Augmenting Learning Activities with Contextual Information Scent

THÈSE N^{O} 6767 (2015)

PRÉSENTÉE LE 27 NOVEMBRE 2015

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS
LABORATOIRE D'ERGONOMIE ÉDUCATIVE
PROGRAMME DOCTORAL EN INFORMATIQUE ET COMMUNICATIONS

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Nan LI

acceptée sur proposition du jury:

Prof. K. Aberer, président du jury
Prof. P. Dillenbourg, directeur de thèse
Prof. P. Olivier, rapporteur
Prof. K. Reinecke, rapporteuse
Prof. D. Gillet, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2015

The will to win,
the desire to succeed,
the urge to reach your full potential…
these are the keys that will unlock the door to personal excellence.
— Confucius

Acknowledgements

This thesis work would not have been possible without the help of many people.

First and foremost I want to thank my thesis supervisor **Prof. Pierre Dillenbourg** for the support and guidance throughout the past five years. I experienced some difficulty at the beginning of my PhD, but you trusted me and guided me through the "darkness". It was you who introduced me to the new but exciting field of MOOC research, where I have learned a lot and made my own research contributions. I could not have had a more patient and supportive supervisor. Thank you for giving me the opportunity to purse a PhD under your supervision. You really have changed my life.

Patrick Jermann, thanks very much for being my "part-time supervisor" for five years, your feedback has always been valuable. **Florence Colomb**, thanks for your help in too many aspects of my life at EPFL and in Switzerland. You are always very patient and nice to me.

Frédéric Kaplan and **Omar Mubin**, thank you for your support and advices during the first years of my Phd. I also would like to thank **Guillaume Zufferey** and **Jan Blom**, you both helped me a lot with my experiments. **Łukasz Kidziński**, thanks for your support in my data analysis and paper writing. My PhD research got a boost since you came to the lab and worked together with me.

Hamed Alavi and **Son Do-Lenh**, I often asked advices from you, and you both have spent very much time in helping me with my research. I learned a lot from you. **Quentin Bonnard** and **Olivier Guédat**, thank you for your technical help when I was working on the Tinkerlamp project. My thanks also go to **Andrea Mazzei** and **Sebastien Cuendet**, you also gave me lots of advices for my research.

Julia Fink and **Himanshu Verma**, I was the first among us to arrive at the lab but I leave the last. We had very good memories in the lab, I still remember we attended the course together in our first year. It is my pleasure to have you as my colleagues. **Kshitij Sharma** and **Mirko Raca**, thank you for helping me with many things from time to time, especially your feedback of talk.

Many thanks go to my thesis juries, including **Prof. Karl Aberer**, **Prof. Denis Gillet**, **Prof. Patrick Olivier** and **Prof. Katharina Reinecke**. My thanks also go to the people who have worked closely with me for some projects, including **Andrei Popescu-Belis**, **Phil Garner**, **Afroditi Skevi** and **Tanmay Sinha**. I also thank all the people who have been willing to participate in my experiments.

I would also like to thank many of my colleagues in CHILI, CEDE and CAPE for your support during my PhD, including (in no particular order) **Luis Pablo Prieto Santos**, **Beat Adrian Schwendimann**, **Lorenzo Lucignano**, **Ayberk Özgür**, **Mina Shirvani Boroujeni**, **Flaviu Roman**, **Francisco Pinto**, **Jeremy Rieder**, **Gwénaël Bocquet**, **Ingrid Le Duc**, **Roland Tormey** and **Siara Isaac**.

Finally, I must express my gratitude to my family, especially to **Shuheng**, for your continued support and encouragement during the past few years. You have made me become what I am.

Lausanne, 23 October 2015

Abstract

Students often have information needs while carrying out a multitude of learning activities at universities. When information is needed for investigating a problem, the student may interrupt the work and switch to an information seeking task. As Internet connectivity becomes ubiquitous, searching information has been routinized and integrated in the learning experience. However, information needs are not always fully recognized, or they can not be well articulated. A MOOC student may perceive a video to be difficult, but fails to express what information can be helpful. Sometimes it is improper to interrupt the learning task for searching information, especially when social factors are concerned, e.g. in a seminar talk. These situations create research potentials for making ambient information cues, hereafter referred to as *contextual information scent (CIS)*, available to address students' situational information needs in learning activities. The CIS is designed to combine context-awareness with information seeking, ambient interaction as well as serendipitous encounter.

In this thesis, we investigate the CIS mainly in collaborative learning activities. We explore three different contexts: conversation, groupware interaction and video content for MOOC learning. *RaindropSearch* investigates capturing conversational words as CIS for building search queries, while the *TileSearch* triggers Web searches based on group discussions and retrieved image and Wikipedia results as CIS for serendipitous interactions. These two explorations both focus on conversation context and provide initial insights into the CIS design practice. Next, we present *MeetHub Search*, which includes three CIS components based on text interactions in a groupware. Our last prototype, the *BOOC Player* employs textbook pages as CIS and links them to MOOC videos during the course of collaborative video viewing. All prototypes show how we manipulated design parameters to reduce distraction, increase relevance and ensure timeliness. The studies also exhibit the influence of group dynamics on the use of CIS. We finally extend our research scope to individual MOOC learning and summarize the design insights obtained from MOOC analytics.

The contributions of this thesis are summarized as (1) a dedicated research framework derived from both research literature and requirement analysis for recognizing the design challenges, design principles and design space of CIS. The framework lays the foundation for us to explore different contexts in this thesis, where we generated (2) design implications that identify the key attributes of CIS. Last but not least, we employed (3) a variety of evaluation methodologies in this thesis for assessing the usability as well as the benefit and appeal of CIS.

Key words: context-aware, ambient information, serendipity, collaborative learning, MOOC

Résumé

Les étudiants ressentent souvent un besoin d'information quand ils font face à une multitude d'activités d'apprentissage dans les universités. Lorsque l'information est nécessaire pour résoudre un problème, l'élève peut interrompre son travail et prendre le temps de chercher les éléments dont il a besoin. Alors qu'Internet devient omniprésent, la recherche d'information en ligne s'intègre de plus en plus dans l'apprentissage au quotidien. Cependant, les besoins d'information ne sont pas toujours pleinement reconnus et peuvent parfois être mal traités. Pour un étudiant inscrit à un MOOC il est possible de percevoir la difficulté d'une vidéo sans parvenir à identifier l'information qui peuvent être utile. Dans d'autres cas il n'est pas concevable d'interrompre la tâche d'apprentissage pour la recherche d'informations, par exemple, lors d'une présentation d'un séminaire. Ces situations créent des potentiels de recherche pour la fabrication d'indices ambiants d'information, ci-après dénommés contextual information scent (CIS). Ces indices sont disponibles pour répondre aux besoins d'informations des élèves en fonction de la situation. Les CIS s'adaptent au contexte afin de favoriser l'accès à l'information.

Dans cette thèse, nous étudions les CIS principalement dans le cadre d'activités d'apprentissage collaboratives. Nous explorons trois contextes différents : la conversation, l'interaction à travers un groupware et le contenu vidéo pour l'apprentissage sur les MOOCs. *RaindropSearch* cherche à capturer des mots dans une conversation en tant que CIS pour aider à construire des requêtes de recherche, tandis que *TileSearch* déclenche des recherches sur le Web basées sur des discussions de groupe et affiche des images et du contenu récupérés sur Wikipédia. Ces deux études se concentrent toutes deux sur le contexte de la conversation et fournissent un premier aperçu de la pratique de la conception de CIS. Ensuite, nous présentons *MeetHub Search*, qui comprend trois composantes de CIS basées sur des échanges de texte d'un groupware. Notre dernier prototype, le *BOOC Player* affiche des pages de manuels scolaires en tant que CIS et les relie aux vidéos d'un MOOC dans le cadre de visionnements en groupe. Tous nos prototypes montrent comment nous avons manipulé les paramètres de conception pour réduire les distractions, accroître la pertinence et favoriser les opportunités de recherche. Les études montrent également l'influence de la dynamique de groupe sur l'utilisation de CIS. Nous avons finalement étendu notre champ de recherche à l'apprentissage individuel sur les MOOCs et nous présentons nos idées de conception obtenus à partir de notre analyse des MOOCs.

Les contributions de cette thèse peuvent se résumer à (1) un cadre dédié issu de la littérature et

de l'analyse des besoins pour reconnaître les défis de conception et les principes de conception des CIS. Le cadre établit les bases d'explorations pour différents contextes, qui fournissent (2) des implications sur la conception et identifient les principaux attributs des CIS. Finalement, la thèse emploie (3) une variété de méthodes pour évaluer la facilité d'utilisation ainsi que le bénéfice global et l'attrait des CIS.

Mots clefs : sensible au contexte, information ambiante, sérendipité, apprentissage en groupe, MOOC

Contents

List of Figures

List of Tables

1 Introduction

It is a regular Monday morning. Alice, a university student in Computer Science starts her campus week by attending a seminar about "Big Data". She is very interested in this trending topic but is new to it. "Data warehouse? Is he referring to the database?" Alice is uncertain about this term when she hears it from the speaker, who does not give further explanation. At first, she leaves it alone. Soon she finds the term is mentioned from time to time during the presentation, so she brings her laptop and looks for an explanation online. The web search temporarily deviates her attention away from the talk. When Alice finishes the information seeking, she finds it more difficult to catch up with the presentation.

After attending the seminar, Alice meets up with her friends to study a flipped HCI course in a group. Following the last week's lectures, their group must brainstorm and finalize a project idea to work on for the semester. "How about a restaurant finder application for travelers?" Alice says. "Can you articulate your idea?", Elisabeth asks. "Sure! I mean, we can design a service similar to Booking.com, but focus on restaurant, and with more social features.", answers Alice. But Elisabeth and others still do not get her point. Alice then takes her pen out and starts sketching an interface. Suddenly, she stops and utters: "Well, let me illustrate my idea with some concrete examples". She opens the website of Booking.com and Facebook on her laptop and compares them. The whole brainstorming session then goes like this, i.e. switching between web searching and discussion. After they agree on the project, the group starts watching the lecture videos together. "I don't get the difference between storyboarding and scenario mapping, do you?" Alice paused the video and asks this question in the group. After a heated debate on this issue, they do not reach an agreement. The group then tries to look for external help. Five minutes later, Alice finds an answer from the textbook and shares it with the others.

In the evening, Alice decides to follow a Web programming MOOC to acquire necessary skills for her HCI course project. She studies the subject alone, because she has never done any web programming before. She feels the syntax awkward compared to the language she is familiar with. Every now and then, she replays some specific video segments multiple times. Sometimes she also checks online documentation and guidelines in order to understand the code examples in the video.

1.1 Motivation

The above story illustrates four learning scenarios that are part of many students' life at universities, including seminar talks, collaborative brainstorming, collaborative lecture video viewing as well as online learning. A common plot in all these scenarios is Alice from time to time looks for information in various learning tasks. The information being seeked, however, serves different purposes. For example, in the seminar scenario, Alice is unfamiliar with a terminology which she thinks is crucial for understanding the presentation; while in the brainstorming scenario, the web search is used to support articulating her ideas. When we look at Alice's learning experiences in the environments that do not involve more advanced technologies than an ordinary laptop, we see that she often interrupts the learning task to seek information. In some scenarios (e.g. collaborative brainstorming, online learning etc.), such interruptions are a matter of "inefficiency", while in other situations (e.g. the seminar talk), interruptions may lead to frustrations for not being able to catch up with the task, which deteriorates the learning experiences.

In fact, human beings have the capability to perform multiple tasks at a time. For example, we can listen to music while jogging or drink while writing a report. These activities happen routinely in our everyday life, and we may not even realize they are happening at the same time. There is no apparent problem in performing them simultaneously, because one of such parallel tasks usually requires no conscious attention, allowing a person to carry it out in the periphery. In contrast, the information seeking in the described scenarios often requires focused attention, so interruptions and attention shifts are unavoidable given existing technologies. Numerous studies (Adamczyk and Bailey, 2004; Czerwinski et al., 2004; González and Mark, 2004; Mark et al., 2005) have recognized managing multiple tasks and interruptions as a challenge for information workers. As our interests lie in the interruptions for information seeking in learning activities, the presented issue easily provokes one to think "What if ready-to-use information is timely prompted to the learners? "

1.2 Calm Technology

Timely prompting useful information implies systems should first be able to capture the context and then appropriately interrupt the task with relevant information. These ideas are not new in human computer interaction, and can be traced back to Mark Weiser's vision of *ubiquitous computing* (Weiser, 1991), also called *pervasive computing* (Satyanarayanan, 2001). It describes "*The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.*" The essence of this vision is however not the physical disappearance of technologies, but rather the seamless integration of technologies into everyday routines so that the users do not necessarily perceive them. Put it in other words, this vision implies that future technologies should work in the background, without demanding focused attention. Weiser and Brown (1997) further articulated this vision and coined the term *calm technology* as a way to minimize the invasiveness of computers.

The concept of *calm technology* has inspired researchers to investigate computing powers that are deployed in the users' periphery of attention, thus breeding the notion of *ambient display* (Mankoff et al., 2003; Stasko et al., 2005) and *peripheral interaction* (Bakker et al., 2012; Hausen and Butz, 2011). The former focuses on subtly presenting information without distracting or burdening the user, whilst the latter explores physical interactions taken place in the periphery of attention. Both concepts emphasize subconscious perception and interaction rather than context-awareness. That is, the information being interacted with may not subordinate to the main task. For example, interactions with an ambient time indicator or a stress monitor while engaging in a task do not directly help accomplish the task.

Devices that work in concert to seamlessly support people in carrying out their activities and tasks are embraced with *ambient intelligence (AmI)* (Zelkha et al., 1998). AmI attempts to capture "all information that is available for a distinct user with a conscious or subconscious desire at a certain place and time" (Lugmayr et al., 2009). Compared to *ambient display* and *peripheral interaction*, this paradigm underlines the pervasiveness of embedded intelligent systems as well as the interconnections between them, whereas unobtrusive task interruptions are not the focus.

1.3 Contextual Information Scent

Calm technologies are seemingly the right paradigm we are looking for, but existing terminologies as described before are all umbrella terms that are either too general or emphasize slightly different aspects. Before proceeding to the main body of our research, we pursue a terminology to frame the context of this thesis, i.e. supporting information behaviors in learning tasks.

Information foraging (Pirolli and Card, 1999) is one of the most important concepts. As users traverse the Web, they encounter "trigger Web elements" that they perceive as meaningful to the task. Such trigger elements may have different forms, such as texts, images or links that drive the users towards the right direction for finding useful information. Information foraging theory describes such information hunting behaviors with an analogy of wild animals hunting for food by following the scent. In this analogy, the "trigger Web elements" become *information scent*, which are proximal cues that help the users to judge distal information sources and to navigate towards them.

Adapted from the concept of *information scent*, *contextual information scent (CIS)* is defined as **ambient information cues, which are perceived from the activity context and made available for a user with a conscious or subconscious situational information need.** The form of CIS is not defined, so theoretically the scents can be of any media of any level of details, e.g. images, videos, Web pages, book pages or text. The gist is that the CIS is provisioned through a smart media environment which aggregates external content with the semantics of certain user contexts, so that the users may potentially be navigated towards the information that may help with the learning task.

1.4 Research Objectives

The goal of designing contextual information scent is to make information interaction an integral part of the learning activity, by enabling effortless access to information pertinent to the situational context. This research field is scarcely explored in literature, so the foremost issue, among other things, is to identify the types of situational information needs and explore the design space for proper support. Second, we aimed at exploring how information can be designed as scents to augment the activities in one's periphery of attention. Finally, with the prevalence of MOOCs in recent years, we intended to deliver data-driven insights in designing contextual information scent for this online learning environment. To be specific, this dissertation aimed to answer the following research questions:

1. What types of situational information needs may arise during learning activities and what are the challenges, principles and potential design space for augmenting the activities with contextual information scent?
2. How can ambient technologies be designed as contextual information scent and what are the benefits and overall appeal of them?
3. Can big educational data (MOOC) provide insights for designing contextual information scent?

1.5 Research Approach

All the research questions have been addressed by different empirical studies or analyses of research literature.

Research Question 1: We first conducted a literature review covering the following aspects: (1) theories about human cognitive capabilities and limitations for attending to multiple tasks and stimuli either concurrently or sequentially; (2) conceptual frameworks and theories about activities as well as learning; (3) well-established information seeking and interaction models, and (4) related interactive systems that applied the discussed theories for presenting or interacting with information. Then we initiated a survey aimed at understanding situational information needs and classifying them. The survey study was conducted with lab participants of six seminar talks, due to the convenience of data collection for these activities.

Based on the literature review and the findings from the survey study, we identified the design challenges, which are in turn transformed to a set of design principles. These principles further informed the design space for contextual information scent. The design space contains five axes : *Privacy*, *Context*, *Information Capacity*, *Information Uncertainty* and *Activation*, each of which includes two or three conditions.

Research Question 2: Given numerous potential solutions suggested by the design space, this thesis cannot deliver an exhaustive exploration on all design possibilities. To answer the research question 2, we built four different prototypes that are derived from design space, and then evaluated them in collaborative learning settings with lab experiments. Both qualitative and quantitative methods were employed in the evaluation. In addition, thorough analyses with the Activity Theory is also presented in order to explain the successful and failed aspects of the designs.

Research Question 3: This research question was addressed by large-scale behavioral data analyses on video lecture viewing, which is the central MOOC learning activity. We associated perceived video difficulty with individual student's video interactions, and identified the occasions when information scent in terms of contextual help can be prompted. Research in this section do not include the implementation and evaluation of concrete designs, but rather shed some light on potential design solutions that could be experimented in future work.

1.6 Thesis Outline

Figure 1.1 illustrates the outline of this thesis, which is divided into 6 block chapters. Each block contains one to three chapters.

Chapter 2, 3 and 4 aim at anchoring contextual information scent by reviewing related work in literature. In **Chapter 2** we first introduce human attention and cognition theories, including selective attention, divided attention, multi-tasking and priming. Then we present a second-wave HCI theory, the Activity Theory and its analytical tools for structuring context. Finally, fundamental learning theories as well as collaborative theories are reviewed. **Chapter 3** starts with a discussion about information seeking behaviors, including its models, principles, contexts and modes. This is followed by a review of help seeking theories in learning context. **Chapter 4** introduces a rich body of theories in ambient information and serendipitous information. We also show example applications for ambient interaction and serendipitous encountering, with a focus on systems designed for learning.

In **Chapter 5** we present a survey study conducted in seminar talks to learn about information needs in learning activities. Based on the research literature and insights obtained from the survey, we further develop a research framework and identify the design challenges, principles and design space of contextual information scent. This chapter answers the first research question of the thesis.

Chapter 6, 7, 8 present a few design prototypes with user studies for exploring the design space with three contexts: conversation, group interaction and MOOC video content. In each exploration, we manipulate the design parameters to reduce distractions, increase relevance and

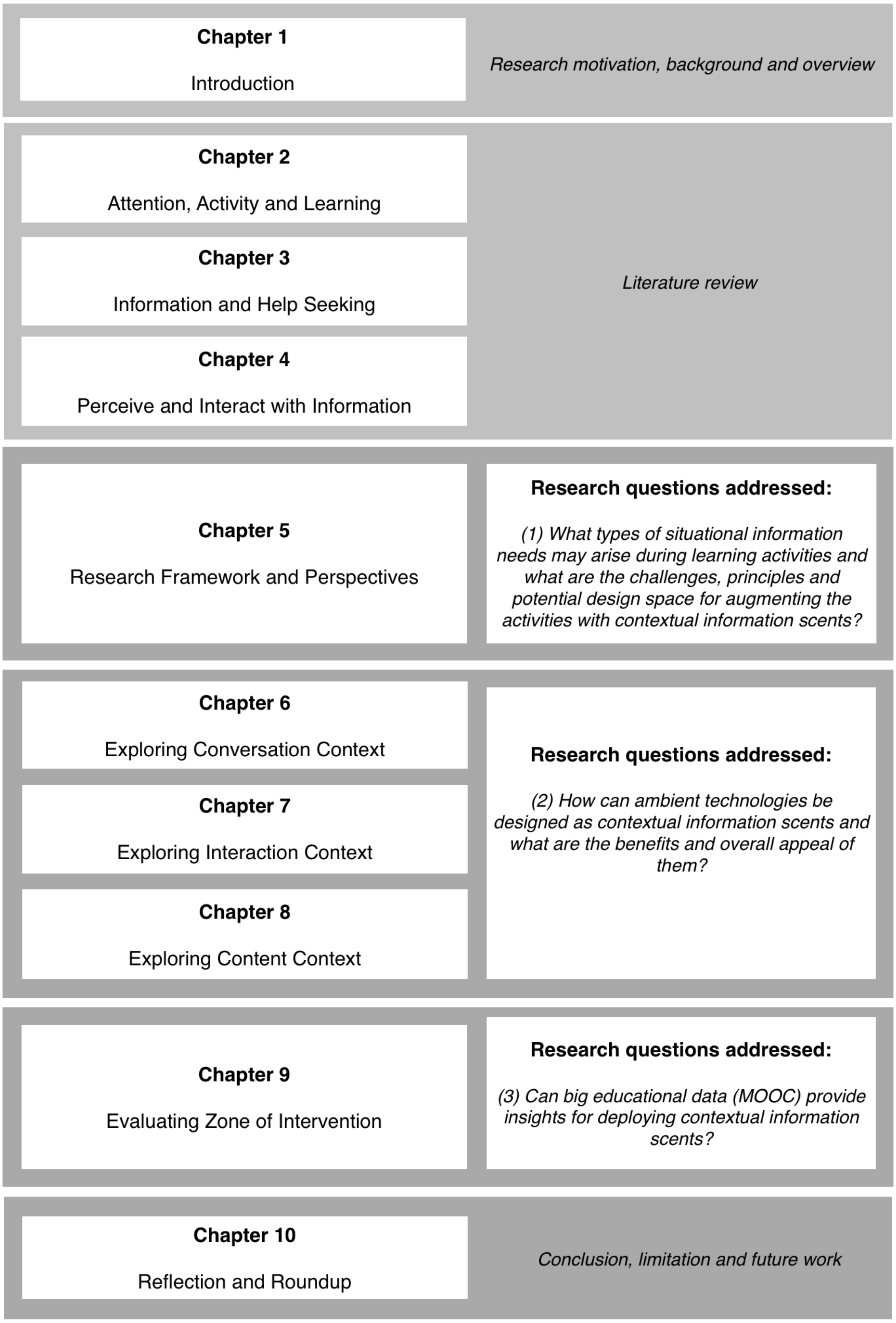


Figure 1.1: Overview of the Thesis Structure

ensure timeliness. In **Chapter 6** we introduce two prototypes of contextual information scents generated with conversation context. *RaindropSearch* investigates capturing and displaying conversational words for building search queries, while *TileSearch* triggers Web searches based on group discussions and retrieves image and Wikipedia results for serendipitous interactions. In **Chapter 7**, we describe *MeetHub Search*, which includes three kinds of contextual information scent based on text interaction in the groupware. **Chapter 8** presents a study of the *BOOC Player*, which displays textbook content corresponding to the MOOC video being viewed. The presented studies answer the second research question raised in the thesis and demonstrate the benefit and overall appeal of contextual information scent.

Chapter 9 answers the third research question in the thesis by employing data science methods to derive design insights for contextual information scent in MOOC learning.

Chapter 10 reflects on the lessons learned from the studies presented in the thesis, and states the contributions, limitations and future research directions of contextual information scent.

2 Attention, Activity and Learning

As a multidisciplinary discipline, human-computer interaction (HCI) lies at the intersection between social and behavioral science on one hand, engineering and computer science on the other. During the course of its development, a number of theories and frameworks have merged, providing us with guidance on designing and evaluating interactive interfaces and techniques. This thesis focuses specifically on *calm technologies* applied in learning activities, which implies the requirement of understanding how people manage attention and perform activities especially in the learning domain. This chapter therefore starts by reviewing a selection of attention theories, which are then followed by related work on high-level abstractions of activities. Finally, cognitive taxonomies of learning and the notion of collaborative learning are reviewed.

2.1 Attention and Cognition

This section is concerned with some of the classical theories that have attempted to explain attention. Before launching into a detailed overview, it would be helpful to establish grounds on which the follow-up discussions will be based upon. This ground is the definition of attention, the earliest of which dates back to the 1890 by William James (James, 1890):

"It is the taking possession in the mind, in clear and vivid form, of one out of several simultaneous possible objects or trains of thought. Focalization, concentration of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others".

Underwood (1993) described James' definition as an "elegant summary". This definition exposes the relationship between attention and consciousness, which many researchers thereafter have made their attempts to conceptualize. During the course of development of attention theories, distinct subcultures have emerged. Some explain man's attention limitation by assuming the existence of structural bottlenecks (bottleneck theories), whereas others believe there is a limit on man's capacity to perform mental work (capacity theories). There are also different aspects of attention research issues to be addressed. For example, the issue of attention in James' definition is what would nowadays be called *selective attention*. In fact, Wickens and McCarley (2007) identify 5 types of attention : focused, selective, switched, divided, and sustained. In this chapter, we briefly summarize selective and divided attention, which we think are the most relevant to our work.

2.1.1 Selective Attention

One of the first researches of selective attention was initiated by Colin Cherry (Cherry, 1953), where he presented the *cocktail party effect*: Suppose you are conversing with someone in a noisy cocktail party, it is fairly easy to hear and follow what your partner says. On the other hand, what other people are talking is barely noticed unless your name is mentioned by someone chatting around. This phenomena indicates that humans selectively attend to certain pieces of information while unattending the rest. The key question is how much of the unattended message could be detected. Cherry found that his participants were unable to recall any specific words in the unattended message. Similar results were produced in a dichotic listening experiement (Moray, 1959), which was widely used in early studies of auditory attention, e.g. in Broadbent (1958) and Treisman (1964). The above findings imply a filtering process where the attended message is let in and the unattended message is filtered out, and this is where the concept of *selective attention* originated.

Selective attention assumes the existence of structural bottlenecks, so that one cannot consciously attend to all of our sensory input at the same time. The related theories are therefore called bottleneck theories (Welford, 1952). Researchers have come up with several bottleneck theories, all of which posit that information must be filtered before entering into the short memory store, so the main question is where the filtering process occurs. Developed by several psychologists, multiple classic attention theories were proposed in history to answer this question.

