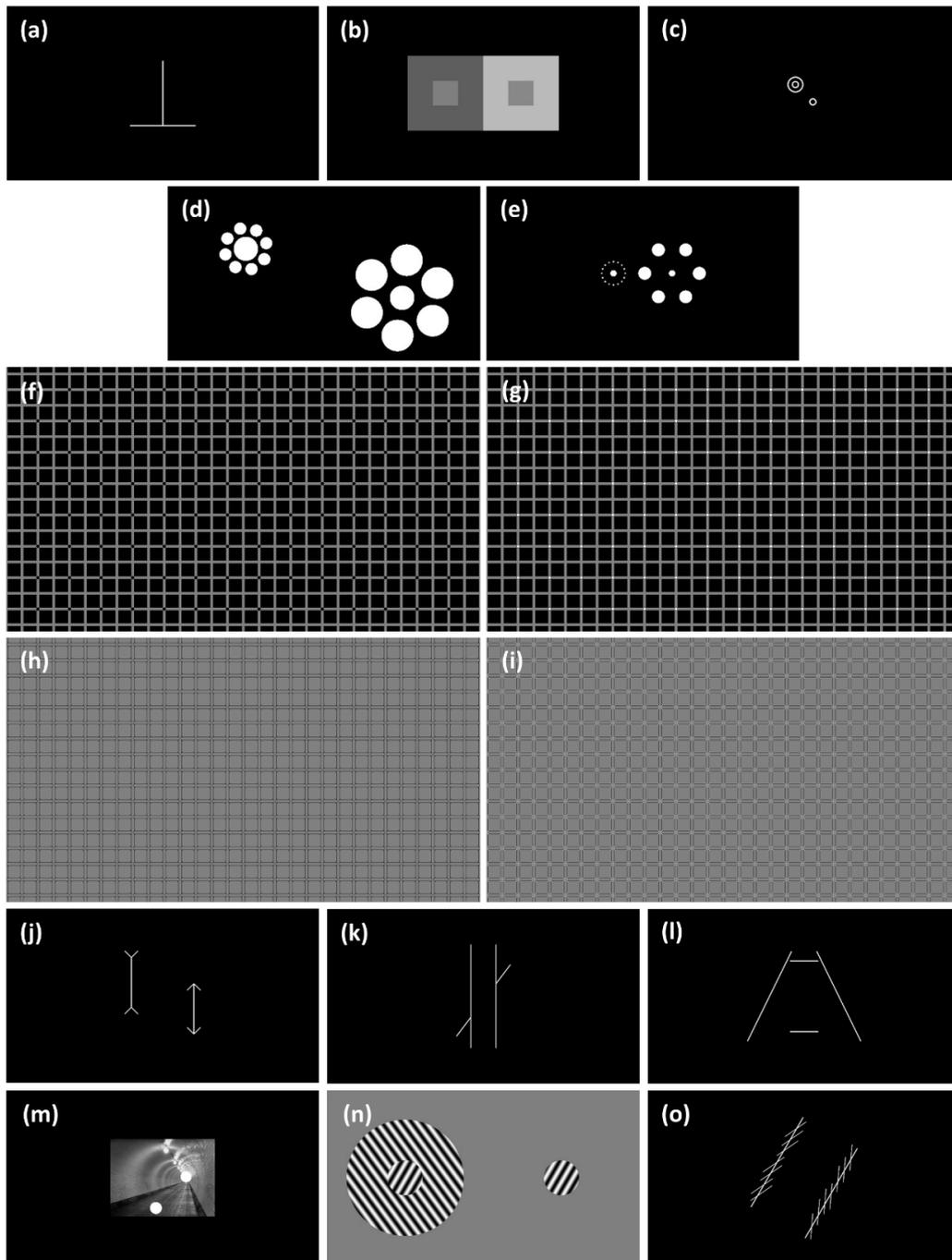


Figure 22. Appendix H. The bisection (a), contrast (b), Delboeuf (c; see Moutsiana et al., 2016), two variants of the Ebbinghaus (d and e; e was reproduced from Schwarzkopf & Rees, 2013), Extinction illusion with (f) black and (g) white dots, Honeycomb illusion with (h) black and (i) white barbs, Müller-Lyer (j), Poggendorff (k), Ponzo (l), corridor Ponzo (m; inspired from Schwarzkopf, Song, & Rees, 2011), Tilt (n), and Zoellner (o) illusions. The images (f) to (i) must be enlarged to see the barbs and dots. Note that the red adjustable ellipse is not represented in images (f) to (i); please refer to Bertamini et al. (2016) for further details. All illusions were tested with two conditions. For example, participants were asked to adjust the length of the vertical segment to match the length of the horizontal segment in the BS illusion, and inversely. Similarly, the Honeycomb and Extinction illusions were either tested with black or white barbs and dots, respectively. Reprinted from Jastrzębowska (2020).

Bisection (BS; also called vertical-horizontal or horizontal-vertical)

Participants had to adjust the length of the vertical segment so that it appeared to be the same length as the horizontal segment or to adjust the length of the horizontal segment to match the length of the vertical one (Figure 22a). The reference segment was 10.4° long and the adjustable segment had a size randomly set in the range of $2-17^\circ$ at the beginning of each trial. The horizontal segment was displayed 5.2° to the bottom of the midscreen and was always touching the vertical segment.

Contrast (CS)

Participants had to adjust the shade of grey of the left inside square so that they perceived it to be the same as the shade of grey of the right inside square, or vice versa (Figure 22b). The inside squares were 4° in side and displayed in the middle of two outside squares, which were 12° in side. The two outside squares were together displayed in the middle of the screen. The luminance of the left and right outside squares was approximately 17 and 75 cd/m^2 , while the reference inside square was $\approx 39 \text{ cd/m}^2$ in luminance.

Delboeuf (DB)

The stimulus was strongly inspired from Moutsiana et al. (2016), where the authors used a Multiple Alternative Perceptual Search (MAPS) procedure, in which participants had to report which of four comparison rings was the more similar in size to a central target. Here, participants had to adjust the size of the upper-left inside circle to match the size of the lower-right circle, or inversely (Figure 22c), by moving the computer mouse on the horizontal axis. The reference and upper-left outside circles were 0.98° and 2.35° in diameter, respectively. The size of the adjustable circle was randomly set at the beginning of each trial, but it never exceeded 2.35° in diameter. The center to center distance between the reference and adjustable circles was 3.92° and the whole illusion was centered in the middle of the screen.

Ebbinghaus (EB1 and EB2)

In both variants of the Ebbinghaus illusion, participants were instructed to adjust the size of the left central disk (i.e., left target) to match the size of the right central disk (i.e., right target), or vice versa.

The susceptibility to the Ebbinghaus illusion (EB1) was measured as previously (Cretenoud et al., 2019). The left and right targets were surrounded by eight small (2.0° in diameter) and six large (5.2° in diameter) flankers, respectively (Figure 22d). The reference target was 4° in diameter. The center of the small and large flankers was 3.36° and 6.08° away compared to the center of the left and right

targets, respectively. The left and right targets were centered 12.41° to the left and right and 4° to the top and bottom, respectively, compared to the center of the screen.

The second variant of the Ebbinghaus illusion (EB2) was strongly inspired from Schwarzkopf and Rees (2013). The left and right targets were surrounded by 16 small (0.26° in diameter) and six large (2.07° in diameter) flankers, respectively (Figure 22e). The reference target was 1.03° in diameter. The center of the small and large flankers was located 1.86° and 4.34° away from the center of the left and right targets, respectively. The left and right targets were located at 4.65° eccentricity compared to the middle of the screen and were vertically centered.

In both variants of the Ebbinghaus illusion, the size of the adjustable target was randomly chosen by the computer at the beginning of each trial with the constraint that it could never touch the flankers.

Extinction (EX) and Honeycomb (HC)

The stimuli were used as in Bertamini et al. (2019; see also Bertamini et al., 2016). In the Honeycomb (HC) and Extinction (EX) illusions, participants are unable to see shapes (barbs in the HC illusion; dots in the EX illusion) in the periphery of a uniform texture. Participants were asked to fixate a red central cross while adjusting the size of a red ellipse on both x and y dimensions, so that they could perceive all barbs (HC) and dots (EX) inside the ellipse. The initial size of the red ellipse was randomly chosen at the beginning of each trial, with the screen size as limits. Two contrast polarity conditions, i.e., black or white barbs (HC) or dots (EX), were tested, making up four conditions (Figure 22f-i). To reduce the aftereffect following a HC or EX trial, 30 random light and dark grey checkerboards (made of squares of 0.52° in side with 0.35 and 0.65 of the maximum luminance) were presented for 0.5 second each.

Müller-Lyer (ML)

Participants had to adjust the length of the segment with inward-pointing arrows so that they perceived it to be as long as the segment with outward-pointing arrows, or vice versa (Figure 22j). The reference segment was 8° long and the fins were 1.5° long, oriented at 45° compared to the vertical. At the beginning of each trial, the length of the adjustable segment was randomly set between 2 and 21° . The left and right segments were centered 2.12° to the top and bottom and 4.97° to the left and right compared to the middle of the screen, respectively.

Poggendorff (PD)

Participants were instructed to adjust the vertical position of the left or right interrupted diagonal so that it appeared to lie on a continuum with the right or left interrupted part, respectively (Figure 22k).

The two main vertical streams were 16.6° long and 4° away from each other. Both parts of the interrupted diagonal were 3.8° long and tilted by 37° compared to the vertical. When the position of the left part of the interrupted diagonal was adjusted, the right part was touching the right main stream 6.29 degrees away compared to the top of the main stream. When adjusting the position of the right part of the interrupted diagonal, the left part was touching the left main stream 5 degrees away compared to the bottom of the main stream. The adjustable element was randomly positioned along the corresponding main stream at the beginning of each trial and the whole illusion was centrally displayed.

Ponzo (PZ) and corridor Ponzo (PZc; also called corridor or hallway)

In the first variant of the Ponzo (PZ) illusion, the task was to adjust the length of the upper or lower horizontal segment to match the length of the lower or upper horizontal segment, respectively (Figure 22l). The reference segment was 4.5° long and the adjustable one was randomly set between 0 and 12° at the beginning of each trial. To induce a trapezoid-like perspective in the illusion, two converging lines were shown with 4° separating them at the apex and 18° at the base. The total height of the imaginary trapezoid was 14.4° and the reference and adjustable segments were 11.3° away from each other. The whole illusion was centrally displayed.

Also called corridor illusion, the second variant of the Ponzo (PZc) illusion was inspired from Schwarzkopf et al. (2011). Participants were instructed to adjust the size of the lower-left disk so that it appeared to be the same size as the upper-right disk, or inversely (Figure 22m). A tunnel image (640 x 480 pixels) was displayed in the center of the screen on a black background. The reference disk was 1° in diameter and the size of the adjustable disk was set in the range of 0 to 2° in diameter at the beginning of each trial. The luminance of the reference and adjustable disks was 150 cd/m^2 . The adjustable disk was fixed at its lowest point, as if it was anchored to the image background.

Tilt (TT)

Participants had to adjust the orientation of the right disk to match the orientation of the left inside disk, or inversely (Figure 22n). The reference and adjustable disks were 6° in diameter and made of a $0.5 \text{ cycles/}^\circ$ full contrast grating texture. The reference disk was tilted 33° clockwise compared to the vertical and the adjustable disk was displayed with a random orientation at the beginning of each trial. The outside left disk ($0.5 \text{ cycles/}^\circ$ full contrast grating texture) was 20° in diameter and tilted 36° counterclockwise compared to the vertical. The center of the left and right disks was displayed 6.98° to the left and right compared to the middle of the screen. The background luminance was 75 cd/m^2 .

Zoellner (ZN)

Participants were instructed to adjust the orientation of one stream so that it appeared parallel to the other stream (Figure 22o). The streams were 16.6° long and 9.94° apart from each other. The reference stream was tilted by 30° compared to the vertical. At the beginning of each trial, the adjustable stream was randomly tilted between 0 and 90° compared to the vertical. Seven segments, 4.15° long, were intersecting each stream. They were tilted by 25° compared to the streams and their position relative to the stream was randomly shifted between $\pm 0.83^\circ$.

MRI data acquisition

Participants completed an MRI session directly following the psychophysical session. MRI data were acquired using a 3 T whole-body MRI system (Magnetom Prisma, Siemens Medical Systems, Germany), using a 64-channel RF receive head coil and body coil for transmission. The acquired data included a T1-weighted (T1w) high-resolution anatomical image and six sessions of 235 volumes each of functional images for pRF mapping. Details can be found in Jastrzębowska (2020).

Population receptive field mapping

The population receptive field (pRF) mapping procedure was based on the one suggested by van Dijk and colleagues (2016) to have comparatively the best intersession reliability among previously suggested stimulus configurations. The stimulus consisted of a simultaneous rotating wedge and expanding and contracting ring. The dimensions of the wedge and ring as well as the timing parameters were very similar to those described in van Dijk et al. (2016).

Six sessions of pRF mapping were conducted. A simple fixation task was used to ensure that participants looked at the central fixation point at all times. Participants were instructed to push a button each time the fixation point changed color. Each session ended with a 45-s period of fixation, with a total of 235 volumes acquired per session.