Bottleneck Theories of Attention

Donald Broadbent was one of the first researchers to characterize the attention selection process. He proposed an early selection model (Broadbent, 1958), which claims that the selection occurs very early, and only one channel of the messages is semantically analyzed and attended to. The selecton process may look like Figure 2.1(a). Broadbent's theory is all-or-nothing, meaning that the unattended messages are completely blocked by the filter. The early selection model was criticized largely due to its inability to explain why the mentioning of names can be attended in the cocktail party scenario. In addition, the filter component in Broadbent's model operates on the basic physical characteristics of the messages (e.g. genders of the speaker, types of sourdns, etc). Gray and Wedderburn (1960) revealed that the meaning of the unattended message can be processed as well in a later study.

In contrast to early selection, Deutsch and Deutsch (1963) proposed another model which

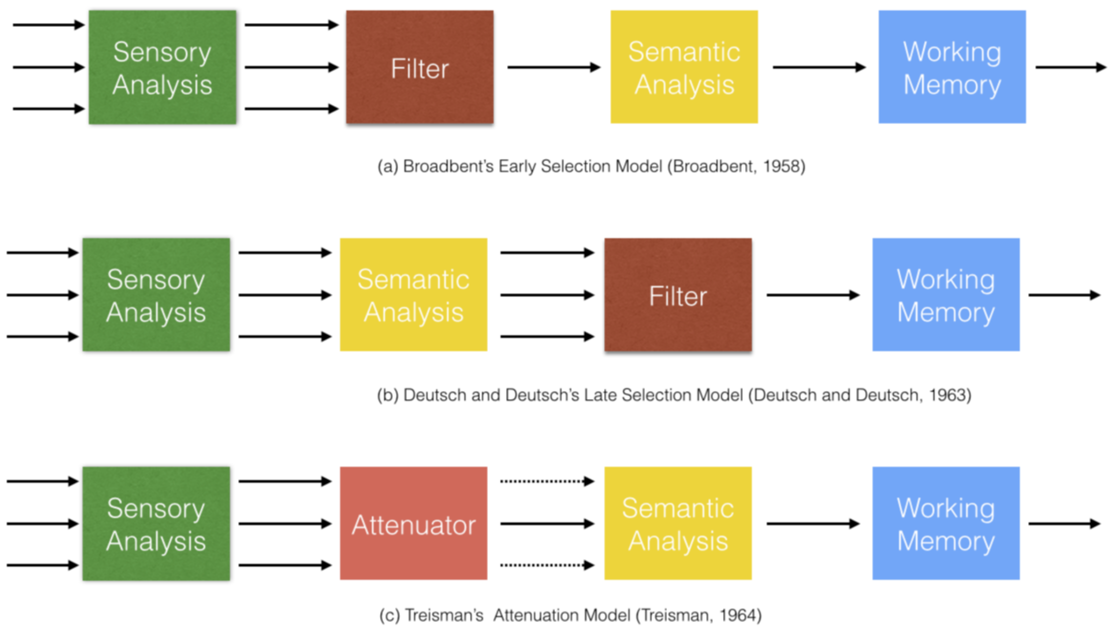


Figure 2.1: Early selection, late selection and attenuation models

argues that the selection and filtering occur much later. As illustrated in Figure 2.1(b), all the messages are processed on the basis of meaning and then filtered based on whether or not the information is pertinent or not. Pertinence refers to the perceived significance level of information, e.g. the utterance of one's names or the sight of those that one cares about or is related to the task. The late selection model is clearly capable of fully explaining the cocktail party phenomena. The main objections to this view is that it is not parsimonious or cognitively economical, i.e. "It is intuitively very unlikely evolution would equip us to process all stimuli to the highest level possible when virtually all of those stimuli are quite irrelevant to our survival, particularly when having the large brain needed to carry our that processing may impose certain physiological costs" (Groome, 2013).

Perhaps the most satisfactory model of attention is the attenuation model proposed by Treisman (1964). Similar to the argument in the early selection model, Treisman also postulated that the selection starts at the physical level and is then followed by a filter, but the unattended information is not completely blocked, it is just attenuated. As a result, highly pertinent information such as one's name will get through the filter and the meaning of it gets processed. The model is depicted in 2.1(c). Treisman further elaborated the model by introducing the concept of threshold to explain how certain information is more likely to be attended. She claimed that every chunk of information had its own threshold depending on, for example, the context or subjective importance, that determines the likelihood to be perceived after attenuation. Like in the late selection model, all the exposed messages in Treisman's attenuation model undergo full processing, but irrelevant stimuli often have high threshold to be fully analyzed, resulting in only physical characteristics rather than semantics of the unattended messages being remembered.

In summary, the recap of the classic attention theories in this section sheds light on how selective attention occurs. Most importantly, we highlight the finding that pertinent information in unattended channels is still likely to be re-attended. This is one of the theoretical basis that dictate why the research in this thesis makes sense.

Visual Attention and Visual Search

While early studies of attention largely involved auditory perceptions, selective attention also holds for visual perception. In the 1970s, Neisser and Becklen (1975) devised a visual anologue of the dichotic listening task, where Neisser and his colleagues found that participants were unaware of the events happening outside the focus of their attention, even when looking right at them. In fact, the participants perceived something was happening, but could not remember the details. A revised version of this study was conducted by Neisser in 1979 with the well-known *Invisible Gorilla Test* (Neisser, 1979), where he asked the subjects to watch a short video of two groups of players, wearing white and black shirts respectively. Subjects were asked to capture whenever the players in white successfully passed a ball, but to ignore the players in black. Partway through the same video, a woman wearing a gorilla suit or carrying an umbrella (depending on the versions of the video) strolls in the scene. Participants were found to intently focus on counting the passes, and 50% of the subjects did not report seeing the gorilla (or the woman carrying an umbrella). This phenomenon is now known as *inattentional blindness* (Mack and Rock, 1998), which refers to the failure to see unexpected stimuli when the attention is focused on something else. *Inattentional blindness* is much affected by the intensity of focused attention. Researchers have shown that the greater the demands on focused attention, the less likely people are to notice objects outside their attention (Macdonald and Lavie, 2011; Simons and Chabris, 1999; Simons and Jensen, 2009).

An important aspect of visual attention is visual search, which requires detection of particular target against a background of other items. Visual search was first investigated in (Neisser et al., 1963; Neisser, 1963, 1964), where Neisser and his colleagues studied the reaction time for locating specific letters presenting in groups of various characters at various locations. The most profound finding was that the recognition of targets against a dissimilar background has a considerable advantage in visual search, and this is called *pop-out effect*. An intuitive real-world experiment he used to demonstrate the *pop-out effect* in his later article (Neisser, 1967) is that his subjects could rapidly recognize the face of the then president John Kennedy among a background of other faces, which were reported by the subjects as simply "blur". The *pop-out effect* indicates no extra effort needs to be devoted to mentally process each of the dissimilar faces. A later study (Treisman and Gelade, 1980) further revealed that for *pop-out effect*, an increased size of distracters does not increase the scanning time accordingly. However this claim holds only if the visual feature of the objects is defined as simple as in the aforementioned example. Conversely, when the objects are complex (e.g. both shapes and colors differ), reaction time is increased. This phenomena can be explained systematically by *feature integration theory* (Treisman and Gelade, 1980; Treisman, 1988), which is regarded

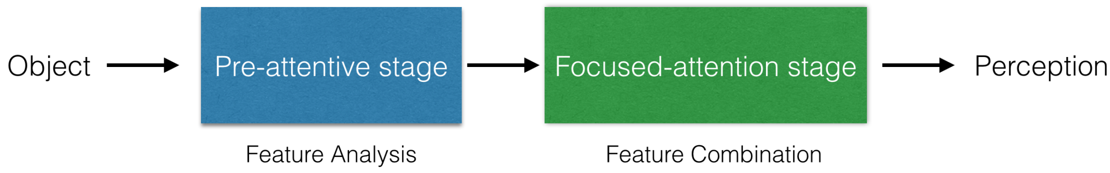


Figure 2.2: Two-stage model of feature integration theory (Treisman, 1988)

as the most influential model of human visual attention. This theory postulates a two-stage process in visual search as illustrated in Figure 2.2 . The first stage is a pre-attentive stage, where visual objects are analyzed in parallel according to their individual features such as shape, color, orientation etc. For the situations where objects have only a single feature, targets will rapidly pop out in this stage. The second stage is a focused attention stage, where the individual features recognized in the first stage are combined to be processed coherently. Objects that are defined by two or more features have to be processed serially for identification, giving rise to increased reaction time.

Though our targeted scenarios do not explicitly imply visual search tasks, the presentation of information scents during learning activities may involve conscious or subconscious scan of pertinent information,requiring the system to reduce the scanning time as minimal as possible.

2.1.2 Divided Attention

Rather than studying how much one can be aware of the unattended information while focusing on something else, *divided attention* studies one's ability to attend to more than one concurrent tasks. The question concerns what kind of tasks can be processed in parallel and how well we can perform them at the same time.

Factors Determining Multi-task Performance

Before converging on a well-founded theory, early research explored the factors related to the performance issue of multi-tasking. With a classic dichotic listening experiment, Allport et al. (1972) found a chance level performance in a subsequent recognition test of the words presented in the unattended channel. However, when the to-be-remembered words were presented visually with pictorial representations, the recognition performance was excellent. The finding suggested that the tasks interfere to the extent if their sensory inputs are from the same stimulus modality (e.g. visual or auditory). Put in other words, employing different modalities has the advantage of improving task performance. Similar findings were confirmed by many other researchers (Mcleod, 1977; Treisman and Davies, 1973). This sensory modality factor is usually referred to as **task similarity**.

It is natural to believe **task difficulty** also influences multi-task performance. In the 1970s,

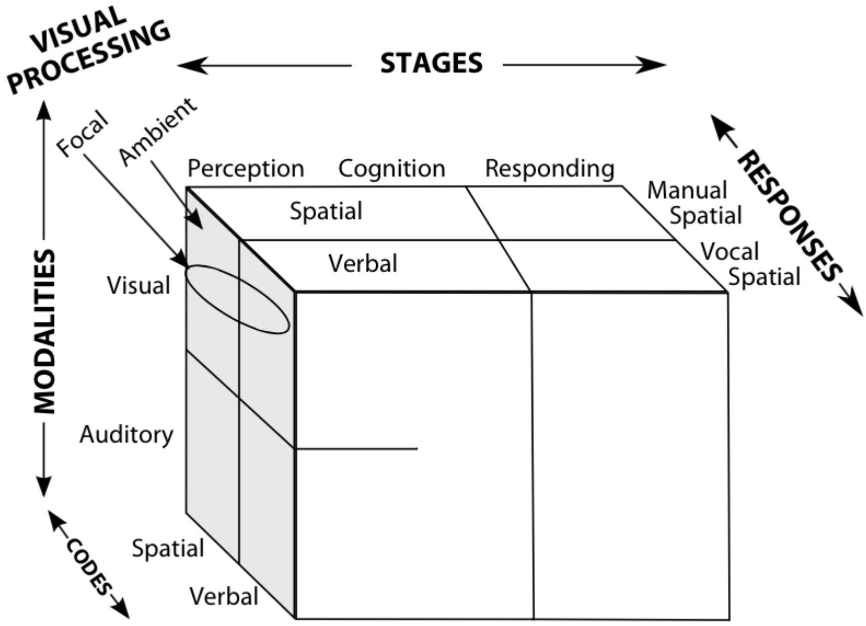


Figure 2.3: The 4-D multiple resource model (redrawn from (Wickens, 2008))

Sullivan and his colleagues designed a dual task Sullivan (1976) , i.e. shadowing an auditory message and detecting words on a non-shadowed message at the same time. When the difficulty of the task was increased by using a less redundant message, fewer words were successfully detected on the non-shadowed message. This indicates that higher task difficulty may deteriorate task performance. However, task difficulty is a very subjective notion and is itself difficult to measure. One of the most influential frameworks is the one proposed by (Navon and Gopher, 1979), where the authors summarized difficult tasks as either data-limited or resource-limited. The former refers to the limitation in the available information. For example, writing a literature review relies heavily on information about related work that is "external" to the writing task. The latter are the tasks that demand cognitive resources, for example, attending a lecture in classroom is cognitively demanding. Clearly, resource-limited tasks are prone to breakdowns in performance.

Several early studies demonstrated **task practice** also affected our ability to multitask. Allport et al. (1972) showed an expert pianist can sight-read at the piano whilst shadowing. Another study (Shaffer, 1975) exhibited proficient typist can type whilst shadowing. The most direct evidence that shows practice improves multitask performance is from Spelke et al. (1976). Two volunteers were trained extensively on an unfamiliar task which is a combination of reading and dictation at the same time. Initially the subjects suffered a lot. After six weeks of daily practice, their writing and reading speed were greatly improved. Four months afterwards, they could even perform another activity, i.e. categorizing dictated words.

Capacity Theories of Attention

Human beings have the ability to divide attention for multitasking, with deteriorated performance on one of the tasks, if not all. The studies presented before are concerned with factors that determine one's ability of divided attention. Such ability is clearly beyond the scope of bottleneck theories which assume one task at a time. Therefore, a more flexible theory, the *single resource theory* is brought forward by Kahneman (1973). Though the new theory also claims attentional resource is limited as the bottleneck theories do, capacity can be allocated to a wide range of activities at the same time. Single resource theory posits that a single pool of cognitive resources (aka. capacity) is shared amongst competing tasks. The limits of the capacity vary depending on the environment, task difficulty, and individual differences (e.g. level of arousal, expert or novice). If the cognitive demands of the combined tasks do not exceed the resources in the central pool, then the tasks will not interfere with each other, otherwise divided attention is detrimental to the performance of one or both tasks.

A notable limitation of the single resource theory is its inability to explain one of the empirical findings presented before: for concurrent tasks in the same modality, allocation of attentional resources is much more difficult. The multiple resource theory (MRT) (Wickens, 1980, 1984) is then developed to account for this limitation. The original MRT model consists of 3 dimensions. The *stages of processing* includes perception, cognition and responding stages. The former two stages may interfere if parallel tasks competing for attention resources in the same pool, but the responding stage, which is concerned with selection and execution of actions uses a different resource pool. *Modalities* dimension comprises visual and auditory senses. As previously stated, one can attend to tasks of different modalities with little interference, but tasks with the same modality would experience great performance decrement. The *codes of processing* dimension consists of spatial and verbal processes, which corresponds to manual control and speech actions. The fourth dimension, *visual processing channels* was later added to the model (Leibowitz and Post, 1982; Previc, 1998), discriminating between focal and ambient vision. The former refers to central vision with high fixations, whereas the latter addresses the peripheral vision. Even when concentrating on something (e.g. reading a book), we are still able to perceive fast changes or movements happening around us (e.g. someone is approaching). The complete 4-D MRT model is depicted visually with a cube in Figure 2.3. This model nicely demonstrates that it is easier to perform simultaneous tasks that require resources from different dimensions. For tasks that demand resources of different levels along the same dimension, a time-sharing scheme is implicitly employed to make task executions more efficient.

Automatic Processes

While capacity theory uses multiple dimensions to model attention resources of different kinds, it does not explain the performance improvement caused by practice, which is an important factor that affects simultaneous task performance. This involves the notion of *automatic processes*, which do not require attention, thus allowing concurrently performing attention-demanding tasks. This theory was proposed by (Shiffrin and Schneider, 1977), where the authors argued that the automatic processes are not capacity-limited so that they are not affected by the limitations of short-term memory. However, fully automatic processes without requiring any attention are rare. One example is in the well-known Stroop task (Stroop, 1935), automatic processing of color words are unavoidably influenced by the semantics of the presented words.

Rather than asserting automatic processes are completely inattentive, Norman and Shallice (1986) proposed three levels of processing. *Fully automatic processing* is totally awareness-free and controlled by schemata; *Deliberate control* involves conscious awareness of the processes; Last, *partially automatic processing* sits somewhere in between the former two levels, involving both automaticity and somewhat control. Norman and Shallice claimed there exists a contention scheduling mechanism that resolves conflicts so that the processes do not interfere with each other. The theory provides a natural explanation for the key phenomena that some processes are not completely automatic. In fact, processes may become automatic through practices because practice leads to the storage of increased information about the stimulus, so "automaticity is memory retrieval : performance is automatic when it is based on a single-step direct-access retrieval of past solutions from memory" (Logan, 1988). This explains why automatic processes affect little on the cognitive capacity available to other tasks.

2.1.3 Work Interruption and Multitasking Continuum

Divided attention mainly concerns how attention is devoted to the execution of parallel tasks. Practically, multiple tasks are not always parallel but also interleaved. Put it differently, multitasking behaviors can be characterized by the time spent on one task before switching to another. If the time is very short (e.g. in seconds), then the tasks are characterized as *concurrent multitasking*, which coincides our discussions in the divided attention section. Researchers attempt to understand how human can divide attention to multiple tasks simultaneously. In contrast, *sequential multitasking* often requires longer time (e.g. in minutes to hours) to be spent on one task before switching to another. Broadly speaking, research in the area of sequential multitasking largely overlaps with the fields of *task switching* (Monsell, 2003) and *work interruption* (Brixey et al., 2007), which concerns shifts of attentions or even changes of goals in the working memory (Altmann and Trafton, 2002; Frese and Zapf, 1994; Salvucci et al., 2009a).

Work Interruption

Typical interruptions may be initiated externally or internally (Miyata and Norman, 1986; Jett and George, 2003). External interruptions are usually from the surroundings, which are probably neither anticipated nor controlled. Such interruptions include pure distractions or the interruptions that lead to secondary tasks. The former can be seen as noise. An example of distraction is when certain objects suddenly fall down to the ground, diverting one's attention away from on-going work. Distractions as such usually last shortly. However, distractions can also be long-lasting, for example, telephone conversations from one's neighbor in the office may constantly distract attention. These long distractions are termed by Hacker (2003) as regulation difficulties. Distractions, regardless of durations do not direct one to a secondary task with new goals, they simply inhibit the primary task from being smoothly executed.

Unlike distractions, some external interruptions may temporarily pull individuals out of their action and divert them to a new goal (Frese and Zapf, 1994). For example, while a student is watching MOOC videos, an in-video quiz may pop out, so that the student has to suspend the video watching and complete the quiz first. Numerous studies have shown that external interruptions as such can result in prolonged time in both primary and the interruption task (Bailey and Konstan, 2006; Eyrolle and Cellier, 2000; Trafton et al., 2003), duplicated work (Wickens and McCarley, 2007), as well as increased anxiety, frustration, and stress (Carton and Aiello, 2009; Mark et al., 2008; Zijlstra et al., 1999). Interruptions are not always negative to the primary task, Speier et al. (2003) found out that interrupting in simple tasks increased the performance while interrupting in complex task experienced the opposite. Speier argued that the subjects perceive the simple tasks as too easy so they did not devote full attention to it. In this case, unpredicted interruptions forced the subjects to focus more on the task which consequently lead to better performance.

In some occasions, interruptions are also self-initiated. Individuals who invoke the interruptions have full control of them. Thus they are conscious about the task switching. As an example, a student engaged in watching a MOOC video may deliberately pauses the video to search relevant study materials online. Internal interruptions can also be breaks, which are defined as resources that help individuals maintain optimal mental and physical performance (Hobfoll, 1989; Sonnentag and Zijlstra, 2006). For example, a MOOC student may want to make a reflection of the knowledge he has learned so far after finish watching 2 videos, and will continue with the rest after the break. A more comprehensive classification of self-interruption can be found in (Jin and Dabbish, 2009), where the authors have identified 7 types of self-interruptions as follows: (1) Adjustment (improving some aspects of the environment, which is intended to increase the productivity of the primary task) (2) Break (temporarily switching to another task to alleviate stress and fatigue with the primary task) (3) Inquiry (seeking external information to facilitate the primary task) (4) Recollection (a prospective memory event which recalls another task that must be performed immediately in case of forgetting) (5) Routine (performing a task as a routine based on prior experiences) (6) Trigger (performing a new task which is stimulated from the current task) (7) Wait (perform another task maximize productivity because some bottlenecks are encountered so that the continuation of the primary task is suspended). These 7-style categorization clearly demonstrates that self-initiated task switching can be employed as work strategies (Konig et al., 2005).

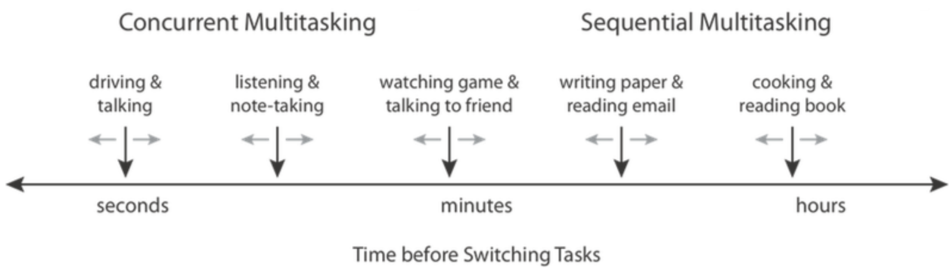


Figure 2.4: The multitasking continuum (redrawn from (Salvucci et al., 2009b))

Multitasking Continuum

Concurrent and sequential multitasking were originally separated areas of research until Salvucci et al. (2009b) proposed a unified theory of *multitasking continuum*. As depicted in Figure 2.4, the left side of the continuum is characterized as concurrent multitasking, exemplified as concurrent tasks such as driving and talking, listening and note-taking. On the other hand, sequential multitasking activities such as cooking and reading book at the same time can be found on the right side of the continuum.

The multitasking continuum encompasses several cognitive theories to explain multitasking. First, the *ACT-R cognitive architecture* (Anderson et al., 2004; Mellon et al., 2007) defines a set of cognitive modules that work in parallel, including (1) a declarative memory module (keeping memories for factual knowledge and task instructions); (2) a goal module (tracking the goal of activity); (3) a problem representation module (holding problem representations such as intermediate steps) and (4) a procedural module (connecting all other modules and controlling the flow of information). According to ACT-R, only one task is permitted per module at a time. However, multitasking continuum theory also combines *threaded cognition theory* (Salvucci and Taatgen, 2008) which allows multiple tasks to work simultaneously across the ACT-R modules with a greedy threading policy. Finally, the multitasking continuum incorporates *memory-for-goals theory* (Altmann and Trafton, 2002) as well. This theory posits that when people need to initiate a new goal, they must strengthen this goal in memory to increase its activation level above the current goal, so as to set the new goal as the primary one. In the meanwhile, the suppressed old goal decays until the interruption goal is completed, when the decayed old goal can be recalled and resumed. In fact, if interruption task requires to process the problem representation module, the old problem representation can not be maintained at the same time, it must be swapped out and stored in the declarative memory, from where it is retrieved later when the interruption task is completed. However, the representation may require time to be retrieved, resulting in resumption lag (Altmann and Trafton, 2004). Moreover, memory retrievals are not always successful. Performance in the primary task recall can be improved by rehearsing the problem representation before storing it, but this usually issue in interruption lag (Altmann and Trafton, 2004). The previous described interruption and resumption stages are analogous to the stack actions in computer science, i.e. push and pop. The whole procedure is illustrated in Figure 2.5.

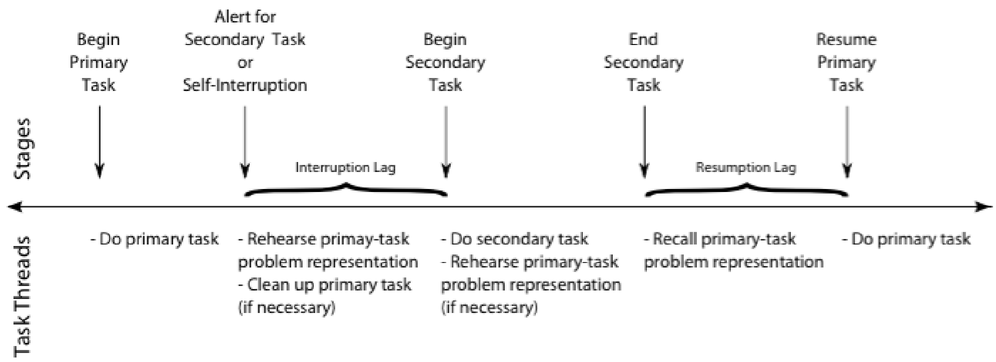


Figure 2.5: The stages of interruption and resumption (redrawn from (Bogunovich and Salvucci, 2011))

2.1.4 Priming Effect

Priming is a concept that stresses the potential interferences between consecutive actions. It is asserted that exposing one stimulus (called *prime*) affects a person's responses to a subsequent stimulus (called *target*). To be precise, priming may increase or decrease the recognition speed of the subsequent item. In the former case, the priming effect is facilitatory, whereas the latter is inhibitory (Tipper and Cranston, 1985). Priming effect was first demonstrated by Meyer and Schvaneveldt (1971). Using a lexical decision task, they found that a word was recognized faster if it is preceded by another related word. For example, it is quicker for a person who sees "car" to recognize the word "train" than the word "apple", because the former two are semantically associated as "transportation means". This type of priming effects is called *semantic priming*(McNamara, 2005). In this section we review *semantic priming* together with an important theory, the *spreading activation model*, which explains how priming effects occur.

Semantic Priming

Semantic Priming is arguably the most common type of priming in word recognition tasks. When we look back at the previously reviewed attenuation model, which claims that unattended stimuli are attenuated and can probably still be identified if they are semantically related, we can now identify it as an example of semantic priming. The semantically related words are actually primed so that it reaches the threshold for being recognized.

Apart from the aforementioned original seminal experiments by Meyer and Schvaneveldt (1971), the priming effect of semantically related words has thus far been investigated in hundreds of studies, most of which were summarized in two meta analyses (Lucas, 2000; Van den Bussche et al., 2009). Researchers have identified two different types of semantic priming based on whether or not the prime and the target words are normatively associative. For example, a prime-target pair of NURSE-DOCTOR will have *associative semantic priming* effects because the concept of nurse and doctor are associated. In contrast, a NURSE-WIFE pair will induce an *non-associative semantic priming* effect because the similarity of the two concepts can be characterized as a sharing of a number of semantic features, though "nurse" does not directly elicit "wife" as an associate (Fischler, 1977).

These two types of semantic priming create different priming effects. In her meta-analysis, Lucas (2000) found that associative priming had boosted effect compared to non-associative priming. Van den Bussche et al. (2009) extend the meta-analysis to include not only word primes, but also other types of stimuli such as pictures, Arabic numbers. They found that the semantic priming effect is larger for symbols than for words.

Spreading Activation Model

Spreading Activation Model, proposed by Collins and Loftus (1975) is an important theory that explains how semantic priming effect occurs. The model assumes that words and their meanings are stored in separate networks in our mind. One network is lexical, storing phonemic and orthographic information about the words whereas the other is semantic, containing concepts of the words as well as their links to the lexical network. Nodes in each network are interconnected based on their lexical and semantic distances (i.e. similarities). Between-network links are equally easily to be activated compared to the their within-network counterpart. The key idea behind semantic priming is that visually presenting words activates the links to semantically associated concepts more quickly.

2.2 Activity and Learning

The preceding section presents a set of cognitive-science based theories, which has provided us with significant background knowledge in understanding how humans manage their attentions. However, for HCI researchers, these findings that are derived or synthesized from simple psychological experiments, often lack considerations of the context, such as why the subjects are performing the tasks and what they can obtain from completing the tasks. As Norman (1980) pointed out:

"The problem seemed to be in the lack of consideration of other aspects of human behavior, of interaction with other people and with the environment, of the influence of the history of the person, or even the culture, and of the lack of consideration of the special problems and issues confronting an animate organism that must survive as both an individual and as a species"

Norman's statement was not specifically in response to applying attention theories in HCI, but to a broader range of cognitive theories that are known as the *information processing psychology* or *first-wave HCI theories* (Kaptelinin et al., 2003). The key message of this problem articulation for HCI research and practice is, rather than focusing merely on cognitive performance on the tasks, the major concern is to understand and design technologies in the context of meaningful activities. Specifically to this dissertation, the context is about *learning activities*. In this section, we briefly review elaborated concepts of both activities and learning in literature. The former is offered by activity theory which will be presented in the first place, followed by the taxonomies of learning as well as its collaborative aspects.

2.2.1 Activity Theory

As the limitations of adopting cognitive theories in HCI was widely acknowledged in the early 1990s (Carroll, 1991), a number of theories were proposed to extend the scope of human capabilities to understanding and supporting meaningful actions and social interactions in everyday contexts (Kaptelinin et al., 2003). These theories include activity theory (Bdker, 1991; Kuutti, 1991; Nardi, 1996), distributed cognition (Hutchins, 1995; Hollan et al., 2000), phenomenology (Flores and Winograd, 1986), situated action (Suchman and Reconfigurations, 1986) etc. We select *Activity Theory (AT)* among others for discussion because it emphasizes and centers on activity, which is a fundamental concept that constitutes our everyday contexts. In addition, the AT also provides many useful tools, e.g. the activity triangle and checklist for understanding and diagnosing interactive systems, and these tools are used extensively in the discussions hereafter in this dissertation.

Activity System Model

Modern Activity Theory is known to originate from cultural-historical psychology developed by Lev Vygotsky and his colleagues in the 1920s and 1930s. Vygotsky (1978) models human activities as subject-object interaction: Any activity is directed by a *subject* towards an *object*, and the interaction is mediated by artifacts, also known as *tools* or instruments. The object refers to either physical objects being interacted or the objective of the activity. Similarly,the tools can be either physical tools (e.g. a hammer) or psychological tools (language and signs). Take MOOC learning as an example, a student (subject) studies a course (object) through interacting with the MOOC learning materials (tools). Vygotsky's idea can be illustrated as a triangle model as shown in Figure 2.6(a).

Vygotsky's model is mainly concerned with individual activities. Alexei Leontiev, one of Vygotsky's students, however mentioned in his work (Leont'ev, 1978) that activities can also be carried out by collective subjects. In other words, the subject-object interaction can mediated by the social entities, known as *community* (Figure 2.6(b)). The concept of community was developed but not explored systematically in Leontiev's original work. A more comprehensive model, called *activity system model* (Engeström, 1987) was developed by Engeström, who expanded Vogotsky and Leontiev's models with more societally constituted forms of mediation: tools, rules, and division of labor. Put another way, Engeström posits that an activity system describes the interactions between subjects and objects, intertwined with the aforementioned elements (Figure 2.6(c)). Again, let us consider the scenario of a student studying MOOCs. The *object* of the activity is to learn the course, and the *expected outcome* is the gained

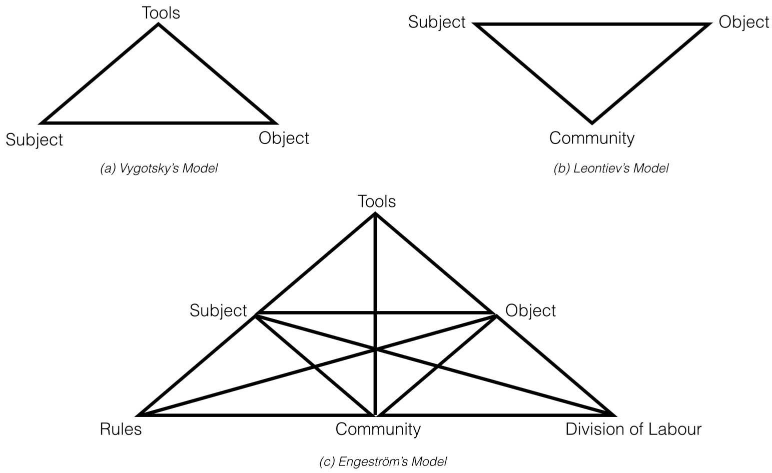


Figure 2.6: The activity system models

knowledge or experience through learning, which may or may not be reflected by the score he or she obtains. The student uses a variety of tools on the MOOC platform to support is learning objective, including Web browser, lecture videos, wikis and discussion forums. The *community* comprises other members of the MOOC learning activity, e.g. the instructors, teaching assistants, his or her fellow students. The community is mediated by many explicit and implicit *rules*, e.g. taking quizzes or exams, peer-grading and forum rules etc. To achieve the learning objective, the student may participate in the discussion forum to get help as well as to help others. The role he plays in the community is actually mediated by the *division of labour*.

Principles of Activity Theory

The main message that the Activity Theory conveys is that our interactions with the world cannot be understood without the context where the interactions take place. The interactions and contexts are indispensable elements of an activity, which is socially cultivated. This message can be elaborated into five basic principles (Kaptelinin, 2006):

1. **Object Orientedness.** Each activity is oriented towards an object. The object can be an entity that objectively exists. It can be a physical object (e.g. a car), a virtual object (e.g. a software application) or even certain properties residing in one's mind (e.g. learning a MOOC). All human activities are driven and directed by the object, and the different activities can be differentiated by their corresponding objects.

2. **Hierarchical Structure of Activity.** This principle is directly derived from (Leont'ev, 1978), which claims that an activity is built from a three-level hierarchy. The top layer includes a *motive* which generates the *activity*. For instance, a student in Computer Science may want to take a course in Computer Graphics, but he lacks background

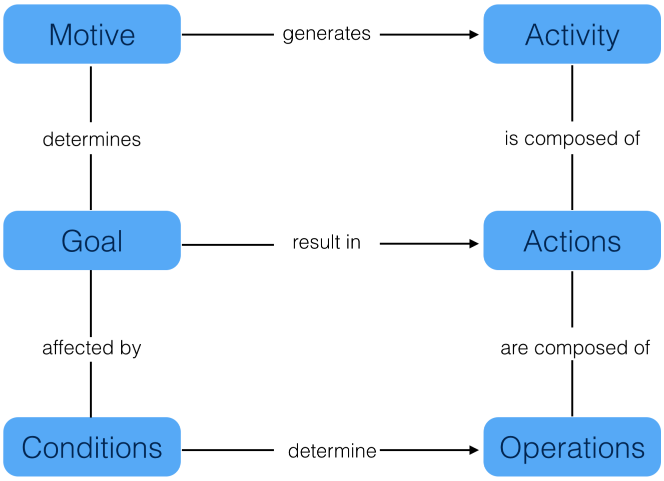


Figure 2.7: Activity, actions and operations (redrawn from (Wilson, 2005))

knowledge in Linear Algebra. He is then motivated to take an Algebra MOOC which generates a MOOC learning activity. In order to fulfil the object, his activity is decomposed into many actions, such as watching video lectures, completing assignments and posting in forums etc. Each *action* is associated with a different *goal* (e.g. the goal of watching videos is to learn the lecture content). Actions are implemented through low-level *operations* oriented towards *conditions*. People are often unaware of the operations. For example, the student watches MOOC videos through combinations of video interactions. At the beginning, the student may need to learn how to make forward jumps in the video. Once familiar, performing forward jumps will immediately become automated processes that do not require conscious awareness.

3. **Internalization and Externalization.** Internalization means external activities can be transformed in such way that people can perform them without interacting with the actual object, i.e. the activities are internalized in one's mind. For example, simple arithmetic calculations can be performed mentally, so the activity becomes internal. On the contrary, internal activities can also be transformed to external ones. Calculations involving complex arithmetic operations may require calculators. In addition, in collective settings, people need to externalize their thoughts for collaboration or cooperation.

4. **Tool Mediation.** Activity Theory stresses that tools represent the accumulated experiences of people who designed and improved them to solve similar problems in the past. Such experiences are reflected on the affordances of the tool as well as the knowledge about how to use the tool. Tools may shape external activities as well as internal ones. For example, MOOCs (tools) foster self-paced individual learning activities whereas classroom teaching cultivates more teacher-student, student-student interactions (external activities). A person's mental calculation processes (internal activities) may depend on whether the person is used to calculating by hand or with a abacus (tools).

5. **Development.** Activity Theory sees activities as continuously developmental processes, so an activity can only be understood when taken the context of its development into account.

Activity Checklist

Activity Theory is not a predictive theory that aims at predicting human behaviors. Instead, it is a high-level analytical framework that helps HCI researchers to systematically frame the contexts of human activities, so that key issues in the design when deployed in real life can be identified. A number of work has been proposed to actualize the concepts and principles developed in the Activity Theory to produce practical usage (Quek and Shah, 2004), including the ArtAD method (Korpela, 1997; Korpela et al., 2001), the Activity Checklist (Kaptelinin et al., 1999), the AODM method (Mwanza, 2002), the Jonassen & Rohrer-Murphy framework (Jonassen and Rohrer-Murphy, 1999) and the Martins & Daltrini framework (Martins and Daltrini, 1999). According to (Quek and Shah, 2004), among these AT-based methods only the ActAD method and the Activity Checklist are concerned with evaluation of interactive systems. Compared to the ActAD method, the Activity Checklist is more strongly coherent with the AT principles presented before, and was widely applied in many evaluation research such as in (Fjeld et al., 2004; Uden and Willis, 2001). In this dissertation we mainly use the Activity Theory as a diagnostic tool for evaluating the various designs of our prototypes, and the Activity Checklist is exclusively applied to support the articulation of complex real-life problems. The Checklist corresponds to four main perspectives:

1. **Means and ends.** This perspective corresponds to the principle of "hierarchical structure of activity". It concerns with the identification of goals and subgoals, and extends the scope to higher level activities or down to lower level operations.

2. **Social and physical aspects of the environment.** This perspective corresponds to the principle of "object-orientedness". It identifies what factors are involved in the activities and constitutes the environment where the technology is deployed, including the tools, division of labor as well as rules and norms that regulate the social interactions.

3. **Learning, cognition and articulation.** This perspective is concerned with how computer systems support externalization or internalization of human actions to facilitate cognition, coordination and problem articulation.

4. **Development.** This perspective involves analyses of potential historical changes in the environment that influences the development of the activities.

Kaptelinin et al. (1999) provides a set of sample questions in each perspective for evaluating interactive systems. However, as the authors claim, the Checklist does not have to be used in a linear way to examine all of the four perspectives one by one. Researchers can focus on relevant items and ignore irrelevant ones.

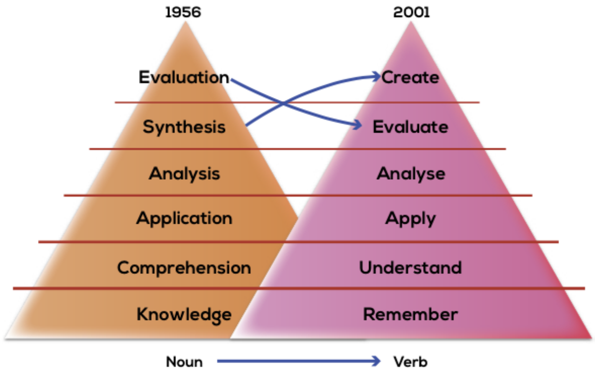


Figure 2.8: Bloom's cognitive taxonomy of learning (adapted from (Bloom et al., 1956) and (Anderson et al., 2001))

2.2.2 Learning Taxonomies and Collaborative Learning

The previous section is concerned with high-level representations of general activities with Activity Theory. In this section we move forward to a very specific type of activity which frames the context of this dissertation, i.e. learning. Research in the field of learning and instruction is a relatively matured field, with a number of theories being developed along its history. We will focus on the cognitive dimensions of knowledge and learning as well as collaborative learning.

Dimensions of Learning and Knowledge

What are the objectives of learning and what are the required skills to achieve these objectives? To answer these questions, in the 1950s a group of educational researchers led by Benjamin Bloom started classifying learning objectives into three domains: (1) Cognitive (mental skills or knowledge) (2) Affective (feelings or attitude) (3) Psychomotor (physical skills) (Bloom et al., 1956). The review in this section mainly is concerned with the cognitive dimensions, which involves knowledge and the development of intellectual skills.

Over the past fifty years, a number of alternative taxonomies (Gagné, 1985; Merrill, 1983; Reigeluth and Moore, 1999) have been proposed to supplement, improve or even replace the original model. Perhaps the most widely accepted revision was the one proposed by Anderson et al. (2001), the major changes of which include changing the names of the levels from nouns to verbs, and reversing the order of the highest two levels, as illustrated in Figure 2.8. The taxonomy includes six levels of learning objectives with increased cognitive complexity that ranges from remember to create:

1. **Remember (Knowledge).** Student reliably recalls or recognizes concepts, principles that were learned previously. For example, the student knows the principles of Activity Theory.

2. **Understand (Comprehension).** Student interprets, exemplifies, summarizes or explains a previously learned concept. For example, the student explains the gist of the Activity Theory.

3. **Apply (Application).** Student transfers the knowledge learned to complete a concrete task. For example, the student uses the principle of *Hierarchical structure of Activity* in the Activity Theory to identify actions and operations in a new activity system.

4. **Analyse (Analysis).** Student distinguishes and organizes the knowledge into structural components that may be better understood. For example, the student compares the Activity Theory with the Distributed Cognition Theory.

5. **Evaluate (Evaluation).** Student assesses, judges or critiques learning materials with specific criteria. For example, the student critiques the weakness of the Activity Theory .

6. **Create (Synthesis).** Student integrates knowledge to produce or construct new ideas. For example, the student comes up with a new theory that improves the Activity Theory.

In addition to the cognitive dimension of learning, Bloom's taxonomy includes a conceptualization of knowledge dimension, which is also revised in (Anderson et al., 2001) by adding a new metacognitive knowledge to the original three-level models. Unlike the dimension of learning which represents increased cognitive level of complexity, the knowledge dimension consists of four levels of knowledge ranging from concrete (factual) to abstract (metacognitive):

1. **Factual.** Must-know knowledge that is fundamental to specific disciplines, such as facts and terminologies etc.

2. **Conceptual.** Knowledge that is constructed by connecting or generalizing the fundamental factual knowledge. Examples are classifications, principles and theories etc.

3. **Procedural.** Methodological knowledge that describes how to do something in specific disciplines, including algorithms, usage critera or specific skills.

4. **Metacognitive.** Awareness and knowledge of one's own cognition. Reflective knowledge when evaluating one's own learning progress or monitoring comprehension is an example of metacognitive knowledge.

The knowledge and cognitive process dimensions of learning do not exist in isolation. The intersection of the two dimensions formulate statements of learning objectives, which contains a **verb** that describes the action associated with the cognitive process and an **object** that depicts the the knowledge students are expected to construct (Anderson et al., 2001). Example objective statements are shown in Figure 2.9. One thing to note is that the statements in the cells are objectives rather than activities. A more appropriate presentation of the statements, for example at the intersection between the "create" thinking skill and "factual" knowledge, is that "student **is able to** generate a log of daily activities".

The Knowledge Dimension

concrete knowledge → abstract knowledge

The Cognitive Process Dimension (lower order thinking → higher order thinking)

	Factual	Conceptual	Procedural	Metacognitive
Create	**Generate** a log of daily activities	**Assemble** a team of experts.	**Design** an efficient project workflow.	**Create** a learning portfolio.
Evaluate	**Check** for consistency among sources.	**Determine** relevance of results.	**Judge** efficiency of sampling techniques.	**Reflect** on one's progress.
Analyze	**Select** the most complete list of activities.	**Differentiate** high and low culture.	**Integrate** compliance with regulations.	**Deconstruct** one's biases.
Apply	**Respond** to frequently asked questions.	**Provide** advice to novices.	**Carry out** pH tests of water samples.	**Use** techniques that match one's strengths.
Understand	**Summarize** features of a new product.	**Classify** adhesives by toxicity.	**Clarify** assembly instructions.	**Predict** one's response to culture shock.
Remember	**List** primary and secondary colors.	**Recognize** symptoms of exhaustion.	**Recall** how to perform CPR.	**Identify** strategies for retaining information.

Figure 2.9: Statements of learning objectives (adapted from (Heer, 2015))

Collaborative Learning

An important paradigm in learning is *Collaborative Learning*, which can be broadly defined as "a situation in which two or more people learn or attempt to learn something together" (Dillenbourg, 1999). The meaning of "learning" in this definition encompasses a variety of activities in research literature. More commonly, however, it specifically refers to joint problem solving activities, where learning comes up as a side-effect of collaboration process, measured by the elicitation of new knowledge or the improvement of task performance (Dillenbourg, 1999).

Collaborative learning is known to have roots in early constructivist theories (Piaget, 1970; Vygotsky, 1978), which essentially claims that humans learn better by constructing knowledge by themselves through interactions with their experiences. The original theory was founded by Piaget, with focus on individual cognitive development, thereby being remembered as cognitive constructivism. Nonetheless, researchers borrowed some concepts in his cognitive development framework, such as conflict (discrepancy between a what a child believes is true and what s/he is experience as true), which is known to trigger learning and intellectual growth to develop theories in collaborative learning (Dillenbourg et al., 1995). The key message is that such conflicts can be facilitated through social interactions in a group, where group mates are expected to possess different knowledge or hold opinions from different perspectives. At this very point, the idea largely coincides with Vygotsky's social contructivist theory, which emphasizes the social context of learning and that knowledge is mutually built and constructed among people. According to Vygotsky (1978), there exists a *zone of proximal development*

(ZPD), which relates to the difference between what a child can achieve independently and what s/he can with guidance from a skilled partner, known as the *more knowledgeable other* (MKO). Through social interactions, students with ZPD may greatly improve understanding compared to those who working alone, as proved by Freund (1990).

Dillenbourg (1999) sees learning as a side-effect of the collaboration process. When adhere to the Bloom's taxonomy of learning and knowledge, numerous researchers argued that collaborative learning is potentially beneficial for developing high-order critical thinking (Bailin et al., 1999; Heyman, 2008; Thayer-Bacon, 2000), because such high-order thinking skills involve the ability to respond constructively to others during group discussion. Additionally, researchers also claim collaborative learning may improve metacognition, since social interactions encourage the construction and refinement of meta-cognitive knowledge (Schraw and Moshman, 1995) by promoting metacognitive discourse (Hennessey, 1999).

3 Information and Help Seeking

The context of this dissertation assumes that in various learning activities students manage their cognitive resources to seek information or help. A review of human cognitive capabilities as well as theoretical models of activities and learning were presented previously. This chapter is devoted to deliver a comprehensive overview of theoretical models and principles of information seeking behaviors, and further extend it exclusively to the theories of help seeking in learning context.

3.1 Information Seeking and Searching

"Knowledge is of two kinds. We know a subject ourselves, or we know where we can find information upon it. When we enquire into any subject, the first thing we have to do is to know what books have treated of it. This leads us to look at catalogues, and at the backs of books in libraries."

This is a quote from a famous English writer Samuel Johnson 240 years ago, as documented by Boswell (1873). Samuel recognized the means and places of the time to find information, i.e. books in libraries, which are still valid nowadays. With the "explosion" of digital information in the past 20 years, computer scientists have blazed a new trail to look for information, i.e. through information retrieval systems, exemplified by those well-known search engines. The act of looking for information, no matter on the Internet or in the libraries, is generally referred to as *information seeking*, which numerous researchers in information science have for years attempted to formally define.

Marchionini and Komlodi (1998) view information seeking as "a process in which humans purposefully engage in order to change their state of knowledge". Case (2002) refers to it as "a conscious effort to acquire information in response to a need or gap in your knowledge". These definitions may date back to the seminal work from Wilson (1981), who coined the term **information seeking behaviors** on the ground that it results from the recognition of some need, perceived by the user. Twenty years later, Wilson (2000) further refines his original defini-

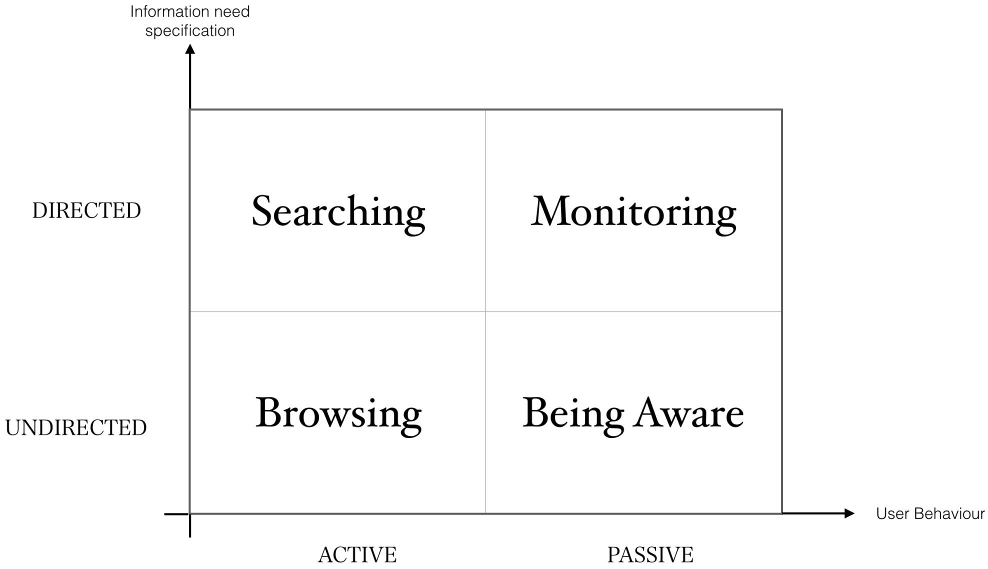


Figure 3.1: Taxonomy of information seeking (Bates, 2002)

tion to view information seeking as "the purposive seeking for information as a consequence of a need to satisfy some goal". This definition clearly underlines that information seeking is goal-oriented.

Wilson (2000) also proposed a nested taxonomy to connect information seeking behavior to other related concepts, including *information behavior*, *information use behavior* and *information searching behavior*. The first term, **information behavior** is often used as an umbrella term that encompasses both **information seeking behavior** and **information use behavior**. The former underlines the seeking process, whereas the latter emphasizes the consumption of information, which can be exemplified by reading information and meanwhile updating the existing knowledge in mind. The last term, **information searching** is viewed as micro-level interactions embraced by information seeking behaviors (e.g. specifying or refining search terms). Nevertheless, information seeking does not only encompass search as in Wilson's nested model. A more elaborated taxonomy of information seeking behaviors was put forward by Bates (2002). As shown in Figure 3.1, Bates' taxonomy includes two dimensions, *user behavior* and *information need specification*. In the former dimension, **Active** and **Passive** refer to whether an information seeker actively acquire/forage information or passively absorb/receive information. In the other dimension, **Directed** and **Undirected** deal with whether an information seeking process is specific or involves random information.

Direct and active information seeking is what we called *information searching* or *searching*, other categories include monitoring, browsing and being aware. This section is concerned with information seeking with particular emphasis on information searching behaviors. We review a selection of prominent models and principles in the field. It should be noted that many theories to be presented in this section are not limited to Web information seeking, but may also apply to going in quest of other resources, such as book or knowledge from others.

3.1.1 Information Seeking Models

Various information seeking models have been proposed. Some focus on the abstraction of different stages of the information seeking process, some emphasize information seeking is an iterative process, and others attach feelings, thoughts and contexts to the seeking process. We review a selection of models that fall into each category.

Stage Model

Robertson (1977) proposed a model of information retrieval, which is perhaps the earliest model aimed at characterizing the information seeking process. In this model, an information seeker first formulates an information need, then s/he turns the need into queries in an information retrieval system, which returns matched documents in a certain representation. Finally the required information is selected. This model abstracts the basic steps for interacting with information retrieval systems, and it has promoted the understanding of the process in its days. However, it overly simplifies the behaviors of the information seeker, which should include more complex activities other than making queries and selecting results.

A number of models were proposed thereafter to rectify the problem of overlooking the information seeker in Robertson's model. Some of the new models describe the information seeking process with concrete human actions from the user's perspective. A oft-cited model as such is the one proposed by Sutcliffe and Ennis (1998), who decomposed the information seeking process into a sequence of actions:

- Problem identification
- Articulation of information need(s)
- Query formulation
- Results evaluation

According to this model, information needs originate from real problems. After querying an information retrieval system, the information seeker also has to evaluate them. It stresses that the information seeking process may go through cycles of the stated actions in case that the returned results do not meet the user's needs after evaluation. Similar models are also proposed in literature (Ellis, 1989; Marchionini and White, 2007; Meho and Tibbo, 2003; Shneiderman et al., 1997). These models mainly differ in the granularity of abstractions, but are common in the assumption that an information seeking process starts from recognizing the need. This is true in most cases, but not always. Sometimes information is encountered and used without an explicit need, e.g. when you browse the web and accidentally find

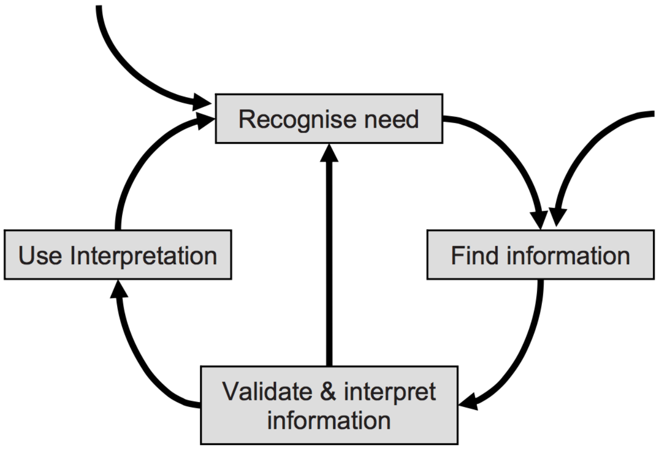


Figure 3.2: The information journey model (Blandford and Attfield, 2010)

information that answers a long-standing question which is out of the current context, you may then deviates the search in a new direction. The information journey model proposed by Blandford and Attfield (2010) takes this into account by allowing multiple entry points, as shown in Figure 3.2. In other words, an information seeking process may start from the *Find information* stage directly. The *Validate & interpret information* stage does not only involve literally interpreting the results, but also include contextualizing them to the current situation. The *User Interpretation* phase refers to making goal-related decisions based on the found information.