Data analysis: MRI

The T1w image was used for cortical reconstruction in FreeSurfer (FreeSurfer software package version 6.0, <http://surfer.nmr.mgh.harvard.edu/>; Dale et al., 1999). fMRI data from pRF mapping were processed according to the standard pre-processing pipeline using the statistical parametric mapping (SPM) software package (SPM12, Wellcome Trust Centre for Neuroimaging, London, UK, <http://www.fil.ion.ucl.ac.uk>) in MATLAB (The MathWorks, Inc., Natick, MA). For details, see Jastrzębowska (2020).

The pre-processed pRF mapping fMRI data were projected onto the cortical surface defined in FreeSurfer. The projected data was then used for pRF mapping with the SamSrf 6 toolbox (Schwarzkopf et al., 2019), which fitted a standard Gaussian 2d pRF model to the data.

We then delineated the visual area regions of interest (ROIs; V1 dorsal and ventral – V1d and V1v, respectively, V2d/v, V3d/v, V4) based on reversals in the polar angle map and restricted our ROIs to realistic eccentricities (i.e., those which were mapped with the stimulus: 0 to 12°) using the eccentricity maps.

Data analysis: illusion magnitude and pRF measures

Both frequentist and Bayesian analyses were computed. In accordance with Jeffreys (1961) and Lee and Wagenmakers (2014), we considered a BF_{10} between 3 and 10 to correspond to substantial evidence in favor of the alternative hypothesis. Correlation coefficients of 0.1, 0.3, and 0.5 were considered as small, medium, and large effect sizes, respectively (Cohen, 1988).

Illusion magnitude

As a measure of test-retest reliability, an intraclass correlation was computed between the two adjustments of each condition (see section 3.2.1). Illusion magnitudes were then computed as a difference compared to the reference, except for the HC and EX illusions, where the area of the adjusted ellipse was considered as the illusion magnitude. Then, both conditions of each illusion were combined, i.e., the illusion magnitude of the condition, which is usually underadjusted, was added (in absolute values) to the illusion magnitude of the condition, which is usually overadjusted. Between-illusion correlations were computed.

Visual surface area

In a first analysis, in an attempt to replicate the results of Schwarzkopf and colleagues (Schwarzkopf et al., 2011; Schwarzkopf & Rees, 2013), the surface area of visual areas V1, V2, V3 and V4 were calculated within eccentricities bounded by the foveal representation and 8 deg. The illusion magnitudes were then correlated to the surface areas of the ROIs. Data points with a z-score larger than 3 in terms of either illusion magnitude or ROI surface area were excluded.

pRF size

It has been suggested that rather than surface area *per se*, it is pRF size which determines illusion magnitude (Moutsiana et al., 2016; Schwarzkopf et al., 2011). Thus, in a second analysis, we extracted

pRF sizes at illusion-relevant eccentricities and computed correlations between these pRF sizes and the corresponding illusion magnitudes.

This analysis can only be done for illusions, in which the targets are at a fixed distance from one another. Moreover, the distance between the targets can only be reasonably determined for circular targets. Therefore, here we focused on the DB and EB2 illusion. In both cases, we assumed that participants used one of two strategies: (1) focusing their gaze at the midpoint location between the two targets, or (2) focusing their gaze directly on one or the other of the targets.

In the DB illusion, the distance between the centers of the two targets was 3.92° , with the targets positioned diagonally from each other – one in the upper left and the other in the lower right. Thus, we extracted pRF sizes from the following eccentricities and ROIs: right ventral (corresponding to the top left target) and left dorsal (corresponding to the bottom right target) ROIs, at eccentricities of 3.92° and 1.96° ($=3.92^\circ/2$, i.e., the midpoint between the two target locations). Similarly, for the EB2 illusion, we considered right and left hemisphere ROIs (since targets were positioned horizontally from each other), at eccentricities of 9.3° (distance between the targets) and 4.65° (midpoint between the targets).

We computed correlations between DB and EB2 illusion magnitudes and pRF sizes at the aforementioned eccentricities for each ROI. We removed any data point which had a z-score above 3 in terms of either illusion magnitude or pRF size.

Slope and intercept of pRF size as a function of eccentricity

The above pRF analysis was limited to just two illusions due to the difficulty of identifying “relevant” eccentricities for most illusions. In order to circumvent this issue and characterize the relationship between pRF size and illusion magnitude for all tested illusions, we conducted a final analysis, in which we estimated the slope and intercept of pRF size as a function of eccentricity. An association between the slope and illusion magnitude would indicate that interindividual susceptibility to the given illusion is predicated on the rate at which pRF size increases with eccentricity. Meanwhile, an association between the intercept and illusion magnitude would imply that the parafoveal pRF size is the determining factor for illusion susceptibility.

For each participant, each ROI and the two hemispheres separately, we fitted the least squares regression line to the scatter plots of pRF size as a function of eccentricity (see Figure 23 for an example). We computed the slope and intercept of the least squares line. We then averaged the slopes and intercepts from the two hemispheres. For some participants, the slope was estimated to be

negative due to noise in the data (see Figure 24). If this was the case in one hemisphere, we only took into account the slope from the other hemisphere. If it was the case in both hemispheres, we omitted the data point from further analysis.

For each illusion and each ROI, we conducted Bayesian linear regression analyses, with illusion magnitudes as the dependent variable and slope and/or intercept as the covariate(s). We tested four models: (1) model that included the slope, (2) model that included the intercept, (3) model that included both slope and intercept, (4) null model that did not include either slope or intercept. We estimated the Bayes factors (BF_{10}) in favor of each of the three alternative models over the null model.

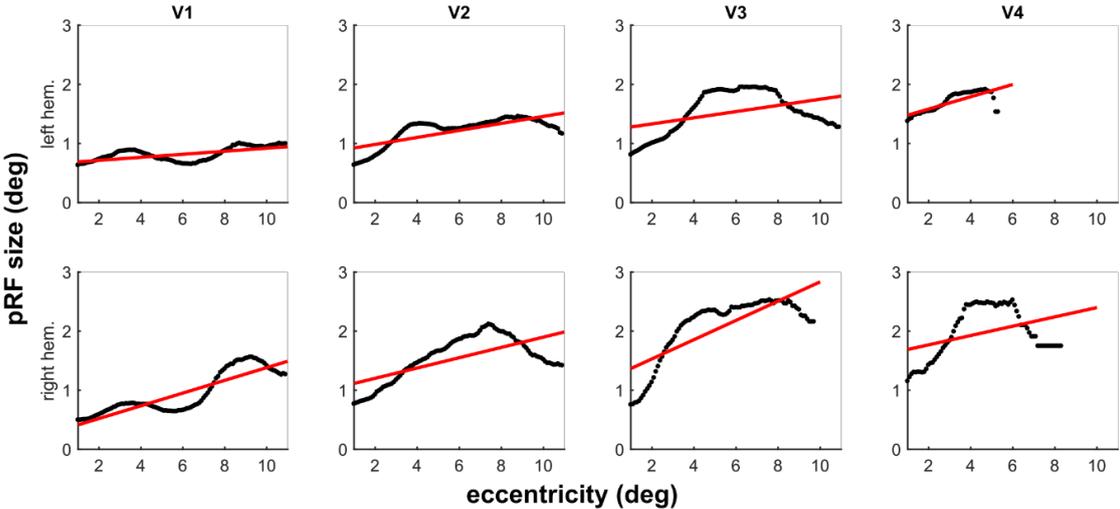


Figure 23. Appendix H. Least squares regression lines fitted to plots of pRF size as a function of eccentricity across ROIs and in the two hemispheres for a sample participant. All slopes are positive, as would be expected based on known properties of retinotopic organization.

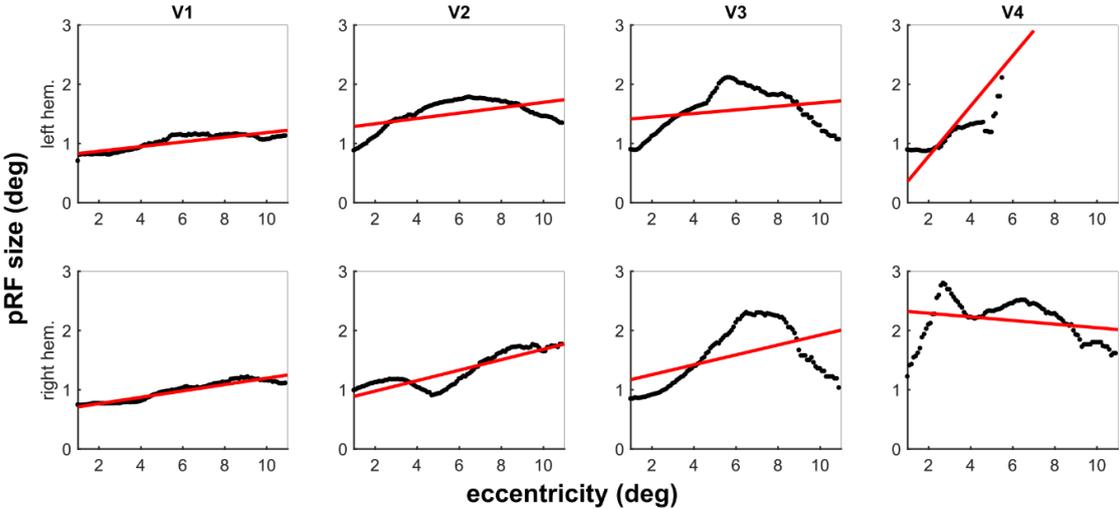


Figure 24. Appendix H. Least squares regression lines fitted to plots of pRF size as a function of eccentricity across ROIs and in the two hemispheres for another sample participant. Here, due to noise, slope is estimated to be negative in V4 of the right hemisphere.

Results

Illusion magnitude

All intraclass correlations were significant, except for the contrast conditions (Table 12), suggesting high test-retest reliabilities. The illusion magnitudes for both conditions of each illusion are shown in Figure 25. As usual, all illusions had an overadjusted and an underadjusted condition, except in the EX and HC illusions.

Table 12. Appendix H. Test-retest reliabilities expressed as an intraclass correlation (ICC) coefficient for each condition of each illusion ($n = 30$). P -values were not corrected for multiple comparison (i.e., $\alpha = 0.05$). Reprinted from Jastrzębowska (2020).

Illusion	Condition	ICC coefficient	95% CI	df	p
BS	horizontal	0.622	[0.394, 0.777]	[29, 29]	<0.001
	vertical	0.540	[0.285, 0.723]	[29, 29]	0.001
CS	left	0.274	[-0.029, 0.531]	[29, 29]	0.068
	right	0.149	[-0.159, 0.430]	[29, 29]	0.212
DB	left	0.478	[0.207, 0.681]	[29, 29]	0.003
	right	0.597	[0.360, 0.761]	[29, 29]	<0.001
EB1	small	0.784	[0.632, 0.878]	[29, 29]	<0.001
	large	0.304	[0.004, 0.554]	[29, 29]	0.048
EB2	small	0.608	[0.376, 0.769]	[29, 29]	<0.001
	large	0.686	[0.486, 0.818]	[29, 29]	<0.001
EX	black	0.763	[0.599, 0.865]	[29, 29]	<0.001
	white	0.906	[0.832, 0.948]	[29, 29]	<0.001
HC	black	0.921	[0.858, 0.957]	[29, 29]	<0.001
	white	0.840	[0.721, 0.911]	[29, 29]	<0.001
ML	outward	0.528	[0.270, 0.715]	[29, 29]	0.001
	inward	0.658	[0.445, 0.800]	[29, 29]	<0.001
PD	left	0.394	[0.105, 0.621]	[29, 29]	0.014
	right	0.699	[0.505, 0.826]	[29, 29]	<0.001
PZ	down	0.713	[0.525, 0.835]	[29, 29]	<0.001
	up	0.654	[0.440, 0.798]	[29, 29]	<0.001
PZc	down	0.850	[0.737, 0.916]	[29, 29]	<0.001
	up	0.642	[0.422, 0.790]	[29, 29]	<0.001
TT	left	0.340	[0.044, 0.581]	[29, 29]	0.031
	right	0.511	[0.249, 0.704]	[29, 29]	0.002
ZN	left	0.534	[0.278, 0.720]	[29, 29]	0.001
	right	0.773	[0.615, 0.871]	[29, 29]	<0.001

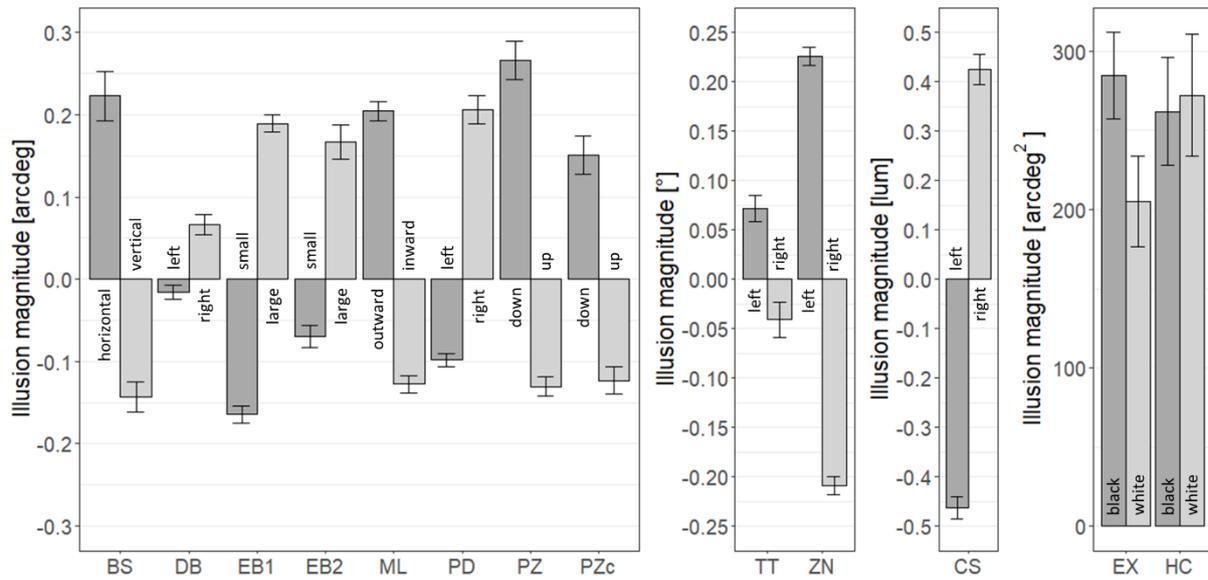


Figure 25. Appendix H. Illusion magnitudes for both conditions of each illusion. For all but the EX and HC illusions, positive and negative magnitudes indicate over- and underadjustments compared to the reference, respectively. Error bars show standard errors (*SE*). Reprinted from Jastrzębowska (2020).