Another model that is worth discussing is the one proposed by Kuhlthau (1991). Though the behavioral stages are logically not dissimilar to the models presented before, the major improvement in this model is that Kuhlthau associated the feelings, thoughts and actions during the information seeking process, resulting in a more phenomenological rather than cognitive model (Wilson, 1999). This is the first model that investigates the affective aspects in the process of information seeking along with the cognitive and physical aspects(Kuhlthau, 2005). Kuhlthau's model is greatly influenced by the personal construct theory (Kelly, 1963). The key idea is that the information seeking involves personal construction in a sense that information seekers actively pursue an understanding or construct meaning from the information encountered during the seeking process. As Figure 3.3 illustrates, in the *Initiation* phase, the information seeker feels uncertain about some knowledge and has vague thoughts about the to-be-searched information, which corresponds to the action of recognizing the information needs. Then the user selects a generally relevant topic to start the searching process (*Selection*), constructing a brief sense of optimism. The optimism gives its way again to an increased level of uncertainty or confusion as the user is exposed to more and more "relevant" information (*Exploration*). The confusion is essentially due to an inability to precisely express what information is needed. As search continues, the user focuses more on specific topics (*Formulation*). As more pertinent information is collected (*Collection*), satisfaction, relief and confidence are increased so that the information seeker can complete the search with a new

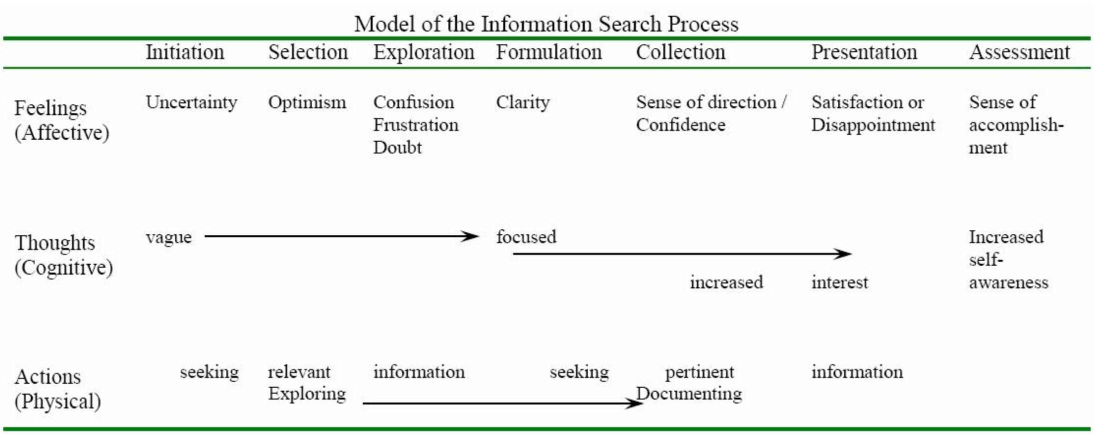


Figure 3.3: Model of information search process (Kuhlthau, 2004)

understanding of the topic that is ready to be presented or explained to others (*Presentation*).

The essential argument in Kuhlthau's model is that an information seeker constructs knowledge during the information seeking process, without which it is difficult for him/her to do it alone. This reminds us of a similar concept reviewed in the previous chapter, i.e. Vygotsky's Zone of Proximal Development (ZPD). Kuhlthau (2004) acknowledges his debt to the ZPD and coined another term called *Zone of Intervention in the Process of Information Seeking*, or simply *Zone of Intervention (ZI)*, which is defined as the "area in which an information user can do with advice and assistance what he or she cannot do alone or can do only with difficulty". Interventions within ZI, no matter realized by technologies or humans, enable information seekers to move along the search process for accomplishing their tasks. However, interventions that are outside ZI may be either overwhelming or unnecessary.

Berry-picking and Information Foraging Model

The stage models presented before focus on abstracting the information seeking process which is decomposed into a sequence of stages or steps. Some models assume that users' information needs keep unchanged during the information seeking process (Robertson, 1977; Sutcliffe and Ennis, 1998). However, as Morville and Callender (2010) have pointed out, the information needs may change in the searching process. Precisely, Attfield et al. (2008) identified a reciprocal relationship between information needs and findings, i.e. "information seeking is shaped by the needs of the task, and yet the evolving task is shaped by the information found". In fact such reciprocal relationship is acknowledged in some of the previously discussed stage models, such as the information journey model and Kuhlthau's model. However, both models center on identifying global seeking behaviors rather than revealing how the search process evolves. The latter aspect is exclusively covered in the Berry-picking model (Bates, 1989), which draws an analogy between seeking information on the web and picking berries in the forest, where the berries are scattered on different bushes, through which a berry-picker moves

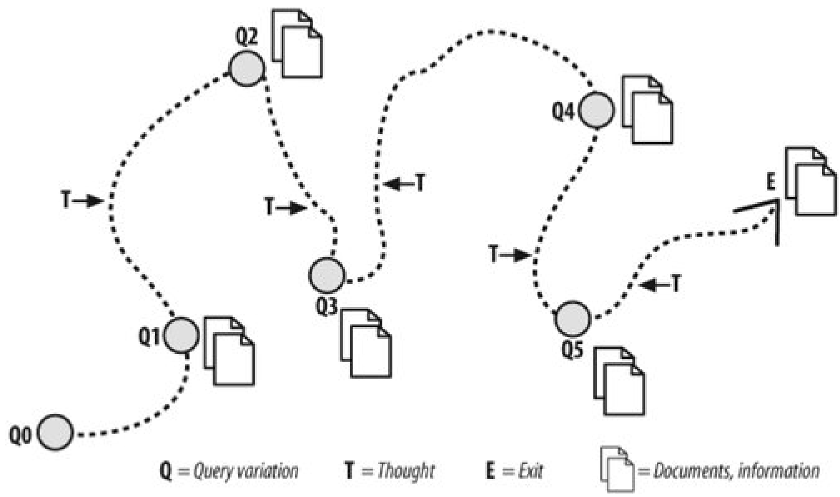


Figure 3.4: Berry-picking model (Bates, 1989)

in order to collect berries. Similarly, an information seeker usually starts with one piece of information, and s/he often has to traverse various resources. The encountered information may either reinforce the original goal or trigger refined goals that lead to new directions in the search process. One thing to note about this model is that the evolving search process implied by the berry-picking model only takes place if the information seeker is also the information user, as "the progression of the information sought is subject to the user making continual judgments regarding its relevancy and interoperability" (Knight and Spink, 2008).

The Berry-picking model acknowledges that information seeking is an iterative process which requires refining information needs and traversing a variety of information resources. But this model does not answer what drive and guide the information seeker's "journey" towards the right information. This realm is covered by the *Information Foraging Model* (Pirolli and Card, 1999). Similar to the Berry-picking model, Pirolli and Card draw an analogy between human seeking information and wild animals hunting for food. The analogous animal behaviors were first studied by a group of biologists, MacArthur and Pianka (1966), who investigated how animals decided what food to eat, where to find them as well as their food foraging strategies. They proposed an *optimal foraging theory*, which asserts that animals forage in an environment scattered with patches of food. After finishing the consumption of food in one patch, the animal moves towards a new patch. The foraging strategy follows the so-called *marginal value theorem* (Charnov, 1976), which states that animals try to maximize the amount of consumed food within a given amount of time. In other words, animals perform a cost/benefit analysis before moving to the next patch in order to achieve maximum benefit with minimum effort. Research study (Pyke et al., 1977) has proved that animals are very good at this and one explanation is that the animals' foraging strategy is guided by the scent of preys.

By analogy, Pirolli (2006) argues that human information interaction systems "tend to maximize the value of external knowledge gained relative to the cost of interaction" and information

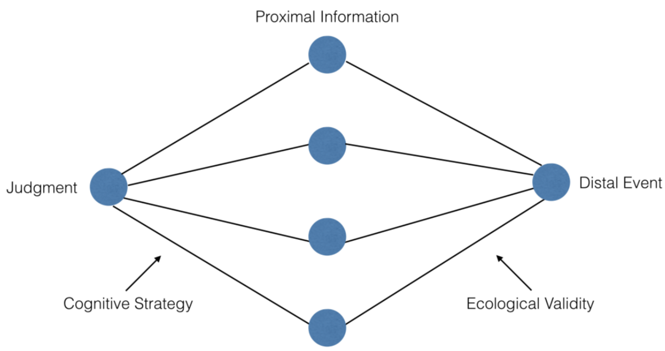


Figure 3.5: The lens model (Brunswik, 1956)

seekers can also be guided by "scent" during their information seeking process. Such "scent" is called *information scent* (Pirolli and Card, 1999). Unlike the scents of preys, information scents are not olfactory cues, but proximal information cues that help the information seeker to judge distal information sources and to navigate towards them. Pirolli (2006) acknowledges the concept of information scents is in debt to the Lens model (Brunswik, 1956), which was originally proposed as an ecological theory to describe how organisms perceive a distal (unobservable) criterion, through proximal (observable) cues (Figure 3.5). In the context of information seeking, the distal object is the information to be sought (e.g. a Website), which is not directly seen by the information seeker. Instead, the distal objects are represented as mediating information that is known as proximal cues that guide the users to make judgment about the potential value of going after the distal object. For example, a picture of classroom (proximal cue) may indicate that navigating through this link would lead to something related to learning (distal object). A book page full of complex mathematical formulas (proximal cue) strongly highlights its connection to the subject of science (distal object). Clearly, the associations between the distal objects and proximal cues are not always direct. And of course, the more direct the association is, the stronger the scents are. Nevertheless, we know that information seekers can make successful judgment under certain conditions of indirect associations. Several cognitive theories can explain this phenomenon. Anderson and Milson (1989) claim that human memory is able to retrieve past experiences that are relevant to the ongoing proximal context, which helps them make judgments. Additionally, Pirolli (1997) proposed a spreading activation model of information scent. The key idea of Pirolli's spreading activation model is not dissimilar to the one reviewed in the last chapter. It also stresses that the proximal cues may activate unobserved features based on the strength of associations stored in the memory.

3.1.2 Information Seeking Principles

Though various forms of information seeking behaviors exist, Buzikashvili (2005) found that any information seeking behaviors could be described by one of the following two principles: (1) *Principle of Least Effort (PLE)* (2) *Principle of Guaranteed Result (PGR)*. The former is regarded as a golden-rule principle in information seeking, whereas the latter was originated in early mediated search in library which is now dying out. Considering mediated search, though not carried out by librarians but by computers, is an important theme in this dissertation. Both principles are reviewed and compared in this section.

Principle of Least Effort

The PLE has a well-known origin in linguistics research. The principle was proposed by Zipf (1949), and was initially known as Zipf's Law, which mathematically states that the frequency of a word decays as a power law of its rank. As an example, the word "the" is the most frequently occurring English word, accounting for nearly 7% of all word occurrences. The second-place word "of" accounts for around 3.5%, followed by the third-place word "and", at about 2.8%. Clearly, the second-common word is used nearly as half as frequently as the top one, and the third-common word is roughly one-third as common. The frequency of words is close to inversely proportion of its rank ($P_n = 1/n^a, a = 0$). Beyond its mathematical beauty, the most significant implication of the Zipf's law is its indication that humans try to minimize their effort when using the words. Manning and Schütze (1999) explain the phenomena as "the speaker's effort is conserved by having a small vocabulary of common words and the hearer's effort is lessened by having a large vocabulary of individually rarer words so that messages are less ambiguous. The maximally economical compromise between these competing needs is argued to be the kind of reciprocal relationship between frequency and rank that appears in the data supporting Zipf's law."

Zipf's law was soon delineated as the principle of least effort, underlying the human nature of "adopting convenience" and was applied in many other domains. In information seeking, it exclusively refers to a user's preference to adopt easier information resources (Connaway et al., 2011; Liu and Yang, 2004). In the study conducted by (Liu and Yang, 2004), the authors investigated a sample of distance education students to study what library resources they used most and why. It was founded that the Internet was the most frequently used resource, followed by libraries. The students reported they chose these resources due to their quickness and convenience to access. However, information being convenient to access does not guarantee its quality. This was highlighted in a similar principle called "satisficing" (Byron, 2004). The word "satisficing" is man-made, composed of "satisfy" and "suffice". It stresses that the information seekers are often in favor of information convince while compromising information quality.

Principle of Guaranteed Result

Before the 1980s, searching information with computers was exclusively carried out by intermediaries, i.e. trained librarians, and this was called *mediated search*. Library users at that time had to ask experts in the library to search documents for them. With the prevalence of user-friendly search interfaces in the recent 30 years, mediated search is giving its way to user-directed *unmediated search.* However, some of the proven principles are still informing. Buzikashvili (2005) stated that in mediated search an intermediary does not abide by the principle of least effort, but follow the principle of guaranteed result, which favors the completeness of query rather than convenience of access. He further identified the differences between how unmediated searchers and intermediaries seek information (Table 3.1) and explained why they use different search tactics.

Searcher	Ability to recognize pertinence	Searcher's aim	Searcher's tactic at each step
Unmediated searcher	Yes	Precision of search results	Least effort tactic
Intermediary	No or Partial	A most complete query	Max coverage tactic

Table 3.1: Model searchers in unmediated and classic mediated search (Buzikashvili, 2005)

In summary, unmediated searchers are fully aware of their own information needs, so they are looking for very precise information. This is the situation where PLE applies. In contrast, though an intermediary, such as a librarian, knows roughly what an information seeker is looking for, the understanding of the information needs is usually partial and uncertain (Kuhlthau, 1993). S/he has to make more complete queries so as not to not leave something out (Nordlie, 1996), and present the user with a more comprehensive set of results that corresponds to all possible interpretations of the end user's information need. Put it in other words, intermediaries adopt max coverage tactic to "compensate" their lack of knowledge in recognizing the actual needs of their patrons. This message is still informing the design of computer-based intermediaries for searching nowadays, i.e. if a computer system cannot clearly understand the patron's information need, it is better to perform general queries and return more comprehensive results, even if the such results inevitably contain redundancy.

3.1.3 Context

As far back as the early 1930s, context was defined by Dewey (1931) as "a spatial and temporal background which affects all thinking and a selective interest or bias which conditions the subject matter of thinking". This is rather a philosophical definition of context, but yet clearly points out its spatiotemporal nature as well as its influence on people. The last few decades have witnessed a tremendously growing trend in recognizing and applying context in information science. In the 1980s, Wilson (1981) proposed an information seeking model that identifies the certain characteristics of the information seeker and the environment influence the seeking behavior. Another influential theory at the time that emphasizes context is *sensemaking* (Dervin, 1983), which suggests that the role of information seeking is to fill the gap between the contextual situation in time and space where the problem arises and the desired situation of the outcome.

Many perspectives of context have been explored in information science literature, but a clear definition of context is difficult to find. As Dervin (1997) complained, "there is no term that is more often used, less often defined, and when defined, defined so variously as context." Some researchers stress the situational nature of context. Schilit et al. (1994) describe context as "where you are, who you are with, and what resources are nearby". Dey et al. (2001) define it as "any information that can be used to characterize the situation of an entity". Other researchers attempt to relate *relevance* to context. For example, Saracevic (2007) viewed context as an element of relevance, and describe it as complex and dynamic "interaction between a number of external and internal aspects" of the human and the environment. According to Mizzaro (1997), context "includes everything not pertaining to topic and task, but however affecting the way the search takes place and the evaluation of results". This definition implies that users do not necessarily express clearly the context in their search queries, but computer systems should otherwise try to model it to help them with searching and evaluation.

In fact, various definitions of context reflect a similar motivation for studying context in information behaviors. Freund and Toms (2013) summarize two aspects of the motivation from (Johnson, 2003) (1) Context serves to disambiguate meaning, which is especially important in human communications. A good understanding of the context makes communications effective. (2) Context also shapes and deliminates social action. Patterns of behaviors are easier to identify among small groups engaging in common activities than among overall population. Both aspects adhere closely to the scope of this dissertation, which explores the possibilities of considering the context of the learning activity for the design of information scents. In the reminder of this section, we review the various spheres into which the context can be deconstructed.

Contextual Sphere

While the notion of context is usually vaguely defined, researchers have attempted to deconstruct it into concrete conceptual spheres. Wilson (1981) proposeed that the basic information needs can be physiological, cognitive or affective, and the context of these needs may be the person himself or herself, or the role demands of the person's work or the environments (political, economic, technological) within which the work takes place. In his later mode of information seeking behaviors, Wilson (1997) reformulated the aforementioned context as follows:

- **Psychological variables,** which describe the information seekers' personal characteristics and emotions. These include attitudes, preferences, interests, styles of learning.
- **Demographic variables**, encompassing sex, age, and economic, education status and professional experiences.
- **Social or interpersonal variables,** including the personal roles, regulations and rules that shape the information seeker's behaviors
- **Environmental variables,** covering the time and space dimensions of the information seeking, interruptions, facilitations as well as information ecology (e.g. individual or collective)

- **Information source characteristics,** which can be seen from three aspects: the ease of information access, the appropriateness and credibility of information, the channel of communication.

Wilson claims that these contextual factors, termed as *intervening variables* in the paper, influence not only the occurrence of information needs, but also the way a need is perceived and satisfied. Wilson is not the only researcher who attempts to make explicit different contextual spheres, but it is one of the most comprehensive one. Most of the other frameworks have similar constituent parts. For example,Melucci (2012) considers four types of contextual variables: content, geographical, interaction and social variables.Myrhaug and Göker (2003) proposed an AmbieSense user context model consisting of five components: environment, personal, task, social and spatiotemporal. Compared to Wilson's framework, these models are largely identical.

3.1.4 Modes

So far we have reviewed the stages of information seeking, the principles that an information seeker follows as well as the context which influences the search behaviors. All of these aspects are behind a motive that drives the user to seek information, and motives as such can be categorized into various *search mode*. In this section we review a few research efforts on search modes. The review not only delivers an understanding of the commonly recognized patterns of information seeking behaviors, but also informs the design of information systems to better fulfill users' needs.

One of the earliest study about search modes was conducted by O'Day and Jeffries (1993), where the authors studied the mediated search behaviors of professional intermediaries on financial and business-related topics and identified three modes of search:

- *Monitoring* a well-known topic or set of variables over time. For example, a financial analyst wants to track revenue and order growth of a company.
- *Following a plan* for information-gathering suggested by a typical approach to the task at hand. For example, a business benchmarking specialist follows a strategy to screening the companies that offer the best service in a certain field.
- *Exploring* a topic in an undirected fashion. For example, a management consultant explores many facets of a company in order to give advices.

These three modes describe generic types of search. In the same paper, O'Day and Jeffries (1993) further categorized six types of search techniques:

- Looking for *trends* or correlations

- Making *comparisons* of different pieces of the dataset
- Experimenting with different *aggregates* and/or scaling
- *Identifying a critical subset* of relevant or unique items
- *Making assessments*
- *Interpreting* data to find meaning in terms of domain or problem concepts

The above six types can be seen as atomic search activities that involve analysis and sense-making, suggesting that search interfaces should be designed to support the above search activities of the intermediaries. O'Day's framework is derived from old-fashioned mediated search, especially in the business field. After all, the intermediaries are in analogy to today's typical web searchers, so the modes and activities identified by the model are still informing. For example, by knowing that potential buyers typically search and compare the features and prices of similar products (i.e. making comparisons as indicated by point 2 of the search types), an e-commerce website should offer the possibilities for the users to perform such activities at ease.

An important mode identified by O'Day's framework is that the web search is usually "exploring a topic in an undirected fashion" (O'Day's search mode 3). This aspect is made explicit in another well-cited framework by Marchionini (2006), who puts an emphasis on exploratory search. Marchionini categorizes search activities into three types as follows:

- **Lookup**, which corresponds to searches that are carefully specified. Usually such type of search returns precise set of results that require little examination. Example search activities of Lookup type include fact retrieval or known item search, as illustrated in Figure 3.6.
- **Learn**, which involves search activities that require users to iteratively comparing and assessing the results, corresponding to the lower layers of the Bloom's taxonomy of learning activities, such knowledge acquisition, comprehension to aggregation and application.
- **Investigate**, which refers to search activities that typically require longer time and high order of cognitive processing, corresponding to the higher layers of the Bloom's taxonomy of learning activities, such as analysis, evaluation and creation.

Lookup searches are viewed as simple "turn-taking" processes, where information seekers input queries and the search system return retrieved results as responses. Both learn and investigate searches are seen as exploratory processes where information seekers are highly involved for making sense of the search results.

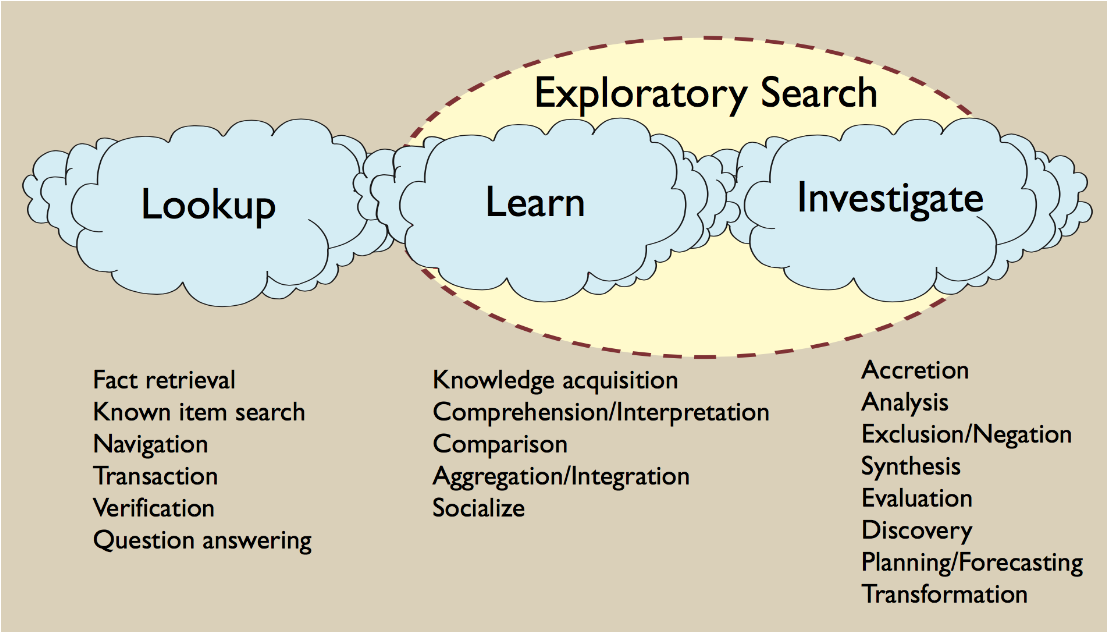


Figure 3.6: Taxonomies of search activities (Marchionini, 2006)

Marchionini (2006) described a holistic taxonomy of search activities, but the modes are too generic. In contrast, Russell-Rose et al. (2011) observed 104 enterprise search scenarios and identified a set of 9 search modes, which are then grouped according to the Marchionini's taxonomy:

- Lookup
  - Locate: To find a specific item (possibly known) item.
  - Verify: To confirm that an item meets some specific criterion
  - Monitor: To maintain awareness of the status of an item for
- Learn
  - Compare: To identify the similarities and differences of two or more items
  - Comprehend: To generate insight by understanding the meaning of an item
  - Explore: To examine an item for the purpose of serendipitous knowledge discovery
- Investigate
  - Analyze: To examine the details of an item to identify patterns and relationships
  - Evaluate: To judge the value of an item with respect to specific benchmark
  - Synthesize: To generate insight by integrating diverse inputs to create a novel artifact or composite view

The modes presented before do not necessarily occur exclusively during the course of a search activity. Instead, two or more modes may form distinct chains. For example, if a business analyst wants to understand market trends, s/he may follow a pattern of Analyze-Compare-Synthesize. Russell-Rose et al. (2011) claims that the true value of the modes lies in that they help recognize common search strategies so that the designers may accommodate new features to improve the corresponding discovery experiences. The modes are derived from the enterprise search domain, but in Chapter 5 we will show how they fit to the learning context.

3.2 Help-seeking

So far we have reviewed a variety of influential theories of information seeking. This section presents a closely related concept, i.e. help-seeking. Most of presented theories in section 3.1 are derived from information searching on the Web. Though many of them may be well applied to seeking help from books, humans or other kinds of resources, research in help-seeking in fact has a relatively standalone development, with a particular emphasis on help seekers' psychological barriers for asking for help.

Ames and Lau (1982) defined help-seeking as "an achievement behavior involving the search for and employment of a strategy to obtain success". Historically, research in academic help-seeking behaviors mostly focus on human intervention, and students who exhibit such behaviors are viewed as immature and incompetent (Nelson-Le Gall, 1985). Researchers also claimed that seeking help from others may cause damage to self-esteem (Nadler and Fisher, 1986). However, help-seeking, if properly conducted, are nowadays often viewed as beneficial to learning (Karabenick, 1998; Lee, 2007; Polson and Richardson, 2013). In literature there exist a variety of interpretations of the benefit of help-seeking. Luckin et al. (1999) take a Vygotskian view on help-seeking behaviors, positing that such behaviors may help the students to deal with complex problems which are beyond their own capabilities. Aleven et al. (2006) interpret help-seeking as a meta-cognitive skill : "The ability to solicit help when needed, from a teacher, peer, textbook, manual, on-line help system, or the Internet may have a significant influence on learning outcomes." In addition to the proposition that help-seeking may influence learning, Aleven's view clearly illustrates various "help resources" that a learner can turn to, from a more knowledgeable other (MKO) to computer-based support. This section reviews the concept of help-seeking behaviors, the factors that impede help-seeking behaviors as well as a selection of models that demonstrate the processes of help-seeking.

3.2.1 Help Seeking Behavior

Help-seeking behaviors can be effective or ineffective, depending on what strategies are employed. Nelson-Le Gall (1985) makes a distinction between *instrumental* and *executive* help-seeking. The former refers to a "mastery-oriented" process, through which learners focus on knowledge acquisition from the helping resources. Indirect help, hints and explanations from the third party (a person or tools) may serve for this purpose. Help-seekers achieve the best effectiveness through *adaptive help seeking* (Newman, 1994) when they know when and why help is needed, what kind of help is needed, whom to ask for help, and how to ask for the help (Ogan et al., 2014). Adaptive help-seeking is a strategy often employed by self-regulated learners, and it is seen to increase the likelihood of long-term mastery and autonomous learning (Newman, 2002).

Executive help-seeking, on the other hand, is often referred to as *help abuse* (Aleven et al., 2006) or *dependent help-seeking* (Newman, 2008). It refers to the situations when students unnecessarily overuse the help by having a third party (a person or a tool) solve problems for themselves, without making sufficient effort on their own. The learners are more interested in achieving the final outcome, i.e. having the task done rather than thinking through the subject by themselves. In contrast to abusing help, another extreme case is to avoid it, and such behaviors are called *avoidance of help seeking*. In such situations, students refuse to ask for help even when they are aware of the need. Instead, they either passively do nothing or attempt other ineffective strategies. Aleven et al. (2006) classifies both *dependent help-seeking* and *avoidance of help seeking* into the category of *nonadaptive help-seeking*, which may lead to less effective learning (Ryan et al., 2001).

Factors impeding Help Seeking from Human

An intuitive question following the discussions above is when and why individuals may choose to avoid help. This is very important for us to design systems that offer helps to the students. Wacker and Roberto (2008) attempted to use well-established psychological theories to explain help avoidance for health problems, and much of the explanations can also be applied in the learning context:

- **Reactance Theory** (Brehm and Brehm, 2013): Individuals value freedom and autonomy, and when their values as such are challenged, negative psychological states (reactance) that attempt may arise to restore the value. In the learning context, students may refuse help from others because it would be tantamount to admitting that they depend on other people to succeed, thus threatening the value of autonomy (Eisenberg et al., 2007; Ferla et al., 2010).
- **Attribution Theory** (Kelley et al., 1972): Individuals formulate attributions to reason about why certain things happen. When an individual is deciding whether or not to receive help from another person, s/he might think about the helper's motive for providing the help. Is it from a genuine concern of him/her, his/her role demands it or other ulterior motives (Fisher et al., 1983)? A student is less reluctant to receive help from his/her teacher, because such help conforms to the teacher's role. In addition, students may be willing to ask for help if they perceived that many others may have similar questions. (Schwartz and Tessler, 1972).
- **Equity Theory** (Walster et al., 1973): Individuals expect to maintain a reciprocal re-

lationship with others, so negative psychological states would occur if they perceive dissymmetric contributions when interacting with others. In the learning context, if a student feels he could compensate the help in other forms, s/he would be more willing to accept help from another person.

- **Threats-to-Self-Esteem** (Fisher et al., 1983): Individuals' attitude towards help-seeking can be either self-defeating or self-enhancing (Elliot and Church, 1997; Skaalvik, 1997). If a student perceives inferiority and believes the helper will look down upon him/her, the student would be reluctant to ask for help (self-defeating attitude). Conversely, if a student feels positive about the help, s/he would be more likely to seek help (self-enhancing attitude).

Other barriers that impede students from asking for help may also attribute to negative perceptions of the usefulness of the help (Eisenberg et al., 2007, 2009) or low self-efficacy, i.e. the learner's self-evaluation of their capabilities to successfully complete the task (Schunk et al., 2008). In fact the learners who need help the most are least likely to ask for it (Ryan et al., 1998).

Factors impeding Help Seeking from Help Systems

In the preceding section we discussed the factors that impede students from asking for help from a presumably *more knowledgeable other* (MKO). The discussions were particularly embraced in an academic learning context. However, help may also be sought while a student is performing a task (e.g. a collaborative brainstorming task) with a computer application. Aleven et al. (2003) claimed that help aimed at supporting task performance does not always lead to better learning, and vice versa. The key differences between pursuing aids from a MKO in an academic learning context and using an application help system when encountering difficulty in a task is that the latter often turns out to be *executive* rather than *instrumental* help-seeking. Given a task to be completed, individuals need immediate remedy of the problematic situations or information that supports the task. Therefore, obtaining a direct answer is much more favorable than achieving mastery learning. Though psychological barriers such as *Threats-to-Self-Esteem* may not come into play in interacting with computer-based help systems, researchers have found that help systems embedded in applications are experiencing low usage frequency (Cool and Xie, 2004; Fisher, 1999). Dworman and Rosenbaum (2004) identified 5 reasons why users do not use help systems within the application:

- **Cognitive blind spots**: Similar to the well-known *banner blindness* (Benway and Lane, 1998) phenomenon in the web, users are experiencing cognitive blindness to the help even it is shown right in front of them.
- **Distraction aversion**: Seeking help often result in diverting away from the current task, which users are often reluctant to do.

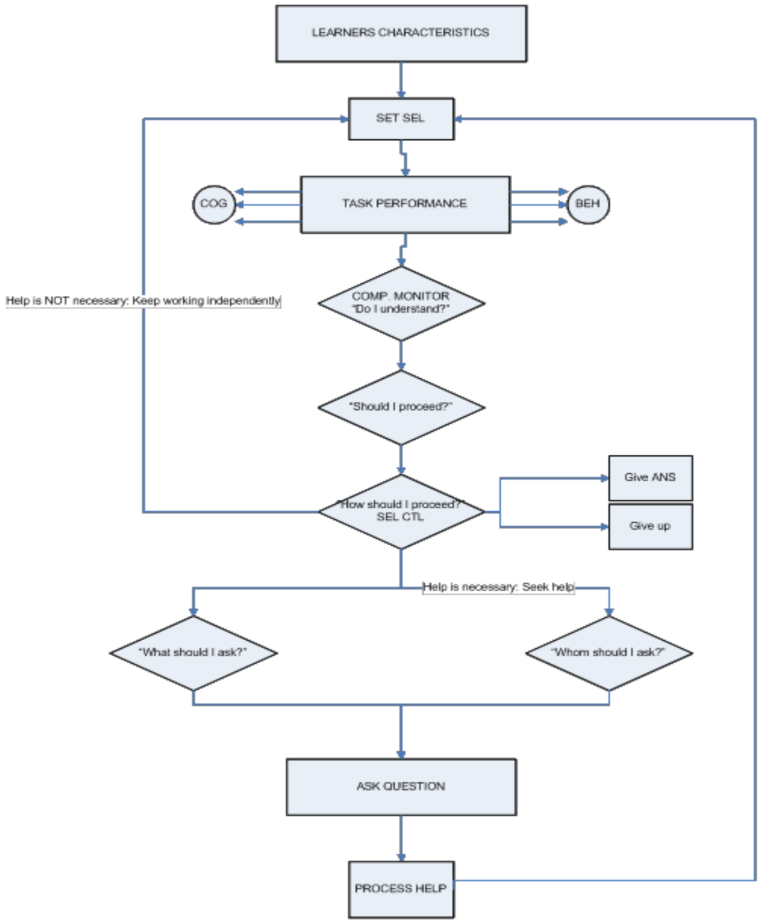


Figure 3.7: Adaptive help-seeking model (Newman, 1994)

- **Fear**: Users may have previously experienced failures with the help system, so they are afraid to fail again.

- **Refusal to admit defeat**: Similar to the concept advocated by the previously discussed reactance theory, users refuse to admit their inability to deal with the problem, and they believe that they can handle it without looking for help.

- **"Rose by another name"**: Users tend to access hints, tips or guides, but they are not willing to click something labeled "help".

Similarly, Purchase and Worrill (2002) reported that users also complained that help systems were often misleading or incomplete, difficult to navigate, or did not contain enough examples.

3.2.2 Help-seeking Models

Decades of research in help-seeking behaviors have contributed a few theoretical frameworks to model help-seeking processes. The most influential model, among others, is the one by Nelson-Le Gall (1981), who posits that a help-seeking process comprises the following steps:

1. **Become aware of need for help**: Individuals realize the problematic situation they are facing as well as the need of help to tackle the problem
2. **Decide to seek help**: Individuals contemplate the environment and the task, and decide to turn for help.
3. **Identify potential helper(s)**: Individuals select available help resources or humans that are believed to offer proper help
4. **Use strategies to elicit help**: Based on the knowledge and experiences of the help seekers, they express their help requests with suitable strategies.
5. **Evaluate help-seeking episode**: As a final step, individuals must retrospectively evaluate the effectiveness and helpfulness of the help.

The above 5 steps are not necessarily as sequential as it appears. For example, steps 3 and 4 may be iterative, i.e. if help elicitation fails, individuals may attempt to identify new helpers. Nelson's model is overly simplified, because it does not depict personal affections and thoughts during making help-seeking decisions. These aspects are considered by the *adaptive help-seeking model* proposed by Newman (1994), which was already described in the preceding section as "individuals know when and why help is needed, what kind of help is needed, whom to ask for help, and how to ask for the help". With the "five Ws" as conditions, the description implies that help-seeking is a complex and constructive decision-making process, which is portrayed in Figure 3.7. In this flow chart, SEL stands for "self-efficacy level", which refers to one's self evaluation of own capability to accomplish the task. In learning, the higher the SEL is, the more efficacious the learner rates his/her own capability. CTL is the abbreviations for "confidence tolerance level", indicating one's own preference for taking challenges or risks. The higher the CTL, the more willingly the learner takes the challenge for resolving difficulties. When engaging in a learning task, a learner constantly poses several meta-cognitive questions : "Do I understand?" ,"Should I proceed?", "How should I proceed?". If the answer to the last question is yes, then the SEL and CTL is compared. If SEL is above the CTL, the learner thinks it unnecessary to seek help and decides to work independently. Otherwise, the learner may perceive external help as necessary, and then starts identifying the helpers and elicit help.

The models proposed by Nelson-Le Gall (1981) and Newman (1994) were originally based on seeking help from human subjects. Given that many computer-based learning tools also provide on-demand help, Newman (1994) proposed a model exclusively designed for computer-based interactive learning environments. Aleven's model shares some common

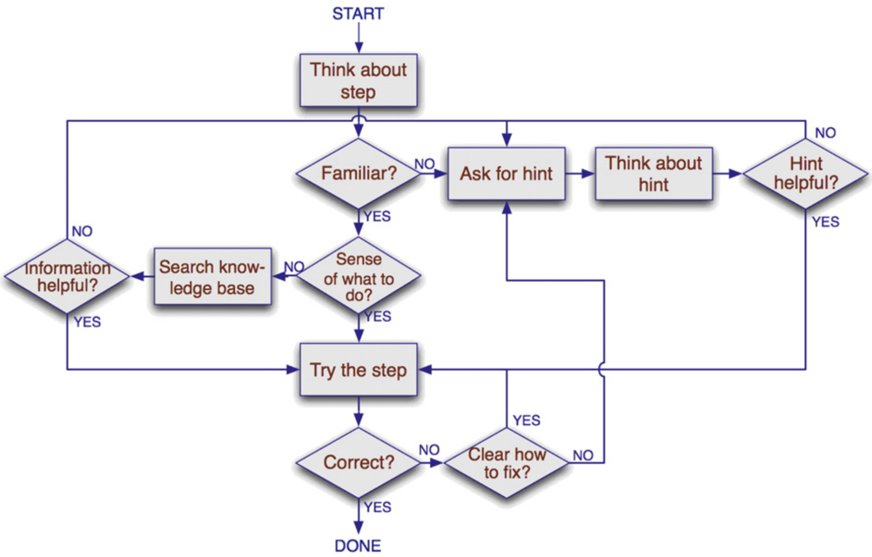


Figure 3.8: Help-seeking model in interactive learning environment (Aleven et al., 2006)

traits of the former two models, but provides a more fine-grained anatomy of the help-seeking process. Newman (1994) put forward two types of on-demand help, context-sensitive hints and de-contextualized knowledge base (i.e. glossary). The hints are provided with different levels of details to give students specific advices, whereas the glossary simply displays definitions, theorems, rules and principles that are not tailored to specific context. Similar to Newman's model, Aleven's model is also characterized by a set of meta-cognitive actions. At different stages of a working task, students constantly make self-evaluations of their knowledge and the effectiveness of help. For example, a learner starts the problem-solving task by first thinking about the steps that should be followed. If s/he perceives little familiarity of the step from the beginning, s/he would ask for hint. Otherwise, the students would sense what to do next. Failing to do so would lead to searching glossary help. The complete processes and conditions of Aleven's help-seeking model is illustrated as flow chart in Figure 3.8.

4 Perceive and Interact with Information

People frequently need information, in various occasions. Carroll et al. (2003) view information needs as "one of the few timeless, transcultural constants" in the world. Chapter 3 is dedicated to outline how people interact with information, yet with a limited scope on the most representative type of information seeking (Bates, 2002), where information seekers take the initiative to look for specific information or help. However, the most prominent power of computer lies in automation. In other words, computer technologies can be designed to make useful information available or even ubiquitous, so that users can directly consume information, which corresponds to the passive information seeking behaviors, according to Bates' taxonomy (2002). In this section, we present a few technologies as such. Ambient display has been explored to allow people to perceive and interact with information on their peripheral attention. Serendipitous interaction emphasizes that useful information that sparks an implicit and longstanding need can be captured "by accident" in the environment, resulting in a "happy coincident". This chapter reviews not only theoretical backgrounds but also practical research efforts made to support and facilitate the aforementioned types of information interactions.

4.1 Interacting with Ambient Information

One way to stay tuned about specific information is to make it easily accessible through repeated checking, known as a strategy of *polling* (Cadiz et al., 2001). Polling interfaces usually rely on the so-called *pull technology*, which involves a user initiates a request to fetch particular information from a computer. Suppose you are reading news articles in Yahoo!, after some time, you might have to refresh the Webpage to see an updated list of articles. According to Bates' taxonomy (2002), polling strategy can be seen as undirected, but active information seeking, i.e. browsing. This strategy is very simple, but potential drawbacks are obvious. Cadiz et al. (2001) summarized three drawbacks as follows: (1) users may miss important events when they cannot pull updates on time, e.g. while engaging in another task. (2) users have to manage increased cognitive burden with polling, since they have to remember to pull information as well as to figure out which of the updated information is new. (3) In case that information is widely distributed, users have to poll from a variety of services.

In the context of article reading, the last drawback can be successfully addressed by technologies like RSS, but the former two remain as thumbscrews. A potential remedy is to employ technologies that are strategically poles apart, i.e. *pushing technology*, which involves computer systems proactively push information to users. Messenger applications such as Whatsapp or Skype adopt push technologies by default, notifying users about arrivals of new messages. In this case, users play a passive role, receiving and consuming information. Depending on whether information need is specified, systems as such can be seen as either *Monitoring* or *Being Aware* type of information behaviors. The proactive notifications of pushing technologies can easily provoke one to think of an obvious negative consequence: users may be distracted from their task.

Therefore, trade-offs between effectiveness and distractiveness of the pushed information must be balanced in the system design. Such considerations gave birth to a special group of pushing technologies, *ambient information systems*, which aim to "convey information via calm changes in the environment, so that the users are more able to focus on their primary work tasks while staying aware of non-critical information that affects them" (Pousman and Stasko, 2006). The key characteristics of ambient information systems are summarized by (Pousman and Stasko, 2006) as they should (1) display information that is important but not critical (2) can move from the periphery to the focus of attention and back again (3) focus on the tangible representations in the environment (4) provide subtle changes to reflect updates in information (5) are aesthetically pleasing and environmentally appropriate. This section reviews a few related terminologies, design patterns, interaction models and evaluation criteria of ambient information systems, and finally presents a selection of ambient information systems particularly in the learning domain.

4.1.1 Disambiguating Terminologies

A variety of terminologies can be found in literature to refer to similar concepts as ambient information systems, including *ambient display* (Ishii et al., 1998; Mankoff et al., 2003), *notification system* (McCrickard et al., 2003), *peripheral display* (Gueddana and Roussel, 2009; Matthews et al., 2003; Stasko et al., 2005), *awareness system* (Visser et al., 2010) and *interruption displays* (Matthews et al., 2003). Some of these terminologies are used interchangeably in literature, while others have distinct emphases.

Ambient display is almost identical to ambient information system, though the former stresses aesthetics and the latter emphasizes information system. Pousman and Stasko (2006) assert that all ambient displays are peripheral displays. This assertion is supported by Matthews et al. (2003), who see peripheral displays as displays that show information that a person is aware of, but not focused on. Matthews et al. (2003) also express the definition from the Activity Theory's perspective and view peripheral display as any information display that (1) is a tool in at least one activity of its user and (2) is used primarily at the operation level rather than the action level. Obviously, ambient displays share the same traits. The authors further articulate that ambient displays usually convey non-critical information, but peripheral displays may include critical information, exemplified by cockpit altimeters.

Pousman and Stasko (2006) claim that only some notification systems are peripheral displays, and others may require focused attention, corresponding to the *Monitor* type in Bates' taxonomy (2002)), e.g. a surveillance system that identifies and notifies suspicious figures. This standpoint is also supported by Matthews et al. (2003), who view notification system as an umbrella term embracing peripheral display, and of course, ambient display as well.

Interruption display, among others, is the most awkward, yet distinct category. Matthews et al. (2003) define it as systems that are intended to attract focused attention to other tasks. In other words, interruption displays intentionally issue alerts that disrupt a user's on-going work, causing an attention shift to another activity. This strategy is fundamentally distinct from that of peripheral or ambient displays, which aim at minimizing distractions. Noteworthily, Matthews et al. (2003) hold that interruption displays include everyday objects such as alarm systems, though they don't necessarily alert visually. Obviously, interruption displays as such serve for different purposes, e.g. alerting important or critical situations. In fact, we cannot avoid being interrupted by telephone rings or home appliances alerts in our everyday life. Sometimes these interruptions are important, but annoying. Section 2.1.3 reviews the potential negative consequences of interruptions at the cognitive level, and several HCI researchers have also pointed out that interruptions can be detrimental to working tasks (Cutrell et al., 2001; Sasse et al., 1999). If a user can not avoid switching his or her attention to another temporarily important task anyway, Matthews (2007) claims that peripheral displays should be applied in such situation, and they should be designed such that a user is not interrupted and can choose to finish her current task first.

To sum up, notification, peripheral and ambient systems appear to be conceptually linked in a top-down hierarchy, whereas interruption displays live apart. Besides, it is not difficult to figure out that alerts themselves do not contain information, so interruption displays do not apply Bates' two dimensional taxonomy (2012), and they are not instances of information systems. In contrast, notification systems as a whole are information systems that support either *Monitor* or *Being Aware* mode of information behaviors. We will keep using the term ambient information system, since the type of systems to be discussed in this dissertation highlights the delivery of information sought.

4.1.2 Ambient Interaction

As one of the first design of ambient technologies, the "Dangling String" (Weiser and Brown, 1996) is hung in an office to indicate network traffic: a busy network would cause the string to whirl. The ambientROOM (Ishii et al., 1998) turns a room into an intelligent environment that displays various subtle information in terms of light, sound and movement, for the sake of background awareness. Sideshow (Cadiz et al., 2001) imposes peripheral awareness of information on the computer screen as sidebar. Reflect (Bachour et al., 2008) reactively visualizes the intensity of speech contributed by each participant in a group discussion for regulating group collaboration. The aforementioned four examples of ambient information systems exploit the potential of conveying information with physical objects, architectural spaces, computer screens and interactive tabletops respectively. Once the users of such ambient information systems are aware of subtle changes in the environment, they selectively react to the information.

Just like these four examples, ambient information systems historically only deal with the perception and monitoring of ambient information, and rarely involve the interaction with them. As research evolves, researchers have realized that ambient information systems can also be interactive, and go beyond simply displaying information.

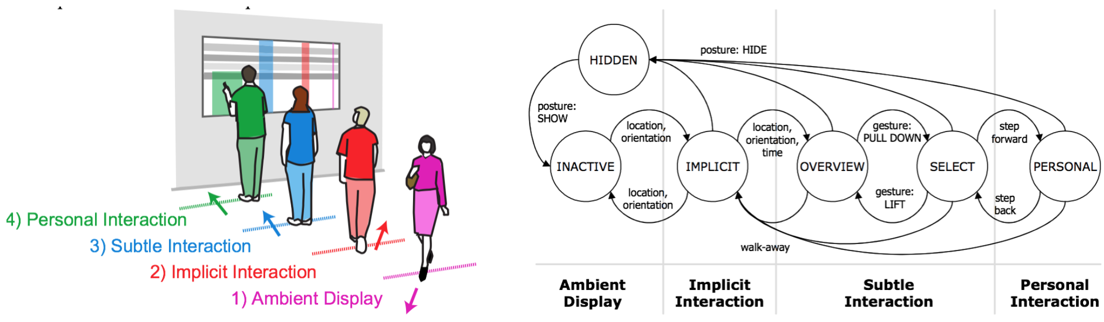
Zone, Phases and Dimensions of Ambient Interaction

The need for interacting with an ambient information systems was first recognized by researchers investigating ambient displays that are deployed in public spaces. The GossipWall system (Prante et al., 2003; Streitz et al., 2003) is one example as such. The GossipWall is a vertical surface that is composed of LED arrays and RFID sensors. The display emits abstract light patterns that can only be decoded by special RFID-enabled hand-held devices. Users holding such devices can be identified to enable informal interpersonal interactions with other people through the wall. The authors posit that ambient artifacts like GossipWall should allow situational interactions that depend on the proximity of people passing by. They distinguish three *zones of interaction*: (1) Ambient Zone: the outer proximal zone where people are simply passing by. General user-independent information is displayed to the passers-by. (2) Notification Zone: An individual enters the zone by approaching to the display. S/he is then identified and notified about "secret" interpersonal messages (3) Interactive Zone: The person can interact with the display when s/he is very close to the wall. The activities of the users within each zone correspond well to the three distinct *activity space* identified by Brignull and Rogers (2003): (1) Peripheral awareness activities (2) Focal awareness activities (3) Direct interaction activities.

The most influential work that discriminates different distance-dependent zones is the interaction phase framework (Vogel and Balakrishnan, 2004), which extends the previously discussed *zones of interaction* to an interaction phase model that is claimed to encompass a wider range of implicit and explicit interaction techniques. As shown in Figure 4.1(a), the framework proposes 4 phases of ambient interaction:

1. **Ambient Display**. Similar to the *Ambient Zone* in Streitz et al's framework (2003), the ambient display should delivery a general sense of the information that anchors potential subsequent interactions.

2. **Implicit Interaction**. This phase generalizes the *Notification Zone* in Streitz et al's



(a) Four interaction phases

(b) State diagram of phase transitions

Figure 4.1: A framework for interaction phases (Vogel and Balakrishnan, 2004)

framework, suggesting that the system should be able to judge the users' openness to receiving information from the display based on their posture, position and orientation. If a user is positive about receiving information, then the system should signal a subtle change in the information presentation style in order to notify the user about something that s/he might be interested.

3. **Subtle Interaction**. If a user is detected to be interested in the implicit information presented in the previous phase, e.g. s/he stops for a short moment and comes even closer, then the system enters into the Subtle Interaction phase, where more detailed and more personalized information should be presented.

4. **Personal Interaction**. In this phase, the user's attention is completely drawn by the display, s/he touches the display and uses gestures to interact with the information, which may last for a longer period of time. This phase, together with the Subtle Interaction phase correspond to Streitz et al's *Interaction Zone*.

Vogel and Balakrishnan (2004) claim that at any given phase, a user may either choose to step back to a previous phase or to give it up and leave the display. The transition of interaction phases can be depicted with a six-state diagram (Figure 4.1(b)). The `HIDDEN` state is when a user explicitly expresses her unwillingness to interact. `INACTIVE` is when a user is far from the display. Usually the information presented on the display has different level of details. When a user is viewing the overall information in the Subtle Interaction phase, s/he is the state of `OVERVIEW`. When s/he decides to query some items, the user `SELECT` a specific piece of information for more fine-grained details.

It is worth mentioning that the 4-phase framework is exclusively based on the distance between the user and the ambient information system. Though the authors elaborated that users' orientation and location can be used to judge their openness of receiving information, the framework does not explicitly include these dimensions. Greenberg et al. (2011) adapted the

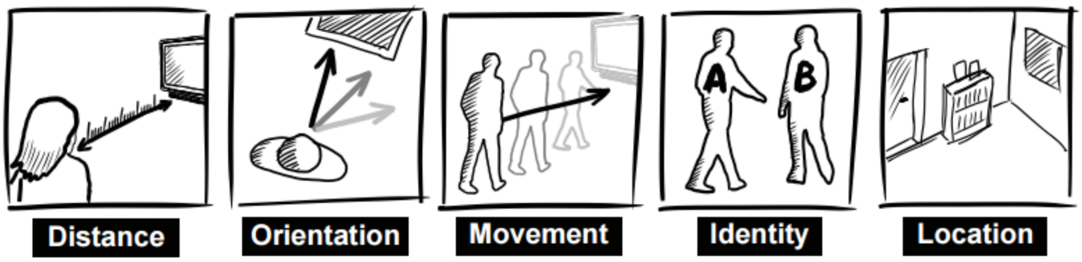


Figure 4.2: Five dimensions of proxemics of Ubicomp (Greenberg et al., 2011)

theory of proxemics from Hall (1969) in the field of cultural anthropology and proposed a slightly different interpretation of *proxemics* in ubiquitous computing: It "concerns inter-entity distance, where entities can be a mix of people, digital devices and non-digital things". In this definition, the inter-entity distance can be characterized in five dimensions as illustrated in Figure 4.2. The *Distance* dimension refers to the longitudinal distance as considered in Vogel and Balakrishnan's four-phase framework. The *Orientation* to the direction to which a user is facing. A system should take actions once it recognizes that the user is looking at the display. *Movement* captures the distance and orientation of a user overtime. Greenberg et al. suggest an ambient system should respond to the user's speed as well as the direction of movement. The *Identify* and *Location* dimensions anchor the context of proxemic interactions. By knowing who the users are and where they are, an ambient system is more capable of tuning its information accordingly.