Table 13 shows between-illusion correlations. As previously observed, correlations were in general weak, except between two variants of the same illusion (e.g., between PZ and PZc). However, the susceptibilities to the EX and HC illusions were strongly intercorrelated, suggesting that both illusions rely on a similar mechanism (see also Bertamini et al., 2019).

Table 13. Appendix H. Between-illusion correlation coefficients (Pearson's *r*). A color scale from blue to red reflects effect sizes from $r = -1$ to $r = 1$. Italics and bold font indicate significant results without ($\alpha = 0.05$) and with ($\alpha = 0.05/78$) Bonferroni correction, respectively. Reprinted from Jastrzębowska (2020).

	BS	CS	DB	EB1	EB2	EX	HC	ML	PD	PZ	PZc	TT	ZN
BS		0.114	-0.280	0.162	<i>0.426</i>	0.098	0.059	0.260	-0.006	-0.156	-0.069	-0.124	-0.133
CS			0.066	-0.196	-0.067	-0.313	-0.359	-0.160	0.090	-0.083	0.048	0.248	0.156
DB				0.100	0.146	0.085	0.071	<i>-0.428</i>	0.132	0.142	-0.060	0.107	-0.001
EB1					<i>0.589</i>	0.204	0.240	0.048	-0.021	0.221	0.207	0.235	-0.048
EB2						-0.013	-0.001	0.029	-0.046	0.053	0.077	0.240	0.070
EX							0.967	-0.126	<i>0.493</i>	0.268	0.165	-0.254	0.079
HC								-0.015	<i>0.521</i>	0.348	0.222	-0.187	0.093
ML									-0.084	0.114	-0.114	0.029	0.063
PD										0.219	-0.079	-0.222	0.073
PZ													<i>0.571</i>
PZc													0.239
TT													
ZN													0.236

Visual surface area

Schwarzkopf and colleagues (2011, 2013) reported negative correlations between V1 surface area and the susceptibility to the Ebbinghaus (EB2) and corridor Ponzo (PZc) illusions. Here, the metrics of these two illusions were matched to those reported in the relevant papers. Correlations between illusion magnitude and ROI surface area for each illusion-ROI pair are shown in Table 14. Bayes factors were computed. BF_{01} and BF_{10} larger than 3.0 indicate substantial evidence for the null and alternative hypotheses, respectively.

Table 14. Appendix H. Correlations between ROI surface areas (mm^2) and illusion magnitudes with outliers removed, expressed as Pearson's r and Bayes factor (BF_{01} and BF_{10}). Significant correlations ($p < 0.05$, uncorrected for multiple comparisons) are marked in blue with an asterisk. Instances of substantial evidence in favor of either the null or the alternative model ($BF \in [3,10]$; Jeffreys, 1961; Lee & Wagenmakers, 2014) are marked in red with an asterisk. Adapted from Jastrzębowska (2020).

Illusion	n	V1			V2			V3			V4		
		r	BF_{01}	BF_{10}	r	BF_{01}	BF_{10}	r	BF_{01}	BF_{10}	r	BF_{01}	BF_{10}
BS	30	0.32	1.08	0.92	0.37*	0.61	1.65	0.42*	0.35	2.89	0.29	1.43	0.70
CS	30	0.11	3.73*	0.27	0.22	2.30	0.44	0.20	2.58	0.39	0.36	0.73	1.38
DB	30	-0.32	1.07	0.93	-0.23	2.17	0.46	-0.28	1.53	0.65	-0.01	4.41*	0.23
EB	30	-0.02	4.37*	0.23	0.19	2.78	0.36	0.10	3.88*	0.26	0.35	0.82	1.23
EB2	29	0.04	4.25*	0.24	0.27	1.68	0.59	0.16	3.19*	0.31	0.17	2.97	0.34
EX	30	-0.23	2.23	0.45	-0.15	3.33*	0.30	0.01	4.40*	0.23	-0.07	4.14*	0.24
HC	30	-0.24	1.96	0.51	-0.19	2.79	0.36	-0.03	4.35*	0.23	-0.09	3.99*	0.25
ML	30	0.03	4.37*	0.23	0.01	4.40*	0.23	0.04	4.32*	0.23	-0.12	3.60*	0.28
PD	29	-0.18	2.93	0.34	-0.30	1.32	0.76	-0.19	2.76	0.36	-0.23	2.23	0.45
PZ	29	-0.13	3.52*	0.28	-0.17	2.97	0.34	-0.37*	0.68	1.47	-0.36	0.76	1.31
PZc	29	0.18	2.89	0.35	0.29	1.40	0.72	-0.02	4.30*	0.23	-0.15	3.26*	0.31
TT	30	0.24	2.00	0.50	0.29	1.37	0.73	0.28	1.57	0.64	0.33	0.95	1.05
ZN	30	-0.08	4.02*	0.25	-0.12	3.59*	0.28	-0.02	4.38*	0.23	0.01	4.40*	0.23

The present results failed to replicate those reported by Schwarzkopf et al. (2011, 2013). Indeed, there was evidence for the null hypothesis in the correlation between the EB2 illusion magnitude and V1 surface area ($r = 0.04$; $BF_{01} = 4.25$). Similarly, rather weak and non-significant correlations were observed for the other variant of the Ebbinghaus illusion (EB1). In addition, there was no significant correlation between the susceptibility to PZc and V1 surface area ($r = 0.18$; $BF_{01} = 2.89$). A medium effect size (Gignac & Szodorai, 2016) was observed between V1 surface area and the Delboeuf illusion magnitude ($r = -0.32$; $BF_{01} = 1.07$), even though the Bayes factor was inconclusive.

Importantly, no illusion-ROI correlation resulted in a BF_{10} larger than 3. Similarly, only three correlations were significant, i.e., between the Ponzo (PZ) illusion magnitude and V3 surface area and between the bisection (BS) illusion magnitude and V2 and V3 surface areas. Note, however, that the latter two were positive, while Schwarzkopf et al. (2011, 2013) suggested that the mediating factor in

the relationship between the illusion magnitude and cortical size is the spatial spread or neuronal connections between the inducers and the target, which favors negative (rather than positive) associations.

PRF size

Correlations were computed between illusion magnitude and pRF size at eccentricities relevant for the the Delboeuf illusion and Ebbinghaus 2 illusions. Bayes factors (BF_{01} and BF_{10}) for each correlation are reported in Table 15, along with Pearson's r . In general, we failed to observe an association between pRF size and illusion magnitude. For the Delboeuf illusion, only one correlation was significant, in the case of the smaller eccentricity and area V2 ($p < 0.01$, uncorrected). In this instance, the Bayes factor indicated substantial evidence for the alternative hypothesis ($BF_{10} = 7.08$). While we observed other correlations of medium effect size (Delboeuf: V1, smaller eccentricity; V2, larger eccentricity; Ebbinghaus 2: V4, smaller eccentricity), none of them were statistically significant ($p < 0.05$ uncorrected) and the evidence for the alternative hypothesis was merely anecdotal as indicated by the BF_{10} .

Table 15. Appendix H. Correlations between pRF size (deg) at relevant eccentricities and illusion magnitudes with outliers removed. Significant correlations ($p < 0.05$, uncorrected for multiple comparisons) are marked in blue with an asterisk. Instances of substantial evidence in favor of either the null or the alternative model ($BF \in [3,10]$; Jeffreys, 1961; Lee & Wagenmakers, 2014) are marked in red with an asterisk.

Visual area	Illusion	n	Eccentricity (deg)	Pearson's r	BF_{01}	BF_{10}
V1	DB	29	1.96	0.31	1.19	0.84
		29	3.92	0.01	4.32*	0.23
	EB2	28	4.65	0.10	3.76*	0.27
		29	9.3	-0.19	2.73	0.37
V2	DB	29	1.96	0.49*	0.14	7.08*
		29	3.92	0.35	0.82	1.22
	EB2	29	4.65	0.04	4.25*	0.24
		29	9.3	0.15	3.25*	0.31
V3	DB	28	1.96	0.11	3.64*	0.28
		29	3.92	0.22	2.30	0.44
	EB2	28	4.65	0.26	1.85	0.54
		28	9.3	0.09	3.90*	0.26
V4	DB	26	1.96	-0.2	2.60	0.38
		29	3.92	0.11	3.75*	0.27
	EB2	27	4.65	0.34	1.02	0.98
		24	9.3	0.10	2.90	0.35

Slope and intercept of pRF size as a function of eccentricity

Bayes factors (BF_{10}) in favor of each of the three alternative Bayesian linear regression models – slope, intercept and slope+intercept of pRF size as a function of eccentricity – are reported in Table 16. We did not find substantial evidence for any of the tested models with respect to the null model for any of the tested illusions.

Table 16. Appendix H. Results of Bayesian linear regression analysis of illusion magnitudes (dependent variable), using slope and intercept of pRF size as a function of eccentricity as covariates. Results consist of Bayes factors (BF_{10}) of three alternative models w.r.t. the null model: (1) model that includes the slope, (2) model that includes the intercept, (3) model that includes both slope and intercept. There was not a single instance of substantial evidence in favor of the alternative model ($BF_{10} \in [3,10]$); Jeffreys, 1961; Lee & Wagenmakers, 2014).

Visual area	Illusion	n	BF_{10}		
			slope	intercept	slope+intercept
V1	BS	27	0.36	1.30	0.75
	CS	27	0.72	0.37	0.35
	DB	27	0.45	0.42	0.22
	EB	27	0.54	0.38	0.38
	EB2	26	0.40	0.38	0.24
	EX	27	0.88	0.62	0.44
	HC	27	0.95	0.73	0.50
	ML	27	0.38	0.37	0.19
	PD	26	0.42	0.88	0.40
	PZ	26	0.53	2.74	1.12
	PZc	26	0.37	1.05	0.56
TT	27	0.44	0.67	0.31	
ZN	27	0.36	1.21	0.66	
V2	BS	29	0.41	0.35	0.24
	CS	29	0.64	0.38	0.30
	DB	29	2.59	1.98	1.45
	EB	29	0.63	0.36	0.57
	EB2	28	0.35	0.38	0.19
	EX	29	0.38	1.05	0.55
	HC	29	0.38	0.78	0.39
	ML	29	0.35	0.36	0.17
	PD	28	0.37	0.77	0.42
	PZ	28	0.35	0.64	0.40
	PZc	28	0.72	0.84	0.42
TT	29	0.56	0.44	0.26	
ZN	29	0.62	0.43	1.34	
V3	BS	29	0.86	0.35	0.41
	CS	29	0.55	0.36	0.25
	DB	29	0.37	0.50	0.24
	EB	29	0.40	0.40	0.23
	EB2	28	0.36	0.50	0.24
	EX	29	0.35	0.57	0.28
	HC	29	0.39	0.43	0.25
	ML	29	0.71	1.74	1.11
	PD	28	0.37	1.22	0.61
	PZ	28	0.36	0.37	0.19
	PZc	28	0.49	0.35	0.23
TT	29	0.36	0.36	0.18	
ZN	29	1.40	0.35	0.62	
V4	BS	25	0.63	0.48	0.30
	CS	25	0.56	0.37	0.43
	DB	25	0.38	0.53	0.28
	EB	25	2.42	1.45	1.19
	EB2	24	0.40	0.52	0.26
	EX	25	0.37	0.39	0.20
	HC	25	0.37	0.40	0.20
	ML	25	0.38	0.88	1.20
	PD	24	0.93	0.65	0.44
	PZ	24	0.38	0.62	0.36
	PZc	24	1.29	0.55	0.59
TT	25	0.37	0.37	0.19	
ZN	25	0.39	0.40	0.20	

Appendix I

Bertamini, M., Cretienoud, A. F., & Herzog, M. H. (2019). Exploring the Extent in the Visual Field of the Honeycomb and Extinction Illusions. *i-Perception*, 10(4), 1–19, <https://doi.org/10.1177/2041669519854784>.