4.1.3 Ambient Design and Evaluation

Decades of ambient system research have given birth to a number of theories about how design and evaluate ambient information systems. This section starts by reviewing a taxonomy of ambient information system, followed by evaluation criteria.

Taxonomy of Ambient Information System

Given that a variety of ambient displays have been proposed in literature, researchers seek a taxonomy to categorize them. Matthews et al. (2003) propose to classify ambient information systems from three perspectives : **notification**, **transition** and **abstraction**. The notification perspective discretizes ambient notifications into 5 five levels : *ignore*, *change blind*, *make aware*, *interrupt* and *demand attention*. These levels reflect the degrees of importance of information. Notifications are usually delivered through animated transitions at the corresponding notification level. For example, critical information requires higher level of notification and should fall in the *interrupt* or *demand attention*. Ambient information is usually not conveyed directly. Rather, ambient information systems display its abstracted form. Matthews et al. (2003) found that most systems display certain extracted features of the original data. For

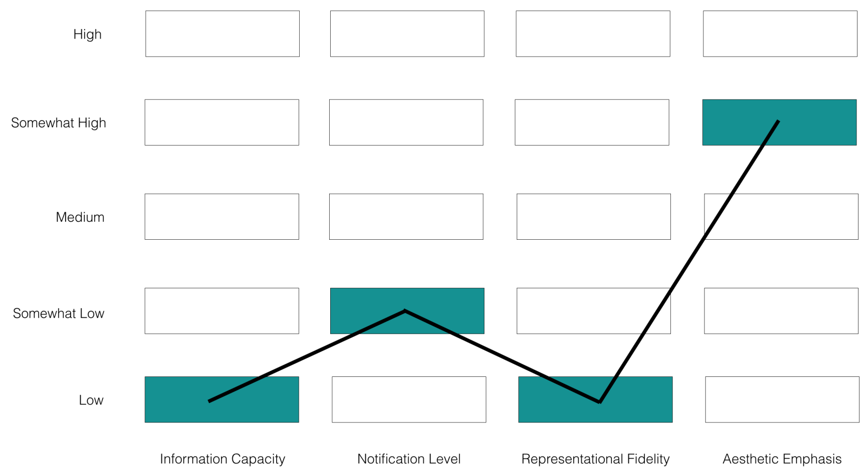


Figure 4.3: Dangling String (Weiser and Brown, 1996) represented in the taxonomy of ambient information systems (Pousman and Stasko, 2006)

example, the "Dangling String" (Weiser and Brown, 1996) extracts the network traffic data and display its strength as twisting forces. Other systems may simply degrade and rearrange the original data. For instance, the Kimura system (MacIntyre et al., 2001) increases a knowledge worker' awareness about past activities by visualizing her activities as a montage of documents in the history log. McCrickard et al. (2003) are perhaps the first researchers who formally defined a true design space for ambient systems (or notification system according to the authors' terminology). They propose a model characterizing three dimensions : **interruption**, **reaction** and **comprehension**. As the names suggest, the interruption dimension shares similar traits as in the Matthew's et al's model. The latter two dimension respective refers to user's near-term reaction and long-term comprehension. McCrickard et al. (2003) propose to discretize each dimension into the levels of `HIGH` and `LOW`, denoted by 1 and 0, so that every system can be positioned in the space, and be represented by three digits.

In debt to the aforementioned work, Pousman and Stasko (2006) then propose a four dimensional model, which is perhaps the most cited ambient system taxonomy in this field. As shown in Figure 4.3, Pousman et al.'s taxonomy has four dimensions, and we will use the Dangling String (Weiser and Brown, 1996) to explain these dimensions:

- **Information Capacity** represents the number of discrete information sources that a system can represent. For example, the Dangling String represents only network traffic information, which carries low capacity of information. On the other hand, a dashboard display, e.g. the dashboard in Mac OS, may contain stock prices, temperatures, time and dates and many others. Such displays carry high capacity of information.
- **Notification Level**. Pousman et al. adapted the five categories of notification levels from (Matthews et al., 2003), with the only change being that the *ignore* type is replaced

by *user poll*, where users have to actively check information. User poll represents the lowest notification level. For the Dangling String, it conveys information based on subtle changes in the whirling string, which can be classified to "Somewhat Low" type of notifications.

- **Representation Fidelity**. Pousman et al. use semiotics theory to categorize representation fidelity. *Indexical* refers to the representations that are close to reality, for example, photographs. *Iconic* concerns with representations with certain level of abstractions. Drawings, scribbles are seen as somewhat high fidelity, while metaphors, such as the montage of images in Kimura (MacIntyre et al., 2001) are considered as medium. *Symbolic* refers mainly to the most abstract representations. Language symbols can be viewed as somewhat low fidelity, whereas systems convey information with more abstract patterns, such as light patterns are of the lowest fidelity. The Dangling String belongs to this category.

- **Aesthetic Emphasis** deals with the importance of aesthetics of the display. As pointed out in section 4.1.1, some ambient displays emphasize that the design must be visually pleasing (Mankoff et al., 2003). The Informative Art (Redström et al., 2000) is an example as such. Even the whirling Dangling String has somewhat high aesthetic emphasis. Other systems, such as the SideShow (Cadiz et al., 2001) however focus more on delivering important notifications.

While Pousman et al.'s taxonomy underlines information presentation, Matthews et al. (2007) take another perspective, by classifying ambient systems based on the users' multiple ongoing activities. The authors claim that, from the Activity Theory's point of view, the operation (Activity Theory's terminology) sequences performed by a person may include some operations that do not service the goal of their primary action. In this regard, human activities can be classified into four classes:

- **Dormant** activities refer to those that are not serviced by any current operations performed by a user. For example, although a student wants to buy a piano, it is a dormant if the student is currently studying math course. Peripheral displays that display piano information is not helpful to the completion of the users' primary action, which is studying. However, dormant activities may be activated in non-working context. For example, public displays displaying an advertisement about a new piano model while a person is walking on the street could potentially attract his or her attention.

- **Primary** activities are those that are serviced by a user's primary action. It is where the user devotes most of her attention to.

- **Secondary** activities are those that are serviced by operations that are in the user's primary action but do not promote the attainment of the primary action's goal. For example, when a group of students are discussing a math problem, each student cares about his/her own contribution to the discussion, but the primary goal is to resolve the problem. A conversation awareness display such as the Reflect (Bachour et al., 2008) does not directly help the student to achieve the goal.

- **Pending** activities are similar to secondary activities, the main differences being that pending activities are monitored with the intent that they will become primary in the near future. Pending activities may be intentionally set aside and resumed soon. Take the scenario where a group of students trying to solve a math problem together as an example. The students may need to search for related concepts on the Web or in the book. The search activity is pending because the students may search from time to time to support their discussion.

To design ambient information systems, designers need to first identify the type of the activities that the system is expected to support with the above taxonomy, then turn to Pousman et al.'s taxonomy to determine the presentation level of each of the 4 dimensions.

Evaluating Ambient Information Systems

Ambient information systems are typically designed for occasional, non-primary and sometimes opportunistic use. This characteristic determines that the evaluation of such systems emphasizes more on the qualitative, rather than quantitative measures. Heuristic evaluation (Nielsen and Molich, 1990) is a popular usability evaluation method for user interfaces, which requires the evaluators to examine whether a design complies with a set of usability criteria (the heuristics). Mankoff et al. (2003) propose that ambient information systems can also be evaluated by this method, with the following heuristics: (1) sufficient information design (2) consistent and intuitive mapping (minimal cognitive load) (3) visibility of state (4) aesthetic and pleasing design (5) useful and relevant information (6) easy transition to more in-depth information (7) peripherality of display (unobtrusiveness) (8) match between system and real world (9) visibility of system status (10) user control and freedom (11) error prevention (12) flexibility and efficiency of use. The first 7 heuristics are derived from user surveys, whereas the last 5 are quoted from Nielsen's heuristics (Nielsen, 2005). In a similar regard, Vogel and Balakrishnan (2004) proposed that public ambient displays must following the principles of (1) calm aesthetics (2) comprehension (3) notification (4) short-duration fluid interaction (5) immediate usability (6) shared use (7) combining public and personal information, and (8) privacy. These principles are much identical to Mankoff et al.'s heuristics. They offer lots of "Do"s for designing the display. Nevertheless, design and evaluation are two sides of the same icon, so evaluations can be done by treating the principles as heuristics.

As discussed previously, ambient information systems encompass a group of systems that service in different situations, so one prominent problem of evaluating ambient information systems with heuristics lies in its lack of context, i.e. the ongoing activity of the user when interacting with the system. This is where the Activity Theory may come into play. Grounded in the Activity Theory, Matthews et al. (2007) specified five criteria to evaluate ambient information systems across various contexts, including appeal, learnability, awareness, effects of breakdowns, and distraction. It is worth mentioning that these dimensions are criteria, rather than specific metrics. In other words, the criteria do not specify, e.g. how much learnability an ambient system should offer. Instead, they emphasize that these dimensions cannot be analyzed without taking user's activities into consideration. Therefore, Matthews et al.'s evaluation criteria go beyond simple statements such as "awareness is important for ambient information system", by suggesting the use of the Activity Theory as an analytical framework to understand what kind of awareness should be provided by the system and how.

4.1.4 Ambient Information Systems in Learning Setup

Most ambient information systems were developed for public (Alt et al., 2012) or daily work use (Röcker, 2009), manifesting their great value in augmenting these everyday situations. We are interested in exploring the design of ambient information systems in learning contexts, which can be exemplified by the learning scenarios presented in Section 1.1 : seminar talks, group learning as well as online learning. Apart from their primary activities, participants in these scenarios are also involved in several secondary or pending activities, which are possibly supported by ambient information systems. In this section we review a few systems servicing learning or collaborative learning activities in classroom and in groups.

Ambient Technologies in Classroom

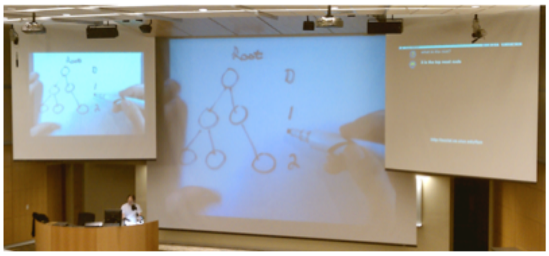
Classroom learning activities are usually characterized by one-to-many relationships between tutors and learners. As primary activity, tutors deliver knowledge, most commonly in the form of presenting in front of students. They may also perform secondary activities such as managing time and pace, monitoring students' affections, responding to students' questions, and orchestrating students' activities etc. On the other hand, the students' primary activity is learning, which may involves comprehending lectures and exercising. They may need to, for example, be aware of other students' progresses etc.

Lantern (Alavi and Dillenbourg, 2012) is designed to support exercise sessions when students need help from tutors. A light object is placed besides each exercise groups (Figure 4.4(c)), and students can use it to call for help. Different colors indicate which exercise the students are working on. The tutor can then make a judgment on which group may need help most based on their progress. Fireflies (Bakker et al., 2013) is also designed as light object similar to Lantern. Each student in a classroom has a Fireflies placed on his/her desk (Figure 4.4(a)), and the tutor can use a centralized control to manage activities by signaling different colors. The Subtle Stone (Balaam et al., 2010) is yet another tangible light-emitting artifact. It is designed for students to express their emotions by changing the color(Figure 4.4(f)). To increase the awareness of the students of others' working progress, Lamberty et al. (2011) proposed a solution to use a public display to show the ongoing work of all the students (Figure 4.4(b)). When lecturing at platform, a tutor may want to get immediate feedback from the students,

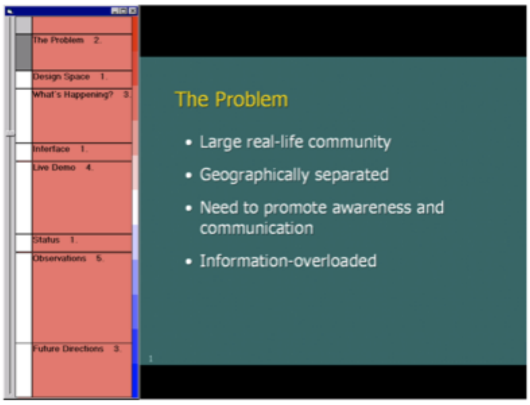
(a) Fireflies (Bakker et. al., 2013)

(b) Peripheral Display for Classroom (Lamberty et. al., 2011)

(c) Lantern (Alavi and Dillenbourg., 2010)



(d) Fragmental Social Mirror (Bergstrom et. al., 2010)



(e) Time Aura (Mamykina et. al., 2001)

(f) Subtle Stone (Balaam et. al 2010)

Figure 4.4: Ambient information systems in classroom

e.g. questions or concerns about the lecture. The Fragmental Social Mirror (Bergstrom et al., 2011) offers a solution by publicly displaying anonymous messages entered by the students during the class. The tutor can then adjust the lecturing based on these feedbacks (Figure 4.4(d)). Time Aura (Mamykina et al., 2001), though originally designed to help presenters to adjust their pacing, can also be applied in the classroom to help tutors manage their pacing (Figure 4.4(e)).

Ambient Technologies in Collaborative Group

Participants in collaborative groups often have to keep track of time and balance social interactions, indicating that ambient information systems can be designed to promote awareness of these aspects. For the awareness of time, Occhialini et al. (2011) designed a halogen spots

(a) Reflect (Bachour et. al., 2008)

(b) Relational Cockpit (Sturm and Terken., 2009)

(c) WeSearch (Morris et. al., 2010)

(d) Conversational Cluster Table (Bergstrom and Karahalios., 2009)

(e) Halogen spots for time awareness (Occhialini et. al., 2011)

Figure 4.5: Ambient Information Systems in Collaborative Setup

prototype which exploits the color, intensity and direction of light beams to reflect meeting status as well as to notify about the elapsed time (Figure 4.5(e)).

As for social interaction awareness, conversation is the most utilized feature. The Reflect (Bachour et al., 2008) visualizes each participant's amount of speech as territories of color LEDs in front of him/her(Figure 4.5(a)). The more one talks, the bigger one's territory grows. The imparity of territories is found to regulate the group discussion in a way that promotes balanced participation. The Relational Cockpit (Sturm and Terken, 2009) also utilizes conversational speech. It projects three circles of social dynamics in front of each participants. As shown in (Figure 4.5(b)), the blue, green and orange circles respectively visualizes the cumulative speaking time, the cumulative attention the person received as listener and speaker. (Bergstrom and Karahalios, 2009b) designed a conversational cluster table, which captures important words during group conversation and automatically constructs topical clusters based on the words' semantic similarity (Figure 4.5(d)). The display thus promotes awareness of the discussion content among the participants.

Speech is not the only exploited collaborative feature for group awareness. Morris et al. (2010) designed WeSearch, a tabletop system used for collaborative search scenarios. The WeSearch interface displays to each searcher a marquee region, where text and images that reflect other participants search activities slide slowly from one end to another (Figure 4.5(c)). The marquee region aims to increase the awareness of other searchers' activity, so as to better support division of labor among group members, or foster collaborations.

4.2 Interaction with Serendipitous Information

Toms (2000) classified the ways people acquire information into 3 categories: (1) seek information about a well-defined and known object(s) (2) seek information that can not be fully articulated, but will be recognized once seen. (3) accidental discovery of useful information. The third category is what we call serendipity. The earliest appearance of this word "serendipity" was found in a letter written by an English historian Horace Walpole in 1754. Walpole illustrated the concept with a story in a fairy tale, which was about three princes, who were "making discoveries, by accidents and sagacity, of things which they were not in quest of" (Remer, 1965). Van Andel (1994) simply call it "the art of making an unsought finding". Note that serendipity was referred to as the accidental discovery of things of many kinds. In fact, the history of science is full of serendipities, with the discovery of Penicillin by Fleming being a notable example.

In information science literature, Erdelez (2004) described serendipity as *information encountering*, referring to a particular phenomenon where an information seeker looks for information on one topic, but accidentally encounters something interesting in another topic. Information encountering is just a typical type of serendipitous information behaviors found in literature. Similar to Bates' taxonomy of information behaviors (2003), André et al. (2009) summarized some of the related research and fit them in a two-dimensional taxonomy.

As shown in Figure 4.6, most of the serendipitous information behaviors found in literature involve finding information irrelevant to the goal of initial activity. *Information encountering* is classified as encountering information that is irrelevant to the goal of a directed browsing. and this type of serendipity often occurs when a user browses information while searching information. *Serendipitous information retrieval* (Toms, 2000) and *opportunistic browsing* (De Bruijn and Spence, 2008) describe phenomena where people are intentionally browsing information without being aware of any goals. An example is when an individual stops by

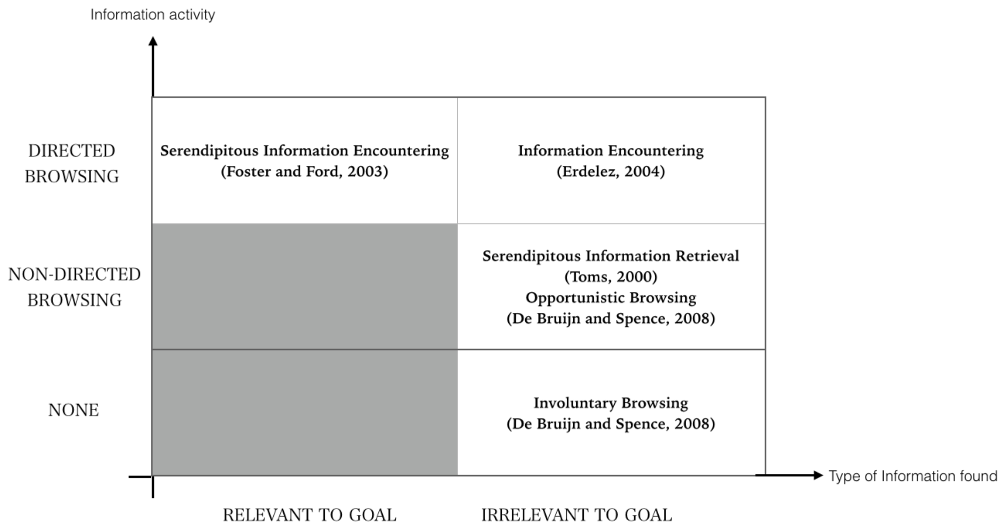


Figure 4.6: Taxonomy of serendipitous information behaviors (adapted from (André et al., 2009))

a public display in the street and checks out what is in there. De Bruijn and Spence (2008) termed *involuntary browsing* to refer to a similar situation with the only difference being that the browsing is unintentional. An example is when a person's eye gaze randomly jumps over a serious of fixations and serendipitously fixates on a piece of information that may answer a long-standing question in his/her mind.

Serendipity can also be relevant to the goal of an initial activity, but research literature reveals this kind of serendipity only occurs during directed browsing activities, which is covered by the concept of *serendipitous information encountering* (Foster and Ford, 2003). In an empirical study on serendipitous behaviors, the authors found examples where (1) the existence and location rather than the value of encountered information were unexpected. In other words, a person intentionally looks for and browses information with a clear goal, and finds the information in an unexpected location. (2) the encountered information has also unexpected value not only by chance, but also by looking in "likely resources". This means a person intentionally attempts to look for something in a revenue where the required information may potentially resides. S/he does it with little expectation but ends it up with an "unexpected" finding.

Given various types of serendipitous information behaviors as discussed previously, (André et al., 2009) propose a definition that only focus on the value the encountered information proves to the person. They define serendipity as (1) the finding of unexpected of information (regardless of its relevance to the goal) while engaged in any information activity (2) the making of an intellectual leap of understanding with that information to arrive at an insight. In the last few decades, many researchers have attempted to understand, model and even design for serendipity, and we review a selection of these research attempts in this section.

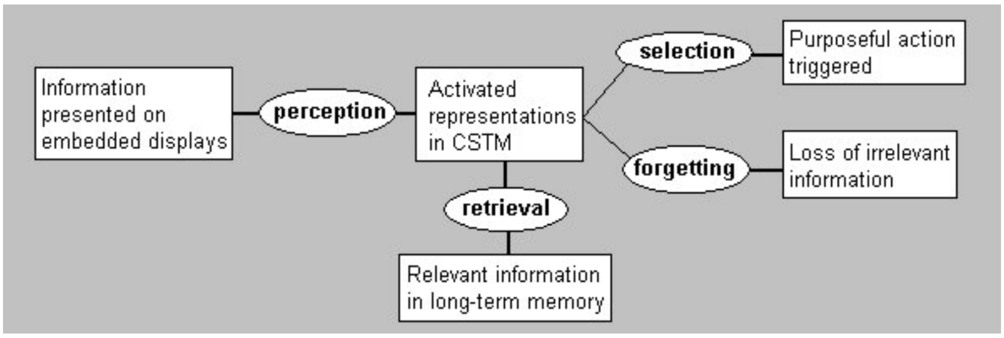


Figure 4.7: Cognitive model of opportunistic browsing (De Bruijn and Spence, 2001)

4.2.1 Models of Serendipity

A variety of models have been proposed in literature to model serendipity. Some of the models are derived from existing theories that are originally targeted for modeling more generalized human computer interaction behaviors. For example, De Bruijn and Spence (2008) modeled serendipity with Norman's action cycles (Norman, 1988), which describe human activities as a sequence of actions of formulating an intention, planning an action, executing the action, perceiving the change, interpreting the effect and finally evaluating the results. The former three actions form the so called *gulf of execution* and the latter constitutes the *gulf of evaluation*. De Bruijn and Spence (2008) claim that the process of serendipitous acquisition of information starts with the gulf of execution without the need to traverse the gulf of execution. In a similar vein, Blandford and Attfield (2010), the two researchers who proposed the information journey model discussed in section 3.1.1, argue that serendipitously encountered information may cause an information seeker to develop a new aganda, leading to new things to find and interpret.

Several models have also been developed exclusively for modeling serendipity. Some focus on the cognitive mechanisms, while others underscore high-level abstractions of serendipitous experiences .

Cognitive Model of Serendipity

A cognitive model of opportunistic browsing was proposed by De Bruijn and Spence (2001). As depicted in Figure 4.7, opportunistic browsing starts from perceiving encountered information from visual or auditory channel. Meaningful content is rapidly extracted and represented in the Conceptual Short Term Memory (CSTM), which further retrieves relevant information in one's long-term memory (LTM). Once the retrieved information is found to be associated with some long-standing problems in mind or an early failure during finding similar information, the person then draws attention to it as well as imposes further actions. Irrelevant information is immediately dropped out. The cognitive process is similar to the one we presented for selective attention and priming in Section 2.1.

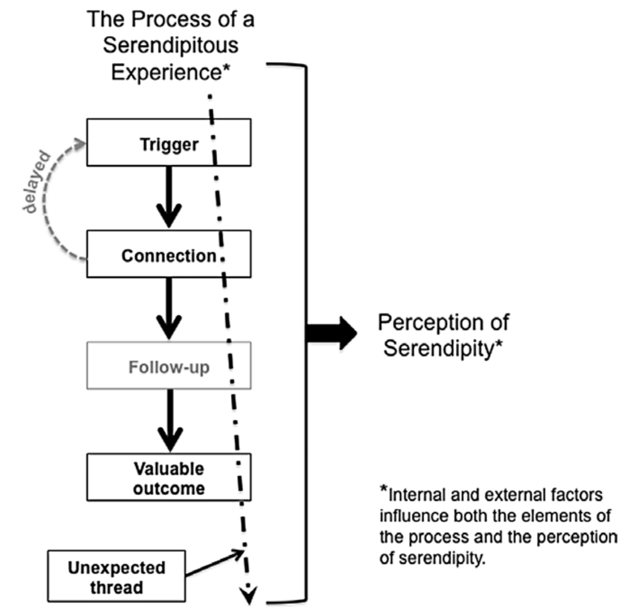


Figure 4.8: Process model of serendipitous experiences (McCay-Peet and Toms, 2015)

Process Model of Serendipity

While the cognitive model captures man's low-level cognitive behaviors such as memory retrieval and attention selection, the sensemaking model emphasizes the key elements involved for the process of serendipity. Several researchers have worked on models as such, including Cunha (2005); Makri and Blandford (2012); McCay-Peet and Toms (2010); Rubin et al. (2011); Sun et al. (2011). Based on these work, McCay-Peet and Toms (2015) summarized and extracted the main components that constitute serendipitous experiences as follows:

- **Trigger** refers to stimulus (e.g. visual, textual, verbal cues) that spark a person's serendipitous experiences. This element was called *Noticing* in the previous models (Cunha, 2005; McCay-Peet and Toms, 2010; Sun et al., 2011)

- **Deplay** refers to the situation when a person perceives the stimulus, but does not immediately associate it with other information or experiences in memory. In other words, there might be an incubation period before connections are made.

- **Connection** means the association between the trigger and one's past experiences, knowledge or problems.

- **Follow-up** refers to the *post connection* in the model proposed by Makri and Blandford (2012). Compared to immediate associations made in the *connection*, follow-up considers the iterative process of "exploiting" the connection to obtain a valuable outcome, e.g. a person may verify the connections s/he has previously built.

- **Valuable outcome** is the positive effect of the serendipitous experience, i.e. after recognizing and exploit the connection, valuable outcome deals with the outcome of the serendipitous finding.

- **Unexpected thread** does not exist on its own, according to the authors. Rather, it is the unexpectedness woven throughout all the prior elements.

- **Perception of serendipity** means an individual is aware of all the prior elements and consider his/her experience as serendipitous.

The above processes are visually depicted in Figure 4.8. *Delay* and *Follow-up* are shown in gray because these elements do not necessarily happen for the perception of serendipity.

4.2.2 Factors influencing Serendipity

McCay-Peet and Toms (2015) argued that certain internal and external factors may influence the elements in the serendipitous process model as well as the perception of serendipity. In a similar vein, Thudt et al. (2012) identified a set of personal traits and environmental factors in literature that may influence serendipity, corresponding to McCay-Peet and Toms' internal and external factors respectively.