Copyright: Held by the authors

Published under CC-BY 4.0

Exploring the Extent in the Visual Field of the Honeycomb and Extinction Illusions

i-Perception

2019, Vol. 10(4), 1–19

© The Author(s) 2019

DOI: 10.1177/2041669519854784

journals.sagepub.com/home/ipe

**Marco Bertamini** 

Department of Psychological Sciences, University of Liverpool, UK

Aline F. Cretenoud and Michael H. Herzog

Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Abstract

There are situations in which what is perceived in central vision is different to what is perceived in the periphery, even though the stimulus display is uniform. Here, we studied two cases, known as the Extinction illusion and the Honeycomb illusion, involving small disks and lines, respectively, presented over a large extent of the visual field. Disks and lines are visible in the periphery on their own, but they become invisible when they are presented as part of a pattern (grid). Observers ($N = 56$) adjusted a circular probe to report the size of the region in which they had seen the lines or the disks. Different images had black or white lines/disks, and we included control stimuli in which these features were spatially separated from the regular grid of squares. We confirmed that the illusion was experienced by the majority of observers and is dependent on the interaction between the elements (i.e., the lines/disks have to be near the squares). We found a dissociation between the two illusions in the dependence on contrast polarity suggesting different mechanisms. We analysed the variability between individuals with respect to schizotypal and autistic-spectrum traits (short version of the Oxford-Liverpool Inventory of Feelings and Experiences [O-LIFE] questionnaire and the Autistic Quotient, respectively) but found no significant relationships. We discuss how illusions relative to what observers are aware of in the periphery may offer a unique tool to study visual awareness.

Keywords

texture, crowding, peripheral vision, suppression, Honeycomb illusion, Extinction illusion

Received 18 February 2019; accepted 8 May 2019

Corresponding author:

Marco Bertamini, Department of Psychological Sciences, University of Liverpool, Eleanor Rathbone Building, Liverpool L69 7ZA, UK.

Email: M.Bertamini@liverpool.ac.uk



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<http://www.creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

For humans, the visual field extends approximately 180° horizontally and 160° vertically. However, only parts of the visual field are binocular, a large blind spot is present in each hemifield, and both resolution and colour vision vary greatly with eccentricity (Helmholtz, 1867). Our experience, nevertheless, is that of a uniform and stable visual field, possibly because of the properties perceived in the central region (Gibson, 1950). This experience of a detailed and uniform visual field can be described as an illusion, perhaps the most striking of all visual illusions. This has been studied by psychologists and commented on by philosophers and is sometimes called the *grand illusion hypothesis* (Blackmore, Brelstaff, Nelson, & Trościanko, 1995; Dennett, 1991; Noë, Pessoa, & Thompson, 2000; Rensink, O'Regan, & Clark, 1997).

A large literature has explored, both theoretically and empirically, to what extent summary statistical information can surmount the limitations imposed by the visual system on the representation of individual elements (Chong & Treisman, 2003; Cohen, Dennett, & Kanwisher, 2016; Dehaene, 2014; Lamme, 2010). A summary or ensemble representation implies that sets of similar items are coded only in terms of summary statistics (Haberman & Whitney, 2012). McClelland and Bayne (2016) have proposed a distinction between summary statistics that observers are aware of, although they may be lost in memory, and a second type of summary representations that are not part of phenomenal experience, although they can affect post-perceptual judgements. It is known that early visual responses are combined in receptive fields that grow with eccentricity. Freeman and Simoncelli (2011) have developed a model and created visual metamers: stimuli that differ physically in the periphery but look the same.

To study what observers are aware of in the periphery researchers have developed various procedures. There are also some visual illusions that provide useful insights.

Honeycomb and Extinction Illusions

Recently, Bertamini, Herzog, and Bruno (2016) have described an illusion characterised by the inability of seeing shapes in the periphery, when these shapes are part of a texture, as shown in Figure 1. The uniform pattern (hexagons and lines) extends over a large part of the visual field, but in the periphery, observers are not aware of some of its features (lines). Therefore, a physically uniform stimulus yields a non-uniform percept. Bertamini et al. (2016) named this effect the Honeycomb (HC) illusion and we will use the acronym HC. In Figure 1, the small lines are visible at fixation, but invisible everywhere else. This is true even after multiple fixations, and the phenomenon is robust to changes in shape and size of the elements. Because the experience does not change over time and with multiple fixations, memory seems to have no role in the build-up and maintenance of the representation of the visual field. Moreover, this representation is not an extrapolation from information available to central vision.

A closely related phenomenon is known as the Extinction (EX) illusion (Ninio & Stevens 2000). This was described as a variation on the Hermann grid (HG), but it is in fact a striking phenomenon on its own. We will use the shorten name EX illusion. As for the HC illusion, features in the periphery completely disappear from awareness. In the case of the EX illusion, they are local disks at the intersection of lines. For a strong effect, the disks have to be light (e.g., light grey) on a mainly dark background (e.g. black) or vice versa.

The experimental evidence about these illusions is limited with a few notable exceptions. McAnany and Levine (2004) and Levine, Anderson, and McAnany (2012) reported a series of experiments on a phenomenon that they call *blanking*. This is closely related to the EX illusion in that a light disk becomes invisible when presented at an intersection of black

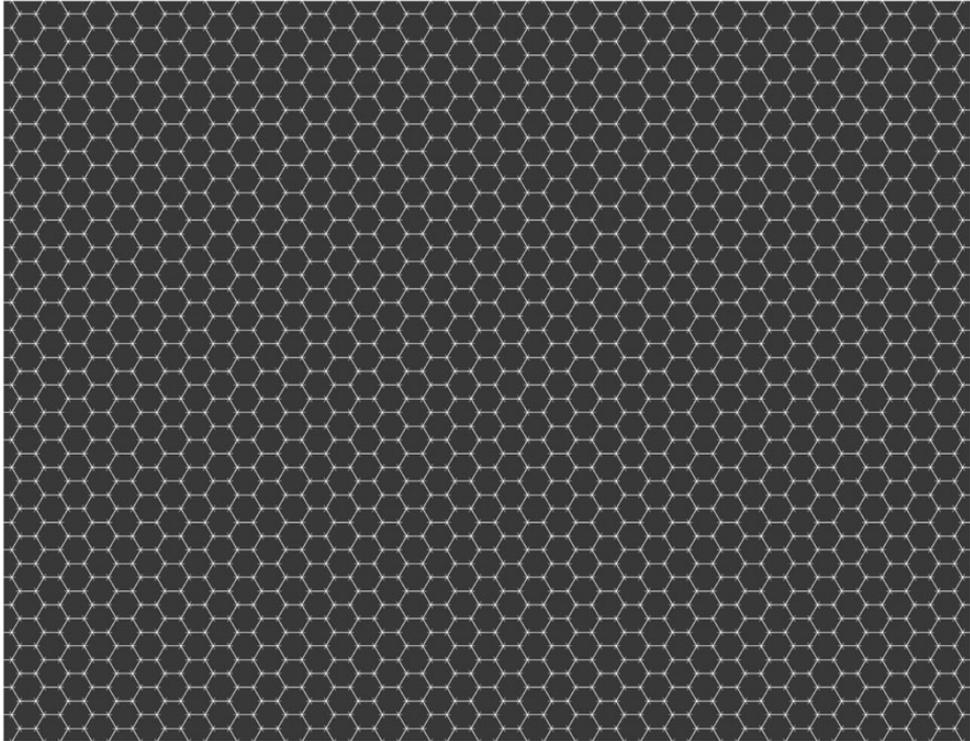


Figure 1. This texture has hexagons and small lines at the vertices of the hexagons. The texture is uniform, but it is not perceived as such because the lines can only be seen at or near fixation. The image needs to be enlarged or looked at closely, so as to fill a large proportion of the visual field; for more examples, go to <https://osf.io/kabyz/>

squares forming a grid. They demonstrated that the disk and at least four squares of the grid must be presented simultaneously. Levine et al. (2012) discovered also that distorting the squares, that is, making the alleys wavy, makes the illusion stronger. Levine and McAnany (2008) have argued that there are two effects for disks at intersections of a grid with dark squares: An obscuring process affects every target in a grid, while blanking only affects the light disks. What that means is that the blanking, in agreement with Ninio and Stevens (2000) observations, is strongest when there is a difference in polarity between the patterns of squares and the disks: When the squares are black, the disks need to be white or vice versa.

Common to all these effects, and in particular to the HC and EX illusions, is that they are instantaneous and do not require any adaptation period. However, their strength may vary over time. Araragi and Kitaoka (2011) found that the EX illusion becomes stronger with longer presentations. At short presentations, they also reported an anisotropy: The illusion occurred more frequently in the upper visual field than in the lower visual field.

Note that these illusions show the opposite of an extrapolation effect. What is visible in central vision is not extrapolated, even though that would produce a veridical percept. Extrapolation effects have been reported, but they seem to require extended presentation time and adaptation (Otten, Pinto, Paffen, Seth, & Kanai, 2017). An example of an adaptation phenomenon is the Healing grid illusion in which the regularity present at central fixation becomes a property of the grid in the periphery over time (Kanai, 2005).

In discrimination tasks, there is also evidence that foveal information can affect judgements of shape (Yu & Shim, 2016) and of brightness (Toscani, Gegenfurtner, & Valsecchi, 2017) in the periphery. Gloriani and Schütz (2019) have reported evidence that observers trust what is perceived in central vision both in photopic and, surprisingly, in scotopic vision.

One of the best known and most studied effects about vision in the periphery is the HG illusion (Geier, Bernáth, Hudák, & Séra, 2008; Hermann, 1870). The full explanation of the HG phenomenon is still under study and is more complex than a process of lateral inhibition (Qian, Yamada, Kawabe, & Miura, 2009; Read, Robson, Smith, & Lucas, 2012; Schiller & Carvey, 2005). Some variants of the HG are especially interesting and constitute distinct phenomena. In particular, blurring the HG changes the dark smudges into scintillating black spots (Bergen, 1985; Schrauf, Lingelbach, Lingelbach, & Wist, 1995; Schrauf, Lingelbach, & Wist, 1997). Other studies have reported shape distortions: Disks become ellipses away from fixation (Qian & Mitsudo, 2016).

Like the HC and the EX illusions, the HG effect is instantaneous and requires an extended grid. It is also similar in that what is seen in central vision is different from what is experienced in the periphery. The fundamental difference is that additional illusory disks or patches are reported in the HG, rather than a failure to report shape properties (features) that are present in the periphery. In this sense, the HC and EX are the opposite of the HG effect: They show a disappearance rather than an appearance of features.

A Dissociation Between HC and EX

We devised a procedure that allows observers to directly report the extent of the visual field within which the lines or the disks are visible. We used this procedure to measure the strength of the illusions and to test the role of contrast polarity. As mentioned there is some evidence that high contrast between the grid and the features contributes to the EX (Levine & McAnany, 2008), but not to the HC (Bertamini et al., 2016). If this dissociation is confirmed within a study that directly compares the two illusions, it would suggest the existence of different mechanisms.

In terms of stimuli, we modified the original HC and EX illusions so as to make them more similar. In both cases, the underlying grid was made of squares. The size was the same and the only difference was that for the HC illusion the squares were outlines, with additional small lines at the corners, and for the EX illusion the squares were filled, and disks were placed at some of the intersections. Examples are shown in Figures 2 and 3.

During the study, there was a presentation of the background grid without the lines or the disks, a brief additional presentation of these features (lines or disks) and then an adjustment of a circular gauge allowing the observer to report the extent of the area in which the features were visible. We acknowledge that this is a type of self-report procedure and therefore subject to possible biases. We chose it so that observers can use the circle to provide a report of their phenomenal experience. Moreover, we interleaved several conditions, including control stimuli for which we expect no illusion, thus providing a baseline.

Individual Differences in HC and EX

There has been great interest recently on the use of visual illusions to study individual differences and the existence of common factors (Grzeczowski, Clarke, Francis, Mast, & Herzog, 2017). Susceptibility to illusions may vary in different clinical conditions, such as autism spectrum disorders (ASD) or dyslexia (Gori, Molteni, & Facoetti, 2016; Slaghuis, Twell, & Kingston, 1996), and there are other factors, including physiological (Peterzell &

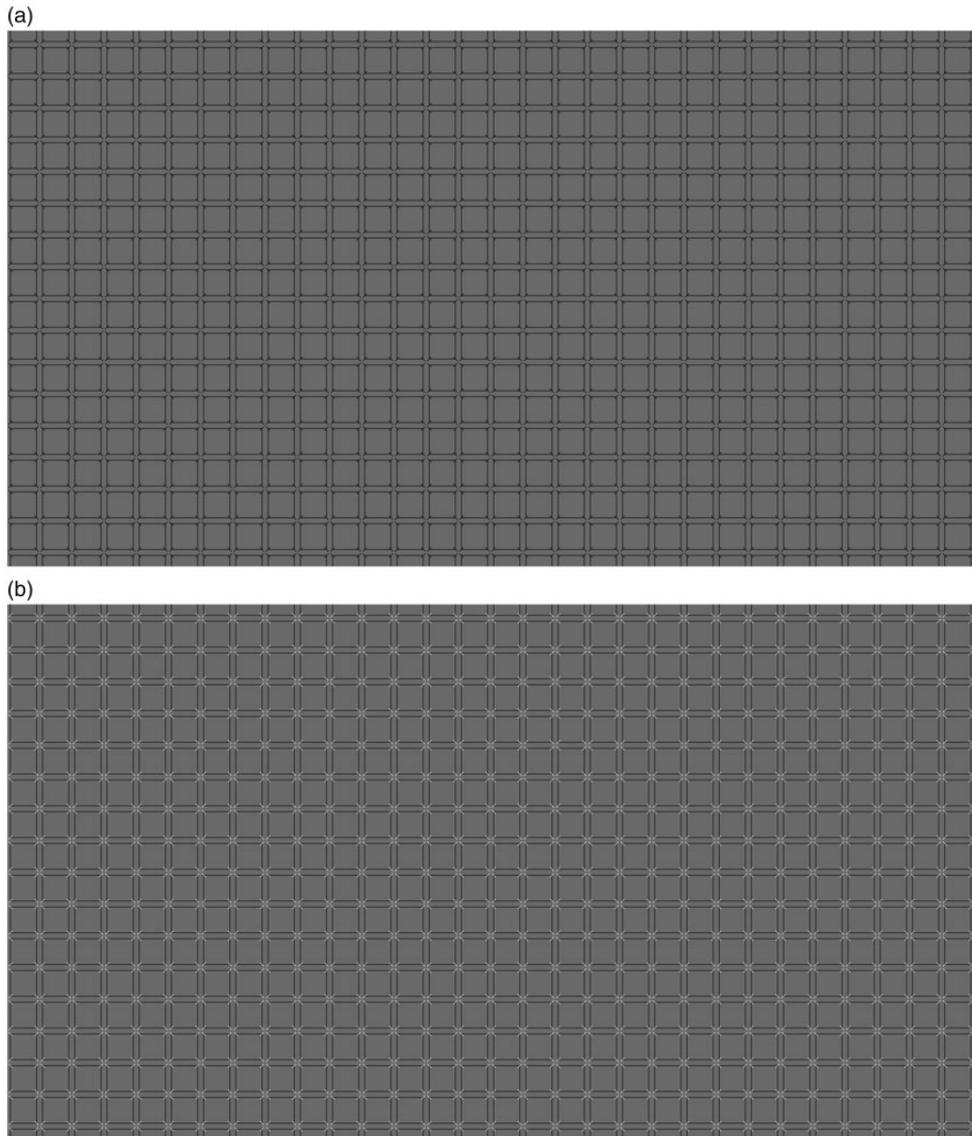


Figure 2. In this version of the Honeycomb illusion, squares replaced hexagons. In one version, the lines were black (a), and in another, the lines were white (b).

Kennedy, 2016) or cultural (de Fockert, Davidoff, Fagot, Parron, & Goldstein, 2007), which contribute to individual differences.

Illusions relating to awareness of stimulus properties in the periphery may be especially useful in relation to individual differences. They tell us something about visual awareness and show how available information, from memory or from central vision, is (or is not) used to construct a representation of the whole visual field experience.

For our study, we tested a sample of 56 observers. They were all undergraduate students, and therefore drawn from a restricted population. Most of them were females. We administered two standard questionnaires: the short version of the Oxford-Liverpool Inventory of

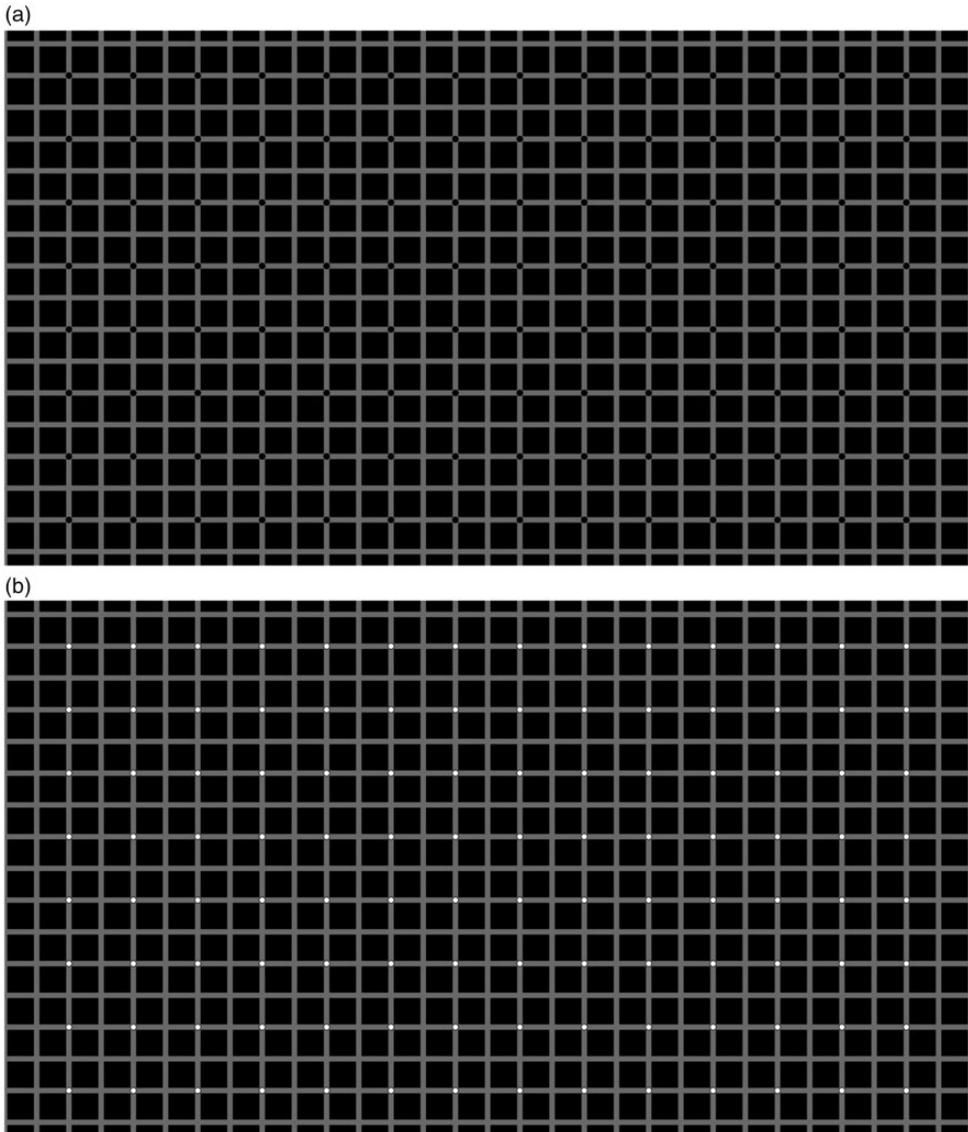


Figure 3. The Extinction illusion (a). In one version, the disks were black (a), and in the other, disks were white (b).

Feelings and Experiences (short O-LIFE) and the Autism-Spectrum Quotient (AQ) questionnaire. The short O-LIFE (Mason, Linney, & Claridge, 2005) is intended for use in the general population to test for schizotypal traits and it includes three subscales of schizotypy; unusual symptoms, cognitive disorganisation and introvertive anhedonia. Some studies have found a decreased susceptibility to visual illusions in schizophrenia (for a review, see Notredame, Pins, Deneve, & Jardri, 2014), others (e.g., Grzeczowski et al., 2018) found no difference in illusion strength between healthy controls and schizophrenic patients (see also Kaliuzhna et al., 2018). In the specific case of these illusions (in the visual periphery), it may

be more relevant that individuals high in schizotypy have a response bias toward seeing something that is not there or that is not visible (Partos, Cropper, & Rawlings, 2016).

With respect to the AQ, there is a vast literature on ASD and visual perception (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). A dissociation between psychosis and autism has been proposed by some authors (Crespi & Badcock, 2008). The AQ includes five subscales: social skills, attention switching, attention to detail, communication and imagination. It has been claimed that ASD individuals have weak global processing and a local over global advantage (Gori et al., 2016; Happé 1996; for reviews, see Simmons et al., 2009). We expect a negative correlation between AQ scores and the strength of the illusion. This prediction follows from the idea that attention to detail may reduce the phenomenal awareness of elements in the periphery. Reduced susceptibility to classic visual illusions has been reported in children with ASD (Happé, 1996), although Manning, Morgan, Allen, and Pellicano (2017) found that this was the case only when a method of adjustment was used. Note that our study does use a method of adjustment.

A more general consideration is that visual illusions are often reported as demonstrations or on the bases of reports from small samples. In the existing literature, this is true for both the HC and the EX illusions. Our study tested the robustness of these phenomena within a relatively large sample, although limited to undergraduate students, and we analysed, in exploratory fashion, the individual differences measured by two standard questionnaires (O-LIFE and AQ).

Methods

Observers

Fifty-six adults participated (54 females). All observers had normal or corrected-to-normal visual acuity. They were all undergraduate students at the University of Liverpool and were unaware of the purpose of the study. The study was approved by the local Ethics committee and written consent was obtained from all observers.

Stimuli and Apparatus

The images were generated using Python and PsychoPy (Peirce, 2007) on a Macintosh computer and displayed on a Sony monitor (52.80×29.70 cm). The images filled the screen and had always a red cross in the centre. A chinrest was used to fix the distance from the screen at 57 cm.

The images for the HC illusion are shown in Figure 2, and the images for the EX illusion are shown in Figure 3. The background was always grey (luminance 44.5 cd/m^2), although in the EX case much of the background was covered by black squares (luminance 2.9 cd/m^2). The squares that formed the grid could be of two sizes: 0.43° or 0.86° of visual angle (25.8 and 51.6 arcmin). We refer to these as small and large square conditions. The features (lines or disks) were black or white in two different conditions. We refer to this as contrast polarity because it is a change in opposite direction relative to the grey background. Figures 2 and 3 show the large squares images. Figures 4 and 5 show the small squares images. The small squares stimuli were control conditions because with smaller squares the lines and the disks are not touching the squares. This is expected to lead to a reduced illusion.

The study included two illusions (HC and EX) two square sizes (large and small), two contrast polarity of the features (black and white) and three presentation durations (250, 500

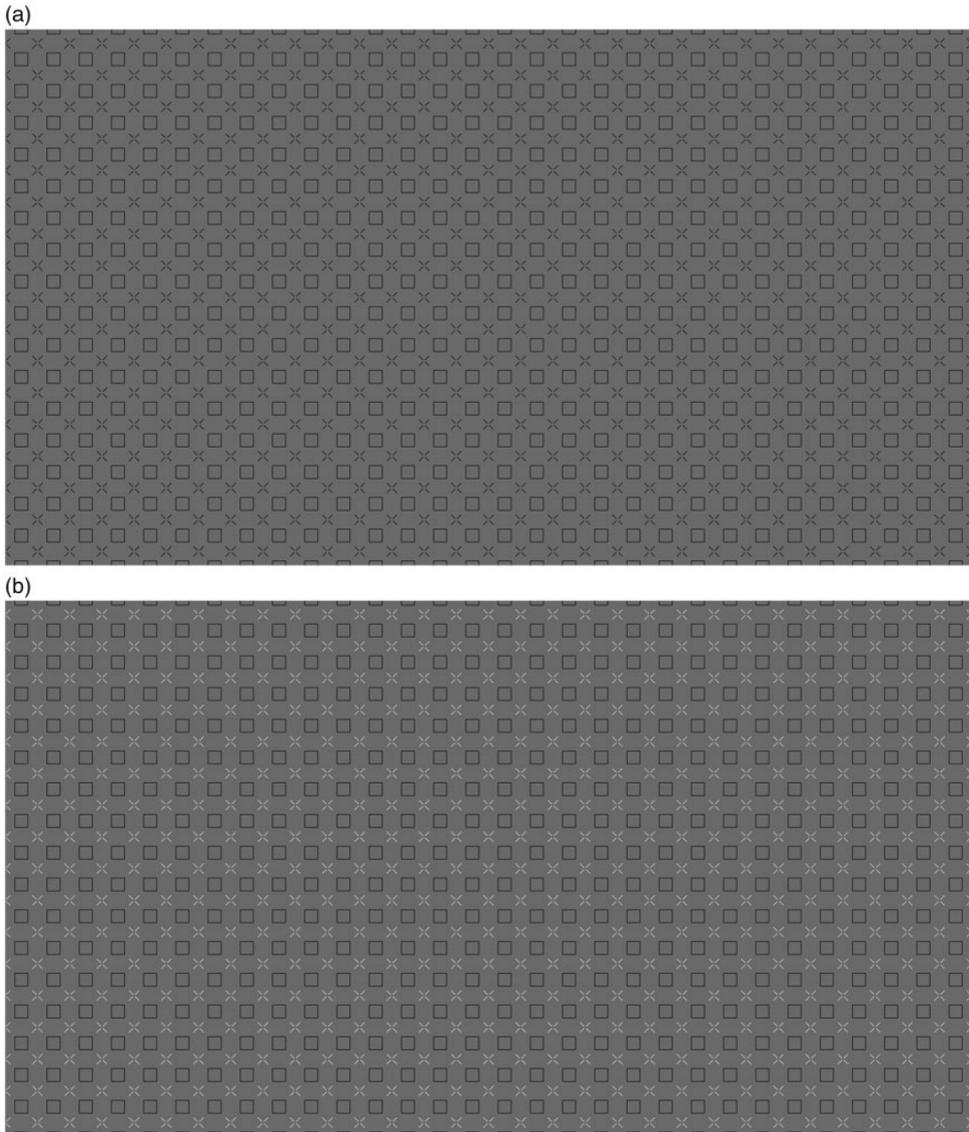


Figure 4. This is a control condition for the Honeycomb illusion in which the squares do not touch the lines. In one version, the lines were black (a), and in the other, lines were white (b).

and 750 ms). There were, therefore, 24 unique conditions. Each was presented 3 times for a total of 72 trials per observer.