Personal Traits

Erdelez (1997) found in her research that information seekers, based on how often they encounter information accidentally, can be classified into four categories: (1) super-encounterers (2) encounterers (c) occasional encounterers and (4) non-encounterers. Those super-encounterers are reported to have not only often experienced information encountering, but also considered it as an integral element of their information behaviors. This clearly shows certain personal characteristics may influence serendipity, and Thudt et al. (2012) summarized four kinds of them : (1) **observational skills** (2) **open-mindedness** (3) **knowledge** (4) **perseverance**. Open-mindedness is also seen as curiosity, or the enthusiasm courage to face challenge, the deliberate seek of information from different perspectives (Liestman, 1992). Knowledge is the concrete manifestation of what Walpole referred to as "sagacity" in his original description of serendipity in the 17th century (Remer, 1965). As the French chemist said, "chance favors the prepared mind". In fact, both open-mindedness and knowledge can be considered as the "prepared mind", which is seen as the strongest personal trait associated with serendipity (McCay-Peet and Toms, 2015).

Environmental Factors

Thudt et al. (2012) identified two environmental factors that favor serendipity, **coincidence** and **influence of people and system**. The former underlines that serendipity is strongly related to accidental, unexpected, unsought or coincidental events that are unpredictable. It assumes that such "bonne chance" has less to do with luck than simple randomness (Liestman, 1992) and could possibly be supported by introducing such randomness in the presented information.

Influence of people and system manifests itself as how information is organized and presented, by people or by systems prior to the serendipitous discovery. This point is termed as *prevenient grace* according to Liestman (1992). For example, a person may look for two totally unrelated books in a library, but she finds the second book on her way to locate the first book. This scenario sounds a matter of luck, but by taking a closer look we may find it attribute to the organization and presentation of books on shelfs. Since both books were written by authors with "T" as their initials, they were placed close to each other. Admittedly, libraries nowadays rarely organize books in this way, but online information systems may present book information in various ways to increase serendipity. The key point is that systematic organization of materials may also lead to serendipity, meaning that serendipity is sometimes more than chances of pure coincidence.

4.2.3 Design Considerations

As serendipity is usually coincidental, Van Andel (1994) negatively reflects on the potential employment of computers to program serendipity:

"Like all intuitive operating, pure serendipity is not amenable to generation by a computer. The very moment I can plan or programme 'serendipity' it cannot be called serendipity anymore".

Statements as such view serendipity as mysterious and seemingly unpredictable experiences. However, more and more researchers tend to believe, though it is impossible to design serendipity, it is practically possible to design "for" serendipity (Campos and Figueiredo, 2001). As André et al. (2009) says, it may be possible "for a computer searching for patterns of association or of related interest to be able to surface something that to its user would be perceived as a serendipitous discovery". In other words, computers have the potential to create opportunities (Makri et al., 2014), to facilitate, or to induce serendipity.

Supporting the Process of Serendipity

The full process of serendipity as discussed in section 4.2.1 can be translated into three key elements (Maxwell et al., 2012) : (1) **making connections** (2) **exploiting the value of connections**, and (3) **reflecting on the value of the outcome**. It is suggested that a serendipity-inducing system can be designed to consider these three elements.

Sun et al. (2011) made suggestions that emphasize the facilitation of making connections. They claim that technologies should consider creating: (1) a resource-rich environment where people are exposed to multiple influences (e.g. visual stimuli) (2) an information environment which contains resources from outside people's habitual data, information or search domain where new ideas can be stimulated (3) a relaxing environment where people are not actively focusing on one thing but where they are open to exploring the things around them, and (4) an environment where people's minds are open and they are used to making many connections between information and their knowledge and experience. In a similar vein, McCay-Peet et al. (2015) propose five facets of a digital environment to create opportunities for serendipity: (1) **Enable exploration** and examination of information, ideas and resources (2) Contain **Trigger-rich** information that sparks users' interest. They adopt the same notion of "trigger" as in (McCay-Peet and Toms, 2015) discussed in section 4.2.2 (3) **Highlights triggers**, which means the triggers need to be delivered in a way that captures user's attention, (4) **Enable connections**, perhaps through visualization tools (Thudt et al., 2012) that connect interesting ideas and information.

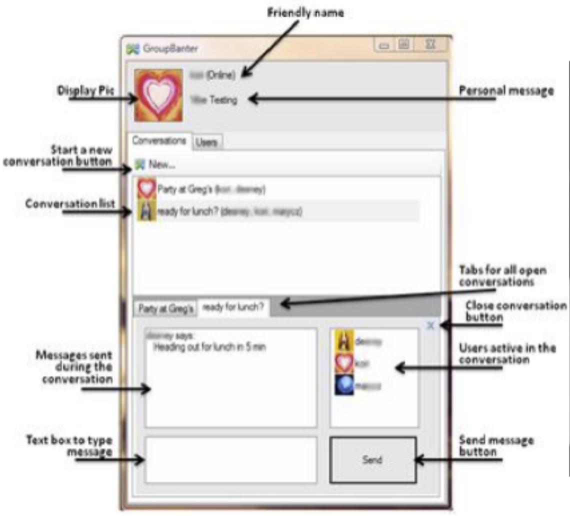
The design considerations from Sun et al. (2011) and McCay-Peet et al. (2015) both focus on designing triggers to facilitate the connection process of serendipity. Makri et al. (2014) further suggest that exploiting and reflecting the value of connections can be supported by (1) **drawing on previous experiences** (2) **looking for patterns**, and (3) **seizing opportunities**. The first point is concerned with remembering users' prior experiences which allow them to make more sense of a new situation. Point two states that people often look for "patterns" in the information space so as to project the value of connections. Therefore, digital systems should highlight or even visualize the connection between the presented information, if they are related, e.g. semantically. The last point underscores the importance of seizing opportunities in order to exploit the value of connections made before. A serendipity-inducing system should assist users to follow up on potentially valuable opportunities. It should also offer an integrated experience of spotting and using the connection.

Considering the Influencing Factors of Serendipity

Researchers have also pointed out the importance of considering the influencing factors of serendipity, i.e. personal traits and environmental factors as discussed in section 4.2.2.

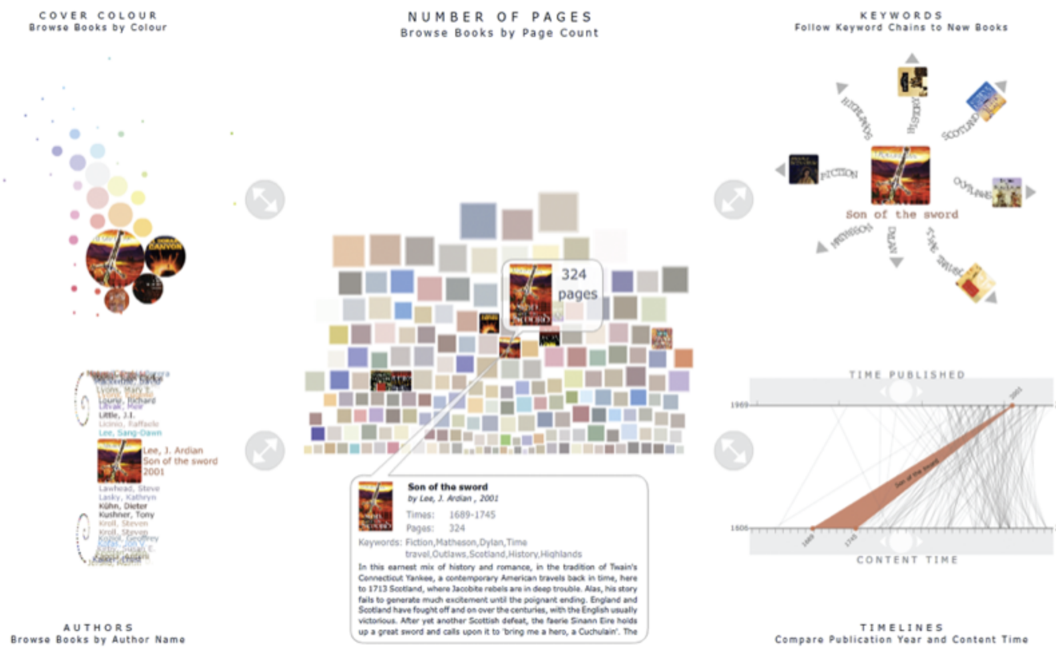
Erdelez (1999) identifies four elements that need to be considered for delivering an information-encountering experience: (1) the **information user** who encounters the information (2) the **environment** where the information encountering occurred (3) the characteristics of the **information encountered**, and (4) the characteristics of the **information needs** that the information encountering addresses. The last two elements imply that digital environments should consider that information encountered can be either problem-related or interest-related, and attempt to understand a user's past, current or future information needs.

André et al. (2009) propose that computers have the potential to (1) deliver **better support chance encounters**. Computers should attempt to present serendipitous content at the appropriate time. This can achieved by e.g. personalized information. They should also support creativity and play through the introduction of random and redundant information (2)



(a) GroupBanter (Inkpen et. al., 2009)

(b) The coffee table (De Bruijn and Spence, 2008)



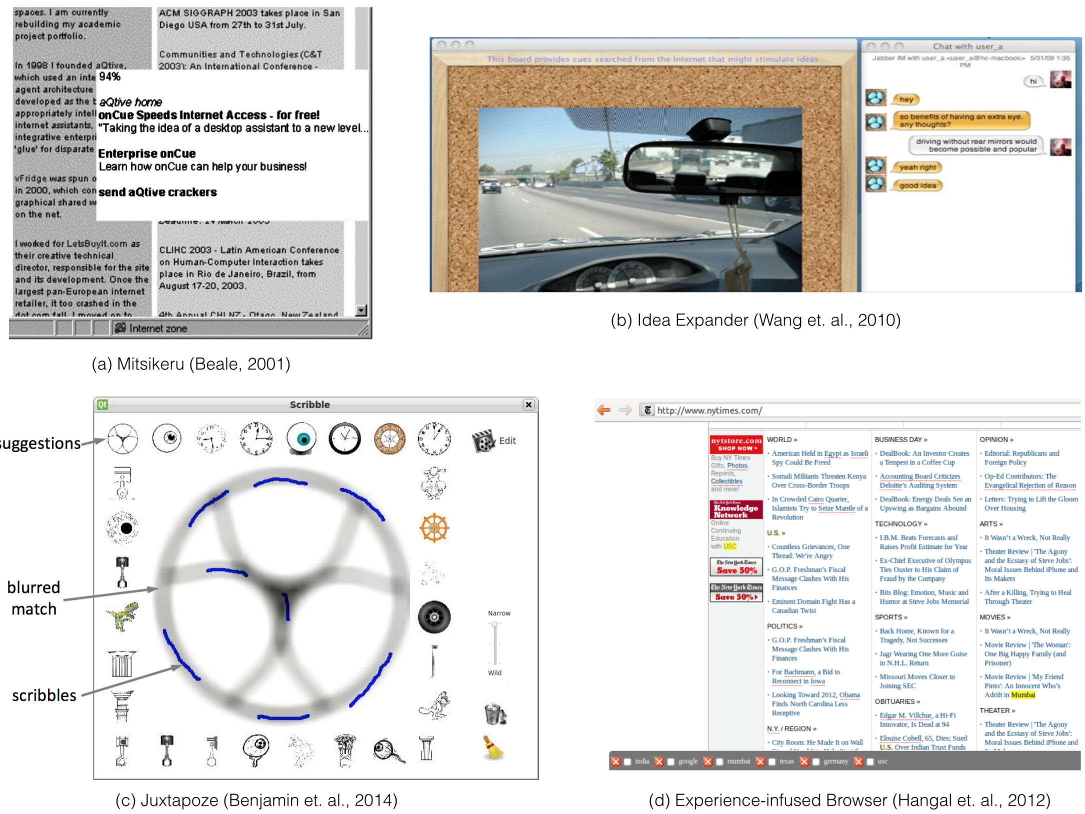
(c) The Bohemian Bookshelf (Thudt et. al., 2012)

Figure 4.9: Systems support coincidental or system-induced serendipity

enhance sagacity. For example, a system should keep track of an individual's existing domain knowledge, and present relevant information (3) build **networks to help serendipity flourish**. This aspect underscores the importance of building community for sharing information with other people who perhaps hold more expert knowledge or different perspectives for recognizing the serendipitous value.

4.2.4 Example Serendipity-inducing Systems

Several researchers have attempted to understand whether pure coincidence and system-induced serendipity can be supported by computers. The coffee table developed by De Bruijn and Spence (2001, 2008) randomly displays information items, which move slowly around (Figure 4.9(b)). In a study where participants were primed with a "national flag identification" task before they gathered together to discuss another topic, they were all found to have occasionally interacted with the table and several participants learned the name of the flag afterwards. GroupBanter (Inkpen et al., 2009) is a group-based instant messenger that allows users to publicize conversations as an implicit invitation for others to join the conversation (Figure 4.9(a)). The authors found that GroupBanter offers benefits of awareness of ongoing conversations and serendipitous conversation compared to other instant messengers. The Bohemian Bookshelf (Thudt et al., 2012) implemented several interlinked visualizations of book collection, offering multiple access points for users to query books from different perspectives, such as book cover colors, publication years, keyword chains etc (Figure 4.9(c)). It is an exploration of how to promote serendipity through information visualization.



(a) Mitsikeru (Beale, 2001)

(b) Idea Expander (Wang et. al., 2010)

(c) Juxtapoze (Benjamin et. al., 2014)

(d) Experience-infused Browser (Hangal et. al., 2012)

Figure 4.10: Systems support serendipity with context

Most serendipity-inducing systems found in literature take into account users' context, including their previous experiences and current activities. Such systems use triggers (mostly in terms of visual stimuli) to spark connections. The Mitsikeru (Campos and Figueiredo, 2001) is an agent-based Web interface to capture and model users' behavior, and then suggest relevant Web pages that will be potentially interested by users in an ambient way. To be specific, the Mitsikeru learns the users' current context and determines the relevance of future pages based on their history of interaction. When a user hovers the mouse over a hypertext, a summary of its hyper-linked page as well as a score that indicates its potential relevance to the current task pops out (Figure 4.10(a)). The goal is to augment users browsing habits in order to help them surf the Internet more effectively. Similarly, the Experience-infused Browser (Hangal et al., 2012) indexes a user's digital history from email and chat archives. When the user views a Web page, the words that match the archives are highlighted in the browser (Figure 4.10(d)). Users found highlighting words like names, products, organizations and places useful. Juxtapoze (Benjamin et al., 2014) is a clipart workflow software supporting serendipitous discoveries and creative expression. When a novice user scribbles in the canvas, the scribble will be matched with a variety of existing illustrations that have similar visual shapes (Figure 4.10(c)), allowing non-artist to easily create interesting artifacts. Idea Expander (Wang et al., 2010) is a tool to facilitate online group brainstorming with pictorial inspirations based on chatting conversation. The system captures conversational words in the text written by group members, and suggest relevant pictures to spark creativity (Figure 4.10(b)). Research has found groups with Idea Expander generated more ideas than those who work without it.

5 Research Framework and Perspectives

The types of learning activities do not only influence what information the participants may need, but also the way they get access to it. Learning context presents the need of managing attention and cognitive resources between information seeking and the learning activity. Specifically for collaborative learning scenarios, factors such as social context, restricted time of information access, or available learning materials all contribute to shaping the information seeking process. Before we discuss how to design concrete systems, it helps to understand the situations where students need information, the associated challenges as well as the kinds of needs that this thesis will address. This chapter starts with the presentation of a survey study aimed at identifying common types of information needs, followed by a derived research framework that guides all the follow-up research projects in the upcoming chapters.

5.1 Gathering Information Needs : A Survey Study

Students often need information while performing learning activities. This is the biggest assumption that motivates us to design technologies for supporting their information needs. As illustrated in the scenarios at the opening of this dissertation, sometimes information needs arise from the content of a presentation, other times they are prompted during the conversation with other students, through the exchange of knowledge. Students have the option to turn to mobile or desktop devices or discuss with a *more knowledgeable other* if condition permits. Chapter 2 reveals a few factors that influence students' information or help seeking behaviors, e.g. some students follow the *principle of least effort* and tend to choose easily accessible means (Liu and Yang, 2004). In terms of asking help from others, they may suffer from a few psychological barriers as listed in Section 3.2.1. Considering information seeking is usually not the primary task in an learning activity, we assume that students' information or help seeking strategy also depends on the types and importance of desired information, the availability of tools and time. Therefore, simply providing students with relevant tools and Internet access and let them search as in everyday information searching activities may neither be effective nor efficient. In this section, we report a survey study aiming at exploring the types of information needs as well as the factors that impede people from actively searching information during learning activities.

5.1.1 Research Methodology

The whole Chapter 3 is devoted to review information and help seeking behaviors rather than the user's *information need*. Before proceeding to the study of information need, it helps to understand its meaning as well as its constitutional parts. Later we will employ a survey study to understand information needs especially in learning context.

Anatomy of Information Needs

Information need is a concept that has been studied for nearly a hundred years, resulting in definitions from various perspectives (Case, 2012). Wilson (1981) claims that information needs are "qualitatively similar" to human needs, which are usually classified into three categories: (1) physiological needs, which are the basic needs for life, e.g. food and water (2) affective needs, which are the needs related to emotions such as the need for achievement, self-actualization etc. and (3) cognitive needs such as the need to plan or to learn something. These needs are interrelated: for example, physiological needs may trigger other kinds of needs, which may in turn induce cognitive needs. As Wilson (1981) pointed out, the performance of tasks and the processes of planning and decision-making are the major generators of cognitive needs. In fact, the information needs we are discussing in this dissertation mainly refer to cognitive needs.

Derr (1983) defines information needs as conditions in which "certain information contributes to the achievement of an information purpose". Two conditions are made explicit in this definition : (1) the presence of an information purpose, and (2) the information itself that contributes to achieving the purpose. Derr's two conditions of information needs have been echoed by Tate and Russell-Rose (2012), who classify information needs into two dimensions: *search motive* and *search type*. Tate and Russel-Rose further explain that search motive corresponds to Marchionini's taxonomy of exploratory search activities (2006), which is reviewed in Section 3.1.4. In addition to the three original types (*lookup*, *learn* and *investigate*), Tate and Russel-Rose suggest to add another type *casual* to represent undirected activities that are irrelevant to the goal of completing a task, e.g. for killing time or for fun. The search type dimension is the "genre of the information being sought", which can be further categorized into four types : informational, geographic, personal information management and transactional. The first two types are self-explanatory. For the other two, personal information management mainly deals with private information such as checking schedule or making plans, and transactional needs focus on actions rather than textual information about a topic, such as price comparison, product monitoring etc.

Information Need Survey: Participants and Procedure

The global goal of this dissertation is to design a computer system that provides timely and relevant information to people engaged in learning. It is necessary to clarify that we focus on satisfying the *informational* type of needs rather than other kinds (Tate and Russell-Rose's terminology), because informational needs are more meaningful for learning. Several scenarios are presented in Section 1.1 to illustrate what we mean by learning in this dissertation, including seminar talks, collaborative brainstorming, collaborative lecture viewing and online learning. People definitely need various information while performing different learning activities, but these needs should be framed in the search motive dimension of Tate and Russel-Rose's framework. We do not intend to study the information needs in every learning scenario. Instead, we start with the most accessible and convenient scenario to investigate, i.e. seminar talks, because weekly seminar presentations take place at our lab. Information behaviors can be observed but information needs are hidden behind the scene. To examine these latent needs, researchers usually employ methodologies that are based on self-reporting, such as diary (Wild et al., 2010; Elsweiler et al., 2010; Sohn et al., 2008; Wilson et al., 2012) and survey (Case, 2012; Khan and Shafique, 2011). We adopted the survey approach and asked the seminar audience to recall their information needs after the talk.

The survey study was conducted in 6 seminar talks that took place at our lab. Not all the talks were given by internal lab members, some of them were presented by invited speakers. Each talk attracted 10-15 lab members, and some of the participants have attended multiple talks. For each internal seminar, we asked the presenter to distribute the slides as printouts right before they started presenting, but this was unfortunately not practical for talks given by invited speakers. After the presentation, the attendees were asked to fill in a survey (cf. Appendix A.1) voluntarily. Our interested questions include:

- Have you had any information needs? If yes, please specify.
- Where is the need from, slides or oral presentation?
- Why did you have the information needs?
- Did you search to address the need? Why or why not?
- Does the need still exist after the presentation?

To answer the first question, the subjects were asked to list their information needs in terms of words or phrases. The last four questions were repeated for each item of the specified information needs. The audience could use the presentation printout, if available, as assistance to recall their information needs arose during the talk. The topics, the number of participants, the number of surveys collected, as well as the number of tools at hand during the presentation are presented in Table 5.1.

Table 5.1: Information about the six seminar talks

Seminar	Topic	Presenter	Participant	Survey	Laptop	Mobile	Notebook
1	Collaborative learning	External	10	7	1	3	2
2	Annotation and gaze pattern	Internal	12	11	0	1	2
3	Eye tracking	Internal	13	12	3	2	5
4	Tangible interfaces & structural mechanics	Internal	13	13	3	1	5
5	Tangible interfaces & learning	Internal	11	10	2	4	8
6	Social robot	External	15	10	2	3	0

Usually attendees bring their laptops, mobile phones or notebooks to seminars. Iqbal et al. (2011) studied the usage of these peripheral computing devices in a similar seminar setup, but in a larger lecture hall and with more attendees. They found that most people reported to have used computing devices for generating unrelated content, such as coding, drawing diagrams, editing text or communicating with others. Few people reported they solely took notes or looked up references with the devices. In our study, we do not focus on what the audience do with their tools. Instead, we consider laptops and mobiles as common technological means that offer search possibilities to address certain information needs. It is clear from Table 5.1 that many attendees did not have devices to make search. This is also part of the current seminar practices. The goal of the study is two-fold, and we want to:

1. examine how informational needs manifest themselves in term of search motives in a concrete learning scenario, i.e.research seminars.
2. understand how people deal with their information needs in current seminar practices.

5.1.2 Survey Results

We collected 63 surveys in total from the 6 seminars. In 41 of them, participants reported 64 information needs, of which 17.5% arose from the oral presentation, 54.9% from the slides and the rest 28.6% from both sources(one information need was not counted because the subject forgot where it came from, so the percentages are computed out of a total of 63 needs). On average 65.9% of the attendees (σ = 25.3%) had ($\mu = 1.6, \sigma = .4$) information needs per seminar. In the following sections, we characterize the categories of information needs based on the motives they specified in the survey. Then, we also discuss how the subjects dealt with their needs and why.

Categorizing Information Needs

From the descriptions of the 64 information needs, we find all of them are informational needs (Tates and Russel-Rose), i.e. the subjects need information about a topic. Additionally, we discriminate *internal* from *external* needs. Internal needs are actually doubts raised from the presented research itself or about the terms invented by the presenter. Such doubts can only be clarified by the presenter. Conversely, external needs are those that can be satisfied by searching information elsewhere, e.g. on the Internet. External informational needs can easily find themselves in the four generic types of search motives: *casual*, *lookup*, *learn* and *investigate* (Tate and Russell-Rose, 2012). We summarize their answers about search motive and we identify 9 concrete modes within the generic motive types, as shown in Table 5.2. Three attendees with information needs did not specify the reason, so the analysis is based on 61 reported search motives.

Table 5.2: Motives and modes of information needs observed in seminar talks

	Motive	Mode	Explanation	Occurence
External	Casual	Curiosity	curious or interesting about an item e.g. "curious about gel robot" (Seminar 6)	13
		Serendipity	pleasant surprise for another goal in mind e.g. "this may be useful for my own research" (Seminar 6)	2
	Lookup	Location	want to obtain an item encountered or to retrieve specific facts e.g. "get the pdf of this paper" (Seminar 4)	3
		Recall	recall the details of known items e.g. "I have heard that name in a similar context" (Seminar 1)	5
	Learn	Knowing	want to find the meaning of an item that is new to them e.g. "did not know the notion" (Seminar 1)	25
		Elaboration	need more information than a definition e.g. "... more information on the psychological terms ..." (Seminar 3)	4
		Exploration	explore a set of data for knowledge discovery e.g. "I want to see more examples ..." (Seminar 5)	1
		Comparision	want to compare one item against another e.g. "I was confused between two terms ..." (Seminar 5)	1
	Investigate	Evaluation	judge the correctness or appropriateness of a statement e.g. "... there was a debate if the system is really TUI ...." (Seminar 4)	3
Internal		Doubt	doubts in the presented research itself e.g. "... not clear what the presenter took into account" (Seminar 4)	4

The namings of our identified modes are mostly in debt to the search modes observed by Russell-Rose et al. (2011) in enterprise search scenarios (cf. Section 3.1.4), with additions of *recall* and *elaboration* and removals of the unobserved "monitor", "analyze" and "synthesize".

The 31 ***learn*** search motives account for half (50.8%) of the information needs reported, where we observed four modes. The *knowing* and *elaboration* modes both correspond to the "comprehend" mode in Russell-Rose et al. (2011), but with varying degrees. The *knowing* mode refers to the situation when a person looks for explanations of a term that is new to him/her. *Elaboration* is similar, but underlines that the required information is more advanced than basic knowledge. As in the example given in Table 5.2, the attendee was not satisfied by the explanation of some psychological terms given by the presenter, and wanted to *elaborate* the notion with "more information". In addition, we observed 1 *exploration* mode and 1 *comparison* modes, which correspond to "explore" and "compare" in Russell-Rose's taxonomy. Among the four presented modes, the *knowing* information needs exhibit the dominance, because most of the participants attended the seminar simply to get exposure to new things, corresponding to the lower layers of Bloom's taxonomy (cf. Section 2.2.2). They mostly do not have the motive to dig into specific problems.

The second largest group of search motives are ***casual***, which occurred 15 times, accounting for 24.6% of the information needs. At first sight, it might be surprising to see many information needs were attributed to *curiosity*. Considering most participants simply wanted to learn about new things, it is probable that the talks sparked their interests in certain aspects. A more interesting finding is that 2 needs arose with *serendipity*. The attendees reported certain messages delivered in the talk triggered connections to their own work, so that they would like to "exploit the value of connections" (Maxwell et al., 2012). This demonstrates seminar talks may induce serendipity. Such serendipitous encounters may not be random, since many attendees have related research experience as the presenter, which share common traits in research methodologies.

Eight ***lookup*** search motives are found, taking up to 13.1% of the total information needs. The *location* mode corresponds to "locate" in Russell-Rose et al's taxonomy. It describes the conditions when a person simply wants to retrieve specific facts. In one occasion, the attendee wanted to find a research paper mentioned by the presenter. Another attendee simply wanted to retrieve the fact about the resolution of human eyes for understanding eye tracking (topic 2). *Recall* is similar to "verify" in Russel-Rose et al.'s framework, but with an emphasis on recalling some already known concepts, probably for verifying the presenters' claim or for consolidating one's own knowledge.