Procedure

Each observer was tested individually in a dark and quite room. The texture made of squares was on the screen before the lines (HC) or disks (EX) were added and also at the time of the response. The sequence of events is illustrated in Figure 6. The grid was first presented for 1,000 ms. High luminance blank transients (white screen) were shown for 33.3 ms and the

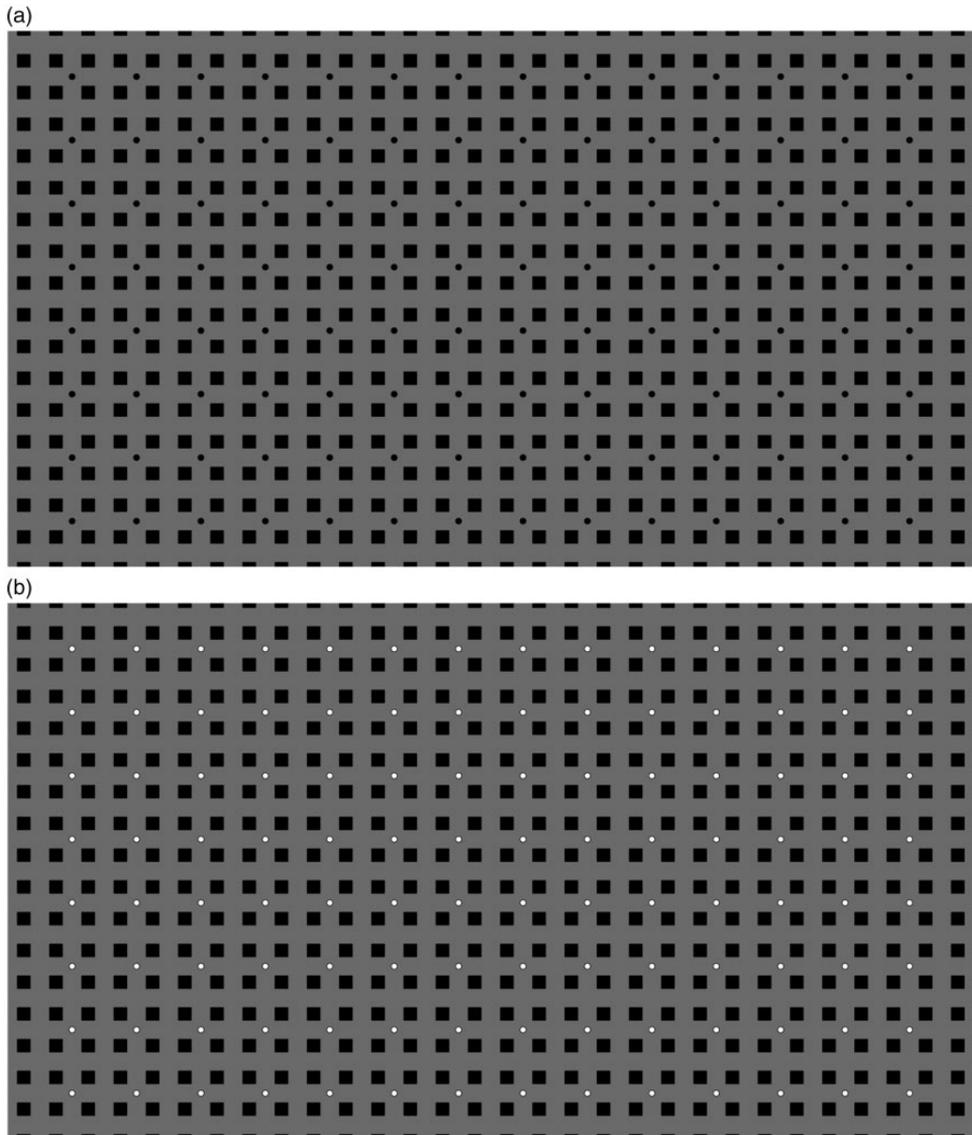


Figure 5. This is a control condition for the Extinction illusion in which the squares do not touch the disks. In one version, the disks were black (a), and in the other, disks were white (b).

texture with the lines/disks for a variable amount of time based on condition. The luminance for the white screen was 230.0 cd/m^2 . There were three presentation durations: 250 ms, 500 ms and 750 ms.

In the response period, observers used a game controller to adjust a circle (the right and left buttons could make it increase or decrease in size). They could use a different button to report that they could see the lines (HC) or disks (EX) over the entirety of the image. There was no time pressure and after they were happy with the response they had to press another button on the controller to advance to the next trial. During the experiment, participants used a chinrest to keep them 57 cm away from the screen. Viewing was binocular.

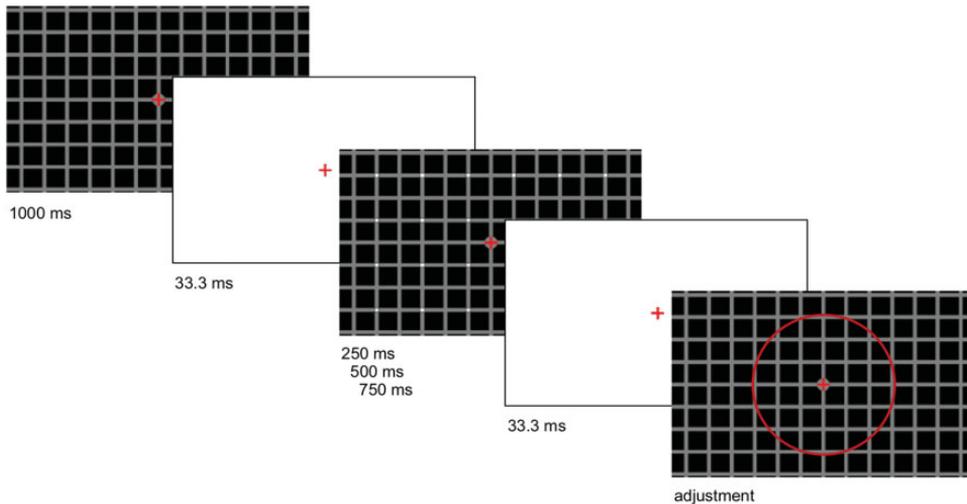


Figure 6. The procedure included a long presentation of the grid and a fixed period of presentation of the additional lines or disks. White screens were also included to minimise the role of luminance onset in the detection of features. The final image shows the red circle used for the adjustment. To make the disks visible in the figure, these images are cropped and show only a central region of the stimulus.

Data Preprocessing

To check for intrarater reliability, we computed two-way mixed-effects models—intraclass correlations of type (3,1) or ICC (3,1)—on the raw data set for each condition, as suggested in Shrout and Fleiss (1979) and in Koo and Li (2016).

For all conditions, the three adjustments of each observer were averaged. We then computed z-scores for each of the two size conditions. All data were within the ± 3 standard deviations of the mean range.

Results

Intrarater Reliability

All intraclass correlations were significant (Bonferroni corrected), indicating that participants made consistent adjustments between all three trials of each condition (Table 1).

Small Square Conditions

Small square conditions were included as a control to demonstrate that the difficulty of seeing the lines or the disks was due to an interaction with the squares and not simply due to eccentricity. Since observers could also report that lines and disks were visible over the whole screen, data were inevitably truncated at screen size.

Figure 7 (left panel) shows the mean extent of the region in which lines and disks are visible for the two illusions (HC and EX) separately for small and large square conditions. Strong ceiling effects were observed for the small square conditions. A paired *t* test (two-tailed) comparing the small square conditions to the large square conditions resulted in a significant difference, $t(671) = 22.458$, $p < .001$, suggesting that the illusion was reduced when

Table 1. All Intraclass Correlations Were Significant (Bonferroni Corrected) for Each Condition With 95% Confidence Intervals Excluding Zero, Indicating Good Intrarater Reliability Over Three Trials.

	EXBlack250S	EXBlack500S	EXBlack750S	EXWhite250S	EXWhite500S	EXWhite750S
ICC (3,1)	0.463	0.574	0.445	0.652	0.595	0.543
F value	3.589***	5.046***	3.407***	6.616***	5.406***	4.565***
95% CI	[0.303, 0.615]	[0.427, 0.704]	[0.283, 0.600]	[0.520, 0.764]	[0.451, 0.720]	[0.391, 0.680]
	HCBlack250S	HCBlack500S	HCBlack750S	HCWhite250S	HCWhite500S	HCWhite750S
ICC (3,1)	0.462	0.558	0.703	0.467	0.664	0.501
F value	3.573***	4.790***	8.094***	3.630***	6.929***	4.010***
95% CI	[0.301, 0.614]	[0.409, 0.692]	[0.583, 0.801]	[0.307, 0.619]	[0.535, 0.773]	[0.344, 0.646]
	EXBlack250L	EXBlack500L	EXBlack750L	EXWhite250L	EXWhite500L	EXWhite750L
ICC (3,1)	0.369	0.408	0.407	0.410	0.433	0.384
F value	2.751***	3.071***	3.056***	3.082***	3.290***	2.868***
95% CI	[0.203, 0.534]	[0.244, 0.569]	[0.243, 0.567]	[0.246, 0.570]	[0.270, 0.590]	[0.219, 0.548]
	HCBlack250L	HCBlack500L	HCBlack750L	HCWhite250L	HCWhite500L	HCWhite750L
ICC (3,1)	0.552	0.394	0.330	0.459	0.548	0.543
F value	4.695***	2.950***	2.479***	3.550***	4.631***	4.569***
95% CI	[0.402, 0.687]	[0.229, 0.557]	[0.164, 0.500]	[0.299, 0.612]	[0.397, 0.684]	[0.392, 0.680]

Note. '250', '500' and '750' indicate the presentation durations (in milliseconds); EX = Extinction illusion; HC = Honeycomb illusion; S = small square conditions; L = large square conditions; ICC = intraclass correlation; CI = confidence interval. *** $p < .001$.

the features (lines or disks) did not touch the squares. Only large square conditions were considered for further analyses.

Large Square Conditions

The mean values from the large square conditions are plotted in the middle and right panels of Figure 7 (within an orange rectangle). As observed in the middle panel, there was a significant interaction between the illusion and contrast polarity. Importantly, this interaction might even be stronger than what we report, because the measured responses for the HC illusion with white lines were affected by a ceiling effect.

Mixed-Effects Model

To account for random variations in baseline among participants, mixed-effects models were used (lmer R package; Bates, Mächler, Bolker, & Walker, 2014) in *RStudio* version 1.1.456 (RStudio Team, 2016). The dependent measure was the standardised scores. Illusion, contrast polarity (black or white lines/disks) and presentation duration were fixed effects—or predictors—and the nature of their relationships, that is, additive or interactive, was tested through log likelihood-ratio comparisons (χ^2). We also tested the impact of the AQ and O-LIFE scores on the illusory effects. The marginal and conditional r^2 effect sizes are reported as measures of the variance explained by the model with the random effect structure included (conditional r^2) and excluded (marginal r^2) from the calculation (Johnson, 2014; Nakagawa & Schielzeth, 2012). They both were computed from the MuMIn R package.

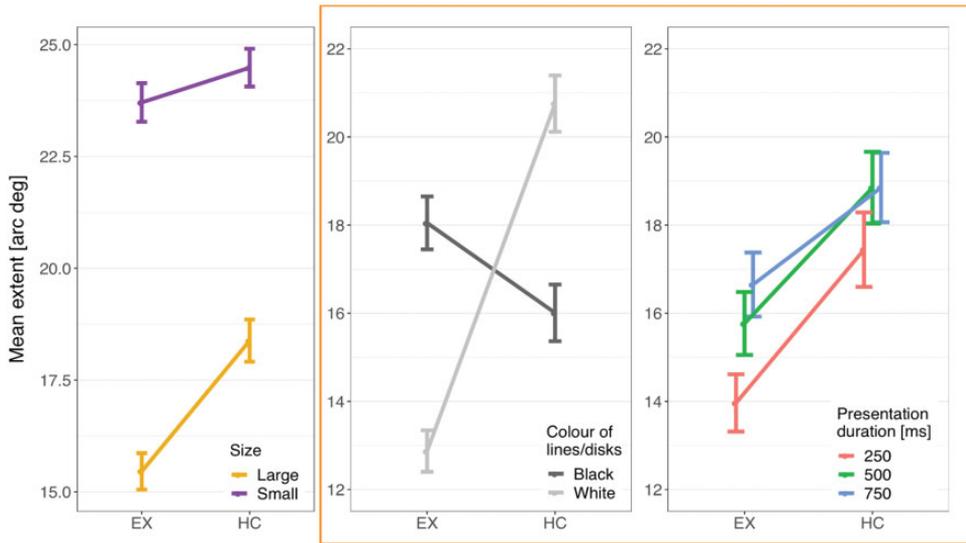


Figure 7. Mean extent of the region (in visual angle) in which lines and disks were visible (left panel) as a function of illusion for small versus large square conditions, (middle panel) as a function of illusion and contrast polarity, and (right panel) as a function of illusion and presentation duration. Only data from the large squares condition were included in the middle and right panels. Error bars show standard error of the mean.

First, the interaction between illusion and contrast polarity was tested. As expected, a likelihood-ratio test revealed a significant difference between additive and interactive models, $\chi^2(1) = 132.470$, $p < .001$, therefore suggesting a significant interaction between illusion and contrast polarity (Figure 7, middle panel).

Second, the model did not significantly improve, $\chi^2(6) = 7.573$, $p = .271$, when presentation duration interacted with illusion and contrast polarity (illusion \times contrast polarity). However, the presentation duration predictor significantly improved the model, $\chi^2(2) = 18.303$, $p < .001$. Hence, there seems to be an effect of presentation duration on the extent of the region in which lines and disks are visible in both illusions (Figure 7, right panel), with a slight increase in the extent of the region visible when presentation duration increases. Third, the total AQ and O-LIFE scores did not significantly improve the model—AQ: $\chi^2(1) = 0.063$, $p = .801$; O-LIFE: $\chi^2(1) = 0.036$, $p = .850$.

The best model therefore included an interaction between illusion and contrast polarity, together with presentation duration. This model accounted for 13.6% of the variance in the data without the random effects, but 58.2% when they were included ($r_m^2 = .136$; $r_c^2 = .582$). The estimates for the individual levels of fixed effects are shown in Table 2.

Correlations

Pairwise correlations were computed between the large square conditions and AQ and O-LIFE scores. Results are reported both with Bonferroni correction and without correction (Table 3). The HC and EX illusions were significantly correlated across contrast polarity and presentation durations, suggesting a similar pattern of individual differences for the two illusions. However, there were no significant correlations between the large square conditions

Table 2. Estimates From The Mixed-Effects Model With Illusion, Contrast Polarity (Black Lines/Disks or White Lines/Disks) and Presentation Duration as Predictors (Interaction Between Illusion and Contrast Polarity).

Fixed Effects	B Estimate	B Standard Error	t Value
Intercept	-0.011	0.108	-0.101
Honeycomb illusion	-0.249	0.071	-3.536
White lines/disks	-0.632	0.071	-8.968
500 ms presentation duration	0.196	0.061	3.210
750 ms presentation duration	0.250	0.061	4.096
Honeycomb Illusion × White Lines/Disks	1.212	0.100	12.157

Only data from large square conditions were included in the model.

and AQ or O-LIFE scores (except for a weak negative correlation between HCWhite500 and O-LIFE score which did however not survive Bonferroni correction: $r = -.266$, $p = .048$).

Discussion

In the HC and the EX illusions, textures that are uniform appear non-uniform. We have introduced a procedure that allows observers to directly report the extent of the visual field over which they were able to see some features (lines or disks). The texture was presented first without lines or disks; then these features were briefly introduced and then removed. After that observers adjusted the radius of a circle to report the percept. This relies on introspection. There are both advantages and weaknesses to this methodology, sometimes referred to as experimental phenomenology (Kubovy, 1999; Kubovy & Gepshtein, 2002).

In our study, it was important to include a baseline condition. We did that by reducing the size of the square elements so that there was no connection between the squares and the lines or disks. We confirmed that this increased the visibility of the features. Therefore, the lines and disks themselves are clearly visible in the periphery and spatial resolution per se is not the reason why the lines are invisible in the HC illusion. Hence, a specific interaction between the squares and the lines must render the lines invisible. The same arguments hold true for the disks in the EX illusion.

In crowding, identification of a target deteriorates in the presence of flankers. Contrast polarity plays a crucial role. Crowding is reduced when flanker and target have opposite contrast polarity compared to when they have the same polarity (Herzog, Sayim, Chicherov, & Manassi, 2015; Kooi, Toet, Tripathy, & Levi, 1994). This role of contrast polarity is similar to what happens with the EX illusion. However, crowding does not render elements invisible. Usually, elements appear as distorted, jumbled or superimposed on each other (Pelli, Palomares, & Majaj, 2004). Although our procedure does not provide direct evidence, we are not aware of any evidence of perceived distortions in the HC or EX illusions. The lines are simply invisible outside a small central region. In addition, crowding is not due to simple flanker target interactions of the type we may have between squares and lines (e.g., Herzog & Manassi, 2015). It may be that the lines are invisible because of some sort of contrast reduction or normalisation, which the larger squares induce. Still, the question arises why the brain does not fill in the lines as it does in some filling-in phenomena or in the healing grid illusion. In the case of filling-in of the blind spot, for example, lines are perceived as complete and not reduced in length (Tripathy, Levi, Ogmen, & Harden, 1995). Patches that differ in colour or motion with respect to the background also are filled in with the background information after fading (Ramachandran & Gregory, 1991), and the lack of

Table 3. Correlations Between Large Square Conditions and AQ and O-LIFE Scores Expressed as Correlation Coefficients (Pearson's r).

r	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	EXBlack250	0.674	0.683	0.717	0.516	0.519	0.745	0.542	0.629	0.493	0.492	0.505	0.047	0.127
2	EXBlack500		0.734	0.604	0.541	0.511	0.623	0.409	0.446	0.469	0.477	0.427	0.146	0.040
3	EXBlack750			0.433	0.375	0.388	0.469	0.348	0.391	0.271	0.386	0.369	0.016	0.050
4	EXWhite250				0.684	0.655	0.725	0.547	0.612	0.559	0.471	0.429	0.129	0.140
5	EXWhite500					0.752	0.569	0.426	0.453	0.473	0.472	0.428	0.080	0.114
6	EXWhite750						0.533	0.485	0.521	0.493	0.442	0.469	-0.048	0.094
7	HCBlack250							0.712	0.768	0.537	0.443	0.340	0.003	0.001
8	HCBlack500								0.818	0.522	0.489	0.364	-0.104	-0.129
9	HCBlack750									0.466	0.464	0.341	-0.066	0.010
10	HCWhite250										0.816	0.703	-0.171	-0.226
11	HCWhite500											0.801	-0.223	-0.266
12	HCWhite750												-0.021	-0.046
13	AQ													
14	O-LIFE													0.678

Note. Light grey indicates significance without any correction, while dark grey indicates significance after Bonferroni correction was applied. Abbreviations are further explained in Table 1.

sensitivity to blue of the human foveola also leads to filling-in (Magnussen, Spillmann, Stürzel, & Werner, 2001). Moreover, a recent study has found that a foveal percept, even when not veridical (central vision filling-in in scotopic vision), was relied upon more than peripheral information (Gloriani & Schütz, 2019). In other words, there are several phenomena where the brain uses available information, in particular at fixation, to fill-in another region of the visual field. In particular, when comparing stimuli near and far from fixation, there is a bias to rely more on the information near fixation (Gloriani & Schütz, 2019). In the HC and EX illusions, we see the reverse, information at fixation is not used for a uniform percept that extends beyond fixation. The mechanism that makes disks and lines invisible in the periphery is stronger than any bias in favour of central vision as well as any prior in favour of uniformity.

In the introduction, we briefly discussed the literature on ensemble perception. In the periphery, similar items may be coded only in terms of summary statistics (Haberman & Whitney, 2012; Whitney & Yamanashi Leib, 2018). In general terms, one could describe the HC and the EX illusions as special cases of ensemble perception. They are special cases because they suggest that the representation of what is in the periphery follows its own rigid rules. It is the rigidity of these processes that creates different representations of the same texture at fixation and away from fixation.

In our study, we have compared directly the HC and the EX illusions. The conditions were similar in that the size of the squares was the same, the task was the same and the trials were interleaved. In line with the original description of these effects, the lines were on each square (HC illusion, Figure 2), while the disks were not present at every intersection (EX illusion, Figure 3). We manipulated presentation time and confirmed that this factor was not critical, although the strength of the illusion decreased with presentation time. We were particularly interested in the role of contrast polarity. As predicted, contrast polarity (white disks among black squares) produced a stronger EX illusion (smaller region of visibility), but contrast polarity (white lines over black squares) produced a weaker HC illusion (larger region of visibility). This type of dissociation suggests that the mechanisms at play are either different in the two cases or, more likely, they dependent on contrast polarity in a complex way that is affected by the spatial relationship between targets (lines and disks) and texture (squares).

Overall in our sample ($N = 56$), the effect was clearer (smaller visibility region) for the EX than HC illusion. This difference is not important as the illusions rely on qualitatively different stimuli. More interesting is the fact that contrast polarity strongly modulated the effect (a cross-over interaction). Responses were consistent over the trials (intrarater reliability was good) and the two illusions correlated with each other. These correlations suggest that individual differences are stable between the two illusions and across changes in presentation duration. Overall, the evidence is that both illusions exist in the majority of individuals, that their effects are larger than in the control conditions when the squares were small, and that they are affected by contrast polarity. The fact that contrast polarity had opposite effects requires further work and provides strong constraints for models. The way information from the periphery relies on summary statistics may play a role here, but we are not aware of a model that can explain the disappearance of some objects and not others, and the way the two illusions depend on contrast polarity.

Finally, we considered individual differences. The sample was entirely constituted by undergraduate students, mostly females. We were interested in whether susceptibility to the illusions (as measured by the size of the visibility region) increases with O-LIFE scores and decreases with AQ scores. The link with schizotypy may be mediated by a response bias toward reporting something that is not visible (Partos et al., 2016), and the link with ASD

spectrum may be mediated by a focus on the local details (Jolliffe & Baron-Cohen, 1997). These hypotheses were not confirmed. Instead, in line with what was found by Russell-Smith, Maybery, and Bayliss (2011), O-LIFE and AQ scores were positively correlated with each other and unrelated to the strength of the illusion. As noted in the introduction, some recent reports have failed to find clear correlations between the strength of illusions and clinical conditions, for instance, schizophrenia and schizotypy (Grzeczowski et al., 2017, 2018).

We conclude by returning to a few general considerations about what makes these illusions so informative and revealing. Both the HC and the EX illusions are counterexamples to a version of the Grand illusion hypothesis. The illusion of seeing a rich and detailed visual field cannot be simply the result of accumulated information over multiple fixations. It is also a counterexample to the idea of an assumption of uniformity for textures, which could explain some of the phenomenal richness of our visual experience. We need to develop more sophisticated models of how visual information determines what observers perceive as extended textures. Surprisingly, even when the textures or patterns are uniform, they may appear as non-uniform.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Marco Bertamini  <https://orcid.org/0000-0001-8617-6864>

References

- Araragi, Y., & Kitaoka, A. (2011). Increment of the extinction illusion by long stimulation. *Perception, 40*, 608–620.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5–17.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R Package Version, 1*, 1–23.
- Bergen, J. R. (1985). Hermann's grid: New and improved. *Investigative Ophthalmology & Visual Science, Supplement, 26*, 280.
- Bertamini, M., Herzog, M. H., & Bruno, N. (2016). The Honeycomb illusion: Uniform textures not perceived as such. *i-Perception, 7*(4), 1–15. doi:10.1177/2041669516660727
- Blackmore, S. J., Brelstaff, G., Nelson, K., & Trościąnko, T. (1995). Is the richness of our visual world an illusion? Transsaccadic memory for complex scenes. *Perception, 24*, 1075–1081. doi:10.1068/p241075
- Chong, S. C., & Treisman, A. (2005). Statistical processing: Computing the average size in perceptual groups. *Vision Research, 45*, 891–900.
- Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience? *Trends in Cognitive Sciences, 20*, 324–335.
- Crespi, B., & Badcock, C. (2008). Psychosis and autism as diametrical disorders of the social brain. *Behavioral and Brain Sciences, 31*, 241–261. doi:10.1017/S0140525X08004214

- de Fockert, J., Davidoff, J., Fagot, J., Parron, C., & Goldstein, J. (2007). More accurate size contrast judgments in the Ebbinghaus Illusion by a remote culture. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 738–742. doi:10.1037/0096-1523.33.3.738
- Dehaene, S. (2014). *Consciousness and the brain*. New York, NY: Viking Press.
- Dennett, D. C. (1991). *Consciousness explained*. Boston, MA: Little Brown.
- Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, *14*, 1195–1201.
- Geier, J., Bernáth, L., Hudák, M., & Séra, L. (2008). Straightness as the main factor of the Hermann grid illusion. *Perception* *37*, 651–665.
- Gibson, J. J. (1950). *Perception of the visual world*. Cambridge, England: Riverside Press.
- Gloriani, A. H., & Schütz, A. C. (2019). Humans trust central vision more than peripheral vision even in the dark. *Current Biology*, *29*, 1206–1210.e4. doi:10.1016/j.cub.2019.02.023
- Gori, S., Molteni, M., & Facoetti, A. (2016). Visual illusions: An interesting tool to investigate developmental dyslexia and autism spectrum disorder. *Frontiers in Human Neuroscience*, *10*, 175.
- Grzeczowski, L., Clarke, A. M., Francis, G., Mast, F. W., & Herzog, M. H. (2017). About individual differences in vision. *Vision Research*, *141*, 282–292. doi:10.1016/j.visres.2016.10.006
- Grzeczowski, L., Roinishvili, M., Chkonia, E., Brand, A., Mast, F. W., Herzog, M. H., & Shaqiri, A. (2018). Is the perception of illusions abnormal in schizophrenia? *Psychiatry Research*, *270*, 929–939.
- Haberman, J., & Whitney, D. (2012). Ensemble perception. In J. Wolfe & L. Robertson (Eds.), *From perception to consciousness* (pp. 339–349). Oxford, England: Oxford University Press. doi:10.1093/acprof:osobl/9780199734337.003.0030
- Happé, F. G. E. (1996). Studying weak central coherence at low levels: Children with autism do not succumb to visual illusions. A research note. *Journal of Child Psychology and Psychiatry*, *37*, 873–877. doi:10.1111/j.1469-7610.1996.tb01483.x
- Helmholtz, H. (1867). *Treatise on physiological optics (Vol. III)*. Mineola, NY: Dover Publications, Inc.
- Hermann, L. (1870). Eine Erscheinung simultanen Kontrastes [Translation: A phenomenon of the simultaneous contrast]. *Pflügers Archiv für die gesamte Physiologie*, *3*, 13–15.
- Herzog, M. H., & Manassi, M. (2015). Uncorking the bottleneck of crowding: A fresh look at object recognition. *Current Opinion in Behavioral Sciences*, *1*, 86–93.
- Herzog, M. H., Sayim, B., Chicherov, V., & Manassi, M. (2015). Crowding, grouping, and object recognition: A matter of appearance. *Journal of Vision*, *15*, 5.
- Johnson, P. C. D. (2014). Extension of Nakagawa & Schielzeth's R^2 GLMM to random slopes models. *Methods in Ecology and Evolution*, *5*, 944–946. doi:10.1111/2041-210X.12225
- Jolliffe, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the embedded figures test. *Journal of Child Psychology and Psychiatry*, *38*, 527–534. doi:10.1111/j.1469-7610.1997.tb01539.x
- Kaliuzhna, M., Stein, T., Rusch, T., Sekutowicz, M., Sterzer, P., & Seymour, K. J. (2018). No evidence for abnormal priors in early vision in schizophrenia. *Schizophrenia Research*. doi:10.1016/j.schres.2018.12.027
- Kanai, R. (2005). *Best illusion of the year contest: Healing grid*. Retrieved from <http://illusionoftheyear.com/2005/08/healing-grid/>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*, 155–163. doi:10.1016/j.jcm.2016.02.012
- Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision*, *8*, 255–279.
- Kubovy, M. (1999). Gestalt psychology. In *The MIT encyclopedia of the cognitive sciences* (pp. 346–349). Cambridge, MA: MIT Press.
- Kubovy, M., & Gepshtein, S. (2002). Perceptual grouping in space and space-time: An exercise in phenomenological psychophysics. In R. Kimchi, M. Behrmann, & C. R. Olson (Eds.), *Perceptual organization in vision: Behavioral and neural perspectives* (pp. 45–85). Mahwah, NJ: Erlbaum.
- Lamme, V. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience*, *1*, 204–220.

- Levine, M. W., Anderson, J. E., & McAnany, J. J. (2012). Effects of orientation and contrast upon targets in straight and curved arrays. *Perception*, *41*, 1419–1433.
- Levine, M. W., & McAnany, J. J. (2008). The effects of curvature on the grid illusions. *Perception*, *37*, 171–184. doi:10.1068/p5691
- Magnussen, S., Spillmann, L., Stürzel, F., & Werner, J. S. (2001). Filling-in of the foveal blue scotoma. *Vision Research*, *41*, 2961–2967. <https://doi.org/10/dm5fz3>
- Manning, C., Morgan, M. J., Allen, C. T. W., & Pellicano, E. (2017). Susceptibility to Ebbinghaus and Müller-Lyer illusions in autistic children: A comparison of three different methods. *Molecular Autism*, *8*, 16. doi:10.1186/s13229-017-0127-y
- Mason, O., Linney, Y., & Claridge, G. (2005). Short scales for measuring schizotypy. *Schizophrenia Research*, *78*, 293–296. doi:10.1016/j.schres.2005.06.020
- McAnany, J. J., & Levine, M. W. (2004). The blanking phenomenon: A novel form of visual disappearance. *Vision Research*, *44*, 993–1001.
- McClelland, T., & Bayne, T. (2016). Ensemble coding and two conceptions of perceptual sparsity. *Trends in Cognitive Sciences*, *20*, 641–642.
- Nakagawa, S., & Schielzeth, H. (2012). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142. doi:10.1111/j.2041-210x.2012.00261.x
- Ninio, J., & Stevens, K. A. (2000). Variations on the Hermann grid: An extinction illusion. *Perception*, *29*, 1209–1217. doi:10.1068/p2985
- Noë, A., Pessoa, L., & Thompson, E. (2000). Beyond the grand illusion: What change blindness really teaches us about vision. *Visual Cognition*, *7*, 93–106. doi:10.1080/135062800394702
- Notredame, C.-E., Pins, D., Deneve, S., & Jardri, R. (2014). What visual illusions teach us about schizophrenia. *Frontiers in Integrative Neuroscience*, *8*, 63. doi:10.3389/fnint.2014.00063
- Otten, M., Pinto, Y., Paffen, C. L. E., Seth, A. K., & Kanai, R. (2017). The uniformity illusion central stimuli can determine peripheral perception. *Psychological Science*, *28*, 56–68. doi:10.1177/0956797616672270
- Partos, T. R., Cropper, S. J., & Rawlings, D. (2016). You don't see what I see: Individual differences in the perception of meaning from visual stimuli. *PLoS One*, *11*, e0150615. doi:10.1371/journal.pone.0150615
- Peirce, J. W. (2007). PsychoPy: Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8–13.
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, *4*, 1136–1169.
- Peterzell, D. H., & Kennedy, J. F. (2016). Discovering sensory processes using individual differences: A review and factor analytic manifesto. *Electronic Imaging*, *2016*, 1–11. doi:10.2352/ISSN.2470-1173.2016.16.HVEI-112
- Qian, K., & Mitsudo, H. (2016). Eggs illusion: Local shape deformation generated by a grid pattern. *Journal of Vision*, *16*, 27. doi:10.1167/16.15.27
- Qian, K., Yamada, Y., Kawabe, T., & Miura, K. (2009). The scintillating grid illusion: Influence of size, shape, and orientation of the luminance patches. *Perception*, *2009*, *38*, 1172–1182.
- Ramachandran, V. S., & Gregory, R. L. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature*, *350*, 699–702.
- Read, J. C. A., Robson, J. H., Smith, C. L., & Lucas, A. D. (2012). The scintillating grid illusion is enhanced by binocular viewing. *i-Perception*, *3*, 820–830.
- Rensink, R. A., O'Regan, K., & Clark, J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*, 368–373.
- RStudio Team. (2016). *RStudio: Integrated development for R*. Boston, MA: RStudio, Inc. <http://www.rstudio.com/>
- Russell-Smith, S. N., Maybery, M. T., & Bayliss, D. M. (2011). Relationships between autistic-like and schizotypy traits: An analysis using the Autism Spectrum Quotient and Oxford-Liverpool Inventory of Feelings and Experiences. *Personality and Individual Differences*, *51*, 128–132. doi:10.1016/j.paid.2011.03.027

- Schiller, P. H., & Carvey, C. E. (2005). The Hermann grid illusion revisited. *Perception, 34*, 1375–1397. doi:10.1068/p5447
- Schrauf, M., Lingelbach, B., Lingelbach, E., & Wist, E. R. (1995). The Hermann grid and the scintillation effect. *Perception, 24*, 88–89.
- Schrauf, M., Lingelbach, B., & Wist, E. R. (1997). The scintillating grid illusion. *Vision Research, 37*, 1033–1038.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420–428. doi:10.1037//0033-2909.86.2.420
- Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., & Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Research, 49*, 2705–2739.
- Slaghuis, W. L., Twell, A. J., & Kingston, K. R. (1996). Visual and language processing disorders are concurrent in dyslexia and continue into adulthood. *Cortex, 32*, 413–438. doi:10.1016/S0010-9452(96)80002-5
- Toscani, M., Gegenfurtner, K. R., & Valsecchi, M. (2017). Foveal to peripheral extrapolation of brightness within objects. *Journal of Vision, 17*, 14. doi:10.1167/17.9.14
- Tripathy, S., Levi, D., Ogmen, H., & Harden, C. (1995). Perceived length across the physiological blind spot. *Visual Neuroscience, 12*, 385–402. doi:10.1017/S0952523800008051
- Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology, 69*, 105–129. doi:10.1146/annurev-psych-010416-044232
- Yu, Q., & Shim, W. M. (2016). Modulating foveal representation can influence visual discrimination in the periphery. *Journal of Vision, 16*, 15. doi:10.1167/16.3.15

How to cite this article

Bertamini, M., Cretenoud, A. F., & Herzog, M. H. (2019). Exploring the Extent in the Visual Field of the Honeycomb and Extinction Illusions. *i-Perception, 10*(4), 1–19. doi:10.1177/2041669519854784

Curriculum Vitae

PERSONAL INFORMATION

Name Aline Cretenoud
Date of Birth 6th of September 1992
Citizenship Swiss
Address EPFL SV BMI LPSY, Station 19
1015 Lausanne, Switzerland
Phone number +41 21 693 72 27
Email aline.cretenoud@epfl.ch

EDUCATION

Ecole Polytechnique Fédérale de Lausanne (EPFL)
2017 – 2021 Ph. D. in Neuroscience
2014 – 2016 M. Sc. in Life Sciences and Technologies
Mention of Excellence (GPA: 5.50/6)
2011 – 2014 B. Sc. in Life Sciences and Technologies

RESEARCH AND PERSONAL SKILLS

Statistical analysis (R, MATLAB, Python, JASP, and GPower)

- Frequentist and Bayesian
- Structural Equation Modeling (SEM)
- Machine learning

Psychophysical experiments

- Design and setup
- Programming (MATLAB and Psychtoolbox)
- Data collection

Writing manuscripts and reports

Organization and supervision of a stand about visual perception at four public events promoting sciences with 20,000+ visitors

TEACHING EXPERIENCES

Ecole Polytechnique Fédérale de Lausanne (EPFL)
2019 – 2021 Supervised a Master student and trained graduate students
2013 – 2019 Teaching assistant for statistics, mathematics, and physics classes (470+ hours)
Etablissement Primaire et Secondaire (EPS), Ecublens
2016 – 2017 Teacher substitute at secondary school for mathematics, physics, and chemistry

LANGUAGES

French (mother tongue), English (fluent), German (elementary)

JOURNAL ARTICLES

Cretenoud, Grzeczkowski, Kunchulia, & Herzog (submitted). Individual differences in the perception of visual illusions are stable across eyes, time, and measurement methods.

Cretenoud, Barakat, Milliet, Choung, Bertamini, Constantin, & Herzog (submitted). How do visual skills relate to action video game performance?

Cretenoud, Francis, & Herzog (2020). When illusions merge. *Journal of Vision*, 20(8):12, 1–15.

Cretenoud, Grzeczkowski, Bertamini, & Herzog (2020). Individual differences in the Müller-Lyer and Ponzo illusions are stable across different contexts. *Journal of Vision*, 20(6):4, 1–14.

Rashal, **Cretenoud**, & Herzog (2020). Perceptual grouping leads to objecthood effects in the Ebbinghaus illusion. *Journal of Vision*, 20(8):11, 1–15.

Cretenoud, Karimpur, Grzeczkowski, Francis, Hamburger, & Herzog (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, 19(14):12, 1–21.

Shaqiri, Pilz, **Cretenoud**, Neumann, Clarke, Kunchulia, & Herzog (2019). No evidence for a common factor underlying visual abilities in healthy older people. *Developmental Psychology*, 55(8), 1775–1787.

Bertamini, **Cretenoud**, & Herzog (2019). Exploring the Extent in the Visual Field of the Honeycomb and Extinction Illusions. *i-Perception*, 10(4), 1–19.

Grzeczkowski, **Cretenoud**, Mast, & Herzog (2019). Motor response specificity in perceptual learning and its release by double training. *Journal of Vision*, 19(6):4, 1–14.

Grzeczkowski, **Cretenoud**, Herzog, & Mast (2017). Perceptual learning is specific beyond vision and decision making. *Journal of Vision*, 17(6):6, 1–11.

Herzog, **Cretenoud**, & Grzeczkowski (2017). What is new in perceptual learning? *Journal of Vision*, 17(1):23, 1–4.

SCIENTIFIC TALKS

Cretenoud & Herzog (2019). When illusions merge. 19th Vision Sciences Society (VSS) Annual Meeting, St Pete Beach, Florida, USA.

Cretenoud, Grzeczkowski, & Herzog (2019). Individual differences in vision: do we all perceive the world in the same way? Meeting of the NeuroLeman Network and Doctoral Schools, Les Diablerets, Switzerland.

CONFERENCE POSTERS

Cretenoud, Ramos da Cruz, Favrod, Shaqiri, Roinishvili, Brand, Chkonia, & Herzog (2019). Schizophrenia: when no difference makes the difference. NCCR-Synapsy, 9th meeting, Villars, Switzerland.

Cretenoud, Grzeczkowski, & Herzog (2018). How to classify visual illusions? The role of intellectual development. 18th Vision Sciences Society (VSS) Annual Meeting, St Pete Beach, Florida, USA.

Cretenoud, Grzeczkowski, Francis, & Herzog (2018). About specific and general factors for visual illusions. 41st European Conference on Visual Perception (ECVP), Trieste, Italy.

- Herzog, **Cretenoud**, Francis, Shaqiri, & Grzeczowski (2018). The Individual Structure of Visual Space. 41st European Conference on Visual Perception (ECPV), Trieste, Italy.
- Rashal, **Cretenoud**, & Herzog (2018). The effect of “objecthood” in the Ebbinghaus illusion. 41st European Conference on Visual Perception (ECPV), Trieste, Italy.
- Cretenoud**, Karimpur, Grzeczowski, Francis, Hamburger, & Herzog (2017). How many factors are there in vision? 40th European Conference on Visual Perception (ECPV), Berlin, Germany.
- Rashal, **Cretenoud**, & Herzog (2017). The effect of “objecthood” in the Ebbinghaus illusion. 20th Conference of the European Society for Cognitive Psychology (ESCoP), Potsdam, Germany.
- Herzog, **Cretenoud**, Francis, & Grzeczowski (2017). The Structure of Visual Space. 17th Vision Sciences Society (VSS) Annual Meeting, St Pete Beach, Florida, USA.
- Grzeczowski, **Cretenoud**, Mast, & Herzog (2017). Double training reduces motor response specificity. 17th Vision Sciences Society (VSS) Annual Meeting, St Pete Beach, Florida, USA.