Only 3 needs (4.9%) are of ***investigate*** motive. *Evaluation* corresponds to "evaluate" in Russel-Rose et al's work, but "analyze" and "sythesize" were not observed in our study. The investigate motive corresponds to higher layers of Bloom's taxonomy. However, since most attendees did not have the intention to obtain deep knowledge about the topics, these category is insignificant in the seminar scenario.

Dealing with Information Needs

In response to our second research question, we take a look at how people deal with their situational information needs arisen from the seminars. Attendees can instantly address an information need by initiating a search. They can also ignore it or deal with it in a later time. This section discusses our findings about these issues.

(1) Were the information needs addressed by immediate searches?

A total number of 25 computing devices (cf. Table 5.1), exemplified by laptops or mobile devices, were taken into the seminars. Table 5.3 shows an overview of Web searches driven by

Table 5.3: Overview of web searches driven by information needs. Needs are the total number of information needs in the corresponding category; Searches are the number of Web searches made, and successes refer to the number of searches that successfully resolve the information need.

Mode	Needs	Searches	Successes
Curiosity	13	1	1
Serendipity	2	1	0
Knowing	25	6	3
Elaboration	4	1	0

information needs. Only four modes of information needs had led to Web searches, including 6 *knowing*, 1 *elaboration*, 1 *curiosity* and 1 *serendipity*. 4 out of these 9 searches successfully addressed the information needs, including 3 *knowing* and 1 *curiosity* needs. Others failed to deliver the required information, though 2 of the information needs with failed information searches also diminished as either the presenter explained it later or the subject figured it out by him/herself. Failing to find the required information does not mean the information cannot be found on the Web. One subject indicated his/her search was not successful, but also expressed his/her willingness to make more thorough searches later on. Similarly, another subject said although s/he "understood a bit" from the already conducted searches, s/he would still search more information later.

(2) What are the factors that impede people from immediately searching during the seminar?

We asked the attendees who did not search information even though they had information needs to specify why they were not doing so. 41 answers were collected and we categorize them into 5 categorizes as follows:

- **No Tools**. 24 needs (from 16 attendees) with no immediate searches were associated with "No tools". This is the primary reason for not having searched information. However, it does not imply they would have searched if they had had computer devices at hand.
- **Uncritical** is the secondary reason which was mentioned 5 times. The subjects realized the needs but did not perceive the urgency to search information immediately.
- **Time-consuming**. 5 entries were associated with the time cost incurred by Web searches. Some expressed this factor as "No fast tools" available for searching.
- **Impolite** were expressed by 2 individuals who thought it improper for using personal devices during seminars
- **Interruption** was indicated 4 times. The subjects were reported to be too busy trying to follow what the presenter said, so they could not shift their attentions to a searching task.

(3) Did the information needs remain after the talk?

As discussed before, 9 information needs led to Web searches, but only 4 (44.4%) of these needs were satisfied by Web searches. Meanwhile, 13 of 48 (27.1%) information needs that did not lead to searches also diminished during the course of the presentation. With Fisher's exact test, the number of resolved information needs by Web searching is not significantly (p=.43) more than those without immediate searches. However, according to 13 entries of feedback collected in the questionnaire, the information needs diminished without Web searches for two main reasons: (1) The presenter clarified it in a later stage (30.8 %); (2) The participants were only interested in the information during the presentation but not afterwards (69.2%). The result echoes our assumption made at the beginning of this chapter: simply making search tools and Internet access available may not be effective and efficient enough for addressing information needs during the performance of a learning activity, leaving plenty of room for technologies to support situational needs.

Summary

Research seminars are common scenarios at universities, not only for researchers but also for students. Our survey study shows that most seminar attendees had in-situ information needs, which may come from either oral presentation or the slides for various motives. Some of them may arise due to interests or curiosity. Most participants were not found to have used computing devices to address their information needs during the presentation, but some of the needs vanished as the presentation went on. For example, the presenter might explain it in a later stage, or the subjects thought the needs were not as important as they were perceived before. Even for those who actually performed searches, they did not always successfully find the right information, perhaps because they had little time to make a more thorough search. The one-to-many communication style in research seminars is similar to that of classroom lectures, for which the findings might be informing as well. In learning scenarios as such, participants simply passively receive information for acquiring knowledge, corresponding to the middle or lower layers of Bloom's cognitive taxonomy of learning. Information needs arose in such situations mostly for understanding certain concepts. We believe that computer systems have the potential to automatize the searches based on learners' situational context. We will discuss more about this issue in the next section.

5.2 Research Framework

The survey study presented in the preceding section was conducted in research seminars due to its convenient accessibility, but this dissertation also deals with other types of learning scenarios. A primary concern that provokes us to consider is "Which of the findings presented before can be generalized to other learning contexts? " In activities such as collaborative learning or online learning, the fact that people do not have tools with searching capability may not be valid any more. In addition, these activities, compared to research seminars,

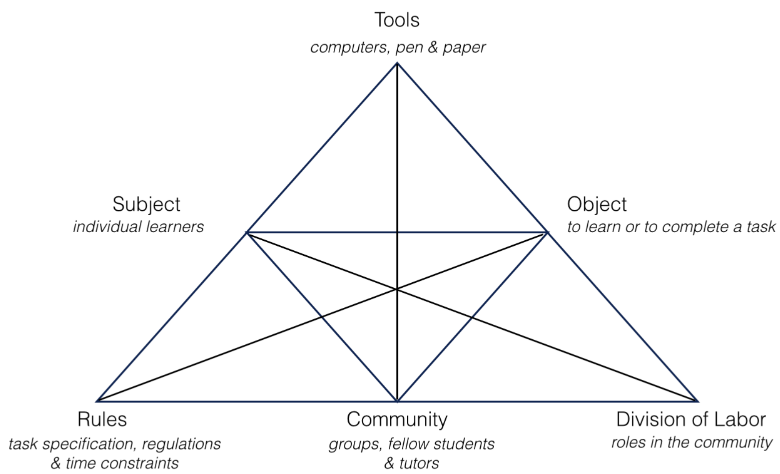


Figure 5.1: Learning activities modeled with Engeström's activity system triangle

may also involve higher cognitive layers in the Bloom's taxonomy, such as evaluating and creating ideas. This leaves us a few fundamental questions that must be answered: What is the research scope of this dissertation? What issues are we addressing? What kind of system are we designing and what are the principles that are followed? The answers to these questions constitute the research framework of this dissertation.

5.2.1 Research Scope, Challenge and Design Principle

Before we proceed to the discussion of what to design and how to design, it makes sense to identify common issues regarding information seeking in various learning activities so as to nail down the design problems. This requires a general framework to deconstruct learning activities. The Activity Theory reviewed in Section 2.2 is one framework as such. This section starts with an analysis of learning activities with the Activity Theory, followed by the design challenges, from which a set of design principles are derived.

Activity Theory's Perspective

There are various learning activities, exemplified by the scenarios illustrated in Section 1.1. The biggest assumption that motivates us is that learners may require information while performing these activities. However, they do not have proper tools to obtain information in current practices, according to our survey study. To better understand learning activities in a general sense, we deconstruct them with Engeström (1987)'s triangle model of activity systems (cf. Section 2.2.1). A schema of common learning activities is illustrated in Figure 5.1.

- **Subject**. The subjects of a learning activity are individual learners. For example, in the seminar scenario, the attendees are the subjects. In collaborative learning scenarios, the group participants are the subjects.

- **Object**. The object of a learning activity is usually to learn something or to complete a learning task. The outcome of this object may be the achievement of learning or certain deliverables.

- **Tool**. Tools in a learning activity can be physical or material, such as pen and paper. In some situations, e.g. MOOC learning, computers are also part of the learning experience. In addition, tools can also be psychological, including the language and signs used for expression. In this thesis, a special emphasis is placed on to the tools that enable information seeking.

- **Rules** There exist a variety of rules in common learning activities. For example, social rules and regulations both support and constrain a learning activity. In the previous seminar study, attendees could not often interrupt the presenter for acquiring additional information. Some of them also found impolite to search while the presenter was talking . This is an example of social regulations. There are other rules such as time limit as well.

- **Communities** In collaborative learning, an example community is the group, but it can also be extended to include other parties that are concerned with the activity, e.g. the teachers or other groups. An individual participates in group discussions to achieve the object. In MOOC learning or traditional classroom learning, the community includes both the teachers and the fellow students.

- **Division of Labor** refers to implicit or explict organization of a learning community. For example, in MOOC learning tutors and students have different roles. In collaborative learning, there also exist specified or emerged roles among the participants. These divisions of labor mediate with other elements and contribute to the transformation of the object into outcomes.

It is clear from the schema that in various learning activities, subjects want to achieve learning or to complete tasks. Although they may encounter information needs, the satisfaction of the needs are usually not the goal of these activities. According to the principle of "hierarchical structure of activity", an activity is composed of actions, and addressing information needs during learning could be a sub-action subordinated to the action of group discussion, video lecture viewing etc. Figure 5.1 does not illustrate the actions, but the schematic graph implies that an information need can be addressed with actions through either tools or community, corresponding to information seeking and help seeking behaviors as discussed in Chapter 3. However, an action is usually performed under conscious attentions, which are typically time consuming and interruptive, requiring allocations of extensive attention resources. In fact, the fundamental problem with addressing spontaneous information needs during learning activities is not the lack of search tools, but the lack of tools that allow searching information with minimal efforts, without interfering with the primary task. To alleviate the problem, innovative technologies can be designed to help learners address their information needs at the operation level.

Elasticity of Information Needs

Addressing all needs arisen in a learning activity at the operation level may sound ambitious, but it is not possible with current technologies. Therefore we must limit our research scope to specific kinds of information needs. This section is devoted to discriminate information needs with the notion of elasticity.

In economic literature, there exists an term called *elasticity of demand*. Marshall (1890) wrote that "the elasticity of demand in a market is great or small according as the amount demanded increases much or little for a given fall in price, and diminishes much or little for a given rise in price". In other words, elastic demand is a type of demands that rise or fall depending on the price of the products. For example, people would buy chocolate bars as snacks, if they find sweets or chips are more expensive than before. In contrast, people buy products with an inelastic demand no matter what the price is. Demands of petrol, salt, water are inelastic, because they are considered as necessities in life.

By analogy, information needs may also be elastic, depending on the incurred costs. The cost of searching information in a learning activity may include attention cost and time cost, which are similar to the price factor in the aforementioned economic examples. We use the term *inelastic information need* to refer to the situation when the required information must be obtained in order to attain the learning goal. In such situations, the needs are strong, so that users must find a way to deal with it regardless of the cost. For example, when a user must solve a problem with concept X, but X is unknown. The student must anyway search relevant information otherwise it is impossible to complete the task. On the other hand, elastic information needs refer to the situation when “whether or not” to address the needs depends on the availability of the required information. In other words, the users may not be aware of the value of the missing information. In such situations, if the information is immediately available and accessible, users are perhaps happy to use it; Otherwise, users may hesitate to search and finally work without it. For example, a student may find it difficult to explain a concept to his fellow student in a study group, but the intention to seek external information for help on this matter may not be perceived necessary. If helpful information is meanwhile prompted, then the availability of the information may facilitate the students’ articulation.

Like the elasticity of demand in economic literature, the elasticity of information needs is subjective. Even when some information is necessary for solving a problem or understanding a topic, the needs can still be elastic depending on the users’ motivation to learn about them, e.g. seminar attendees who are not interested in the topic may have more elastic demand of information.

Design Challenges and Principles

This dissertation is devoted to design technologies for supporting elastic information needs. Inelastic needs are not specifically targeted, though they may still be supported through serendipitous encountering. This decision was made based on the assumption that inelastic needs usually arise when an individual realizes a definite need of certain information and strives for it. In this case, the individual becomes a true information seeker who is likely to favor self-control, autonomy, and freedom in the seeking process, which results in a more thorough search. In contrast, elastic information needs may sometimes **not be fully aware of, not be able to articulate, improper or inconvenient to search for under social or time constraints**. Designing a system to help people recognize and address elastic information needs has potentials to enhance the learning experiences.

The "internationalization and externalization" principle of the Activity Theory provides theoretical insights for understanding the research goal of this thesis. Kuutti (1996) claims that any activity has both internal and external sides. Externalization puts subjects' thoughts into visible or tangible form. Usually externalization is performed by subjects, e.g sketching an idea. In this thesis, we envisage externalization to be realized by a tool that presents to the subjects information pertinent to the context. This process is termed by us as *augmented externalization*. According to the Activity Theory, when a tool is seamlessly integrated into the subject's activities, the interactions with these tools become the subjects' own attributes (the internalization process). Therefore, our goal can be described as an exploration of **how augmented externalization of learning activities can be designed so as to extend the individual's abilities for recognizing and addressing situational information needs**. Such systems are bound to face the following challenges:

- **Distraction**. If the system proactively suggests information to users in a learning activity, the users have high probability of experiencing distractions, which could be detrimental to the overall learning experiences. Technologies should balance the attention allocated between the supporting information and on-going activity [Design Challenge]
- **Relevance**. It is difficult to determine the relevance of automatically sought information, since the contexts are not constant. The supporting information must be made as pertinent as possible to the in-situ learning context [Technical Challenge]
- **Timeliness**. Information can only make sense if arriving at the right time. Supporting information must be made available and the system should notify the users at the right moment in the right form, though the timeliness of information is difficult to determine algorithmically [Design and Technical Challenge]

Precision is arguably another technical challenge. Ideally the system should be able to deduce the precise information required by a user, with the availability of context. Practically this is not feasible, since even the users may not know the what exact information is required. We thereby do not strive to elicit precise information, which explains why "precision" is not listed as core design challenges.

To guide the designs hereafter, we transform the listed challenges to a set of design principles as follows:

- **Calm**. In response to the "distraction" challenge, information should be presented in a way that it is unobtrusive to the learning task, but meanwhile maintains an appropriate level of awareness.

- **Context-aware**. The information should be reactive to the ever-changing context of the learning task. This principle corresponds to the "relevance" and "timeliness" challenges.

- **Redundant**. The target system essentially performs mediated search as computer intermediary. It is not fully certain of user's information needs. As discussed in Section 3.1.2, such systems should follow the *principle of guaranteed result*, by adopting max coverage tactic to return the most complete query results. In other words, the system actually focus on recall rather than precision. It should attempt to retrieve more potentially relevant results, even some of them may be redundant.

- **Trigger-rich**. The target system is also a serendipity-inducing system, with trigger-richness as one of its key features (cf. Section 4.2.3). Information can be delivered as visual stimuli that allow users to easily comprehend it as well as to make connections.

- **Multi-phase Interactive**. The information should have different levels of details, requiring different intensities of attention as well as user engagement.

5.2.2 Design Space and Interaction Phases

As reviewed in Chapter 4, ambient information systems emphasize the awareness of peripheral information that is not intrusive to an on-going activity, whereas serendipity-inducing systems underlines the presentation of trigger-rich context-relevant information to enable connections. The design principles presented before clearly exhibit some of the key facets of both ambient information systems and serendipity-inducing systems. We therefore have reasons to believe that a combination of these two types of systems has the potential to service elastic information needs during learning activities.

We coin a term *contextual information scents* (CIS) to crystallize the "augmented externalization of learning activities" mentioned before, and the CIS should inherit the core properties listed in the design principles. A formal definition of CIS is given as **ambient information cues, which are perceived from the activity context and made available for a user with a conscious or subconscious situational information need**. Parts of the term's naming is borrowed from Pirolli and Card's notion of information scent (1999)(cf. Section 3.1.1). The original notion was defined in an information seeking context, to represent proximal information cues that help an information seeker to judge distal information sources and to navigate towards them. In our definition, the scents are cues that help people trigger the connections between the presented information and certain situational information needs. The scents can manifest themselves either as stimuli that guide the navigation to the desired information, or simply the desired information per se.

Figure 5.2: A supermarket places goods along the escalator as "scents"

Presenting contextual information scents to people engaging in learning activities is in analogy to a popular marketing strategy in supermarkets as shown in Figure 5.2. The supermarket places piles of goods alongside the escalator. When a customer steps on the escalator, his/her goal is simply to go up or down. During his or her travel, the goods constantly emits "scents" that may catch the customer's attention, so certain demands of customer may be activated and strengthened. For example, while traveling with the escalator, a customer may encounter some chocolate bars, which trigger the demands for some of candies. But the demands may be elastic depending on the ease of access to the goods. In other words, the customer may not be willing to go back to the second floor to get them on the shelfs, but is more than happy to grab them during the travel with escalator. Here, the chocolate bars "emit scents" along the customer's travel with escalator, and trigger an elastic need with serendipitous encounter. In a similar vein, the contextual information scents discussed before, are designed to present useful contextual information to support elastic information needs of the learners in the learning process.

The design of contextual information scent is concerned with the design of information presentation as well as information interaction. We propose a design space and a model of interaction phases for guiding the explorations in this thesis.

Where	Privacy	Personal		Public
What	Context	Conversation	Interaction	Content
	Capacity	Low	Medium	High
	Uncertainty	Low	Medium	High
When	Activation	Proactive		Reactive

Figure 5.3: Design space of contextual information scent

Design Space

Similar to other design spaces in HCI (Card et al., 1990; Pousman and Stasko, 2006), the goal of constructing a design space for contextual information scent is to provide a descriptive tool for exploring how information scent can be presented. The following questions inform the design space illustrated in Table 5.3

- *What* sort of information scents should be provided to the users?
- *Where* should the information be displayed?
- *When* should the information be delivered?

The "where" question is concerned with privacy issues, i.e. the information presented publicly available or tailored for individuals? The rationale behind this dimension is that displaying information in public or not may influence the way people interact with it.

The "when" question deals with notification styles. The reactive style means the system retrieves contextual information scents in response to users' immediate behaviors, whereas systems with proactive style notify even when the users are idle.

The key facets to characterize contextual information scents are concerned with the "what" question, which considers 3 aspects: *context*, *capacity* and *uncertainty*. Section 3.1.3 delivers a comprehensive overview of contextual spheres that influences information seeking behaviors. In principle, all the contextual variables can be considered, but in this dissertation we focus on 3 contextual variables for generating contextual information scents. The first variable is *conversation*, which is part of the social sphere discussed in Chapter 3. With conversation context, computer systems capture the semantics of discussions to generate informational content to address elastic information needs. *Interaction* refers to user interactions with the system, which reflects how they work with the tools. *Content* refer to the materials or resources that are worked on. For example, in a MOOC scenario, the content context would be the video content that is being viewed. Information capacity is in debt to the corresponding dimension in Pousman and Stasko's taxonomy of ambient information systems (2006). It represents the number of discrete information sources that the information scents can represent. For example, the capacity of a word is low since it only carries a single meaning, but a full page of words may contain high amount of information. The information scents aim at deducing users' information needs from the ever-changing context, so there is uncertainty about which exact pieces of contextual information should be taken into account for retrieving useful information and which should be presented to users. For example, in the conversation context, if one participant says "The president of EPFL is going to New York to attend a meeting in MOOCs", then various elements, such as the person (the president) the city (New York), the object (attend a meeting) or the condition (MOOCs) can be augmented with abundant information. Computer systems may randomly combine these elements to retrieve information, or simply concatenate the terms in order, resulting in high uncertainty. Certain rules may also be applied to filter or prioritize the information, leading to medium level of uncertainty. Information selection can be pre-determined before the activity takes place as well. For instance, in classroom or MOOC lecture viewing scenarios, the lecturing content is prepared in advance, so enriched information may be elicited in advance as well, which leads to information scents with low uncertainty.

Interaction Phases

The presented design space portrays how the target system presents information scents to users. The next step is to interact with the information. We envisage the system to visualize information scents in the periphery of the users' attention during the learning activity, to spark serendipitous interactions. Whenever interesting information is encountered, users have the opportunity to explore it in more details. We adapt the interaction phase framework (Vogel and Balakrishnan, 2004) (Section 4.1.2) to conceptualize user interactions:

- **Implicit Interaction**. Visual changes that capture users' attention for potential interesting information.
- **Subtle Interaction**. When the display is within a user's peripheral view, it allows the user to get the information by glancing at it.
- **Personal Interaction**. If a user is interested to know more about certain information, they can select the information to view more details.

In principle, the first two phases of interaction can be operationalized (i.e. they are subconscious operations), but the last phase is involved with conscious interactions with the goal of attaining the target information. In the following chapters, different designs of the contextual information scents will be presented and evaluated. The exploration will be framed within the design space and interaction phases presented in this chapter.

6 Exploring Conversation Context

If we say the seminar talk scenario is most convenient to run a survey study for gaining insights about information needs, then collaborative learning activities are most suitable for us to conduct experiments for exploring the design of contextual information scents. Here we refer to the broadest definition of collaborative learning, where a group of participants attempt to learn together. The notion of "learn" includes not only "studying a course" or "working on a problem", but also "learn from collaborative work practices", i.e. the learners construct knowledge through collaboratively working on a specific task. This would broaden our scope to many kinds of collaborative working scenarios that involve building collective knowledge, such as "brainstorming", "problem solving" or "decision making", so that we have more freedom to design tasks, select participants, and observe their interactions. This chapter, together with the two that follow, present our explorations of information scents that utilize different types of contextual variables in the design space (cf. Section 5.2.2). We built two systems to gain insights about the design of contextual information scents generated from conversation in collaborative scenarios. The following sections present two explorations: RaindropSearch [1] in Section 6.1 and TileSearch [2] in Section 6.2. The former is a system that captures and displays conversational keywords as contextual information scents, which can be used as query terms to search information. Similarly, conversational words are also captured in the latter system, but Web searches are made automatically by combing the words as queries. The search results are displayed to users as information scent that guide them towards further interactions. Both projects combine ambient information interaction and serendipitous encountering.

6.1 Explorations : RaindropSearch

In collaborative groups, participants share knowledge and build common understanding through argumentation and cognitive elaboration. In the meanwhile, certain information needs may arise during the collaborative work, and computer technologies can help them address these needs. With people searching information in a collaborative environment, our

[1] based on paper [8] and [12] in the publication list
[2] based on paper [8] and [13] in the publication list

research is easily associated with a well-established field in HCI, i.e. *collaborative search*. Morris (2013) views collaboration search as a kind of social search where participants work together to satisfy an information need. In other words, collaborative search systems are designed exclusively for group participants engaging in a search-oriented task, where the information needs are more inelastic. This is the main distinction between a collaborative search system and our target system, which places more focuses on elastic information needs.

Since conversation constantly shapes the collaborative context in collocated collaborations, it is quite sensible to assume that many of the information needs can be deduced from it. Therefore, "real-time conversation" stands out as a candidate for generating computer support of group's information needs. To start with, we envisage a scenario where a system captures the keywords in conversations and visualizes them in a certain way, so that the group participants have a chance to easily make use of them for searching information. The visualization of conversational words is what we called *contextual information scent*, because they are designed as information cues navigating the users towards potentially desired information. By mirroring social signals as contextual information scents, the system attempts to realize the so-called "augmented externalization" to increase the awareness of information needs.

6.1.1 Related work

To our knowledge, little research has endeavored to exploit conversational features for supporting information needs, but many researchers have attempted to visualize conversational features as feedback to group participants for other purposes. A selection of such projects are presented in this section.

Mirroring Non-linguistic Conversational Features

Research work that attempts to mirror non-linguistic features of group conversation as social signals are far from rare. Several research projects have presented ambient systems that visualizes conversational behaviors in meetings, including the Reflect (Bachour et al., 2008), Second messenger (DiMicco et al., 2004; DiMicco and Bender, 2007), Relational Cockpit (Sturm and Terken, 2009), Meeting Mediator (Kim et al., 2008), and Conversation Clock (Bergstrom and Karahalios, 2009b). The most often utilized conversational features are the speaking time and intensity contributed by each participant, researchers found mirroring these features to help the groups to self-regulate.

Mirroring Linguistic Conversational Features

Compared to the employment of non-linguistic conversational features, mirroring linguistic conversational features in group work finds itself in a smaller body of research. WordPlay (Hunter and Maes, 2008) is a multitouch tabletop that aims to support collaborative activities, such as brainstorming and decision making. The system has a dedicated speech recognition component to capture and display speech input from group participants. The recognized words are linked with a semantic knowledge database to aid idea generation. It should be noted that WordPlay does not extract words from natural conversations. Speech recognition simply works as an input modality in addition to keyboards. The Conversation Cluster Table (Bergstrom and Karahalios, 2009a) also employs a speech recognition component for capturing conversation. Compared to WordPlay, it automatically extracts conversational words and performs semantic analysis on the words. As a result, the table visualizes distinct topics emerged from the conversation. The goal is to support group activities with content recall and idea formulation. Rocchi et al. (2008) also reports a tabletop system that attempts to understand the conversation of the people around it, and present related visual stimuli to promote new topics for discussions. The system is targeted for groups of museum visitors to reflect their museum visiting experiences. Note that the last two of the aforementioned systems, which attempt to capture real-time verbal conversation, were only studied with Wizard-Of-Oz to simulate speech recognition, perhaps due to technological constraints of current speech recognition technologies.

Several projects have attempted to mirror linguistic features from text messaging in online collaboration, and it is more reliable to capture text-based conversation. Groupmeter is such a tool that provides linguistic feedback, such as proportions of "agreement word" use, to online collaborators (Leshed et al., 2009). It is reported that such feedback raised awareness of each participant's own language use. Idea Expander (Wang et al., 2010) is an example system that supports group brainstorming by presenting pictorial stimuli pertinent to the conversation. It is shown that the employment of the Idea Expander has increased the number of ideas produced by individuals.

6.1.2 System Design

As a very first exploration of conversational information scents, RaindropSearch is designed as a tabletop system to provide group users with a visualization of conversational words that can be used as query terms for Web searches.

The Projection-camera System

Following the interaction phases proposed in Section 5.2.2, the system visualized conversational words in an ambient way and allowed further interactions with them, i.e. searching and eliciting information. Our design publicized the display of conversational words but personalized the search experience, since different users may have various information needs. Upon these considerations, the RaindropSearch system adopted a top-projection augmented tabletop surface for displaying conversational words as contextual information scents cuing for Web searches. On one hand, a top-projection system can be easily deployed in most existing working environment with a table, which yields a large surface for interaction. On the other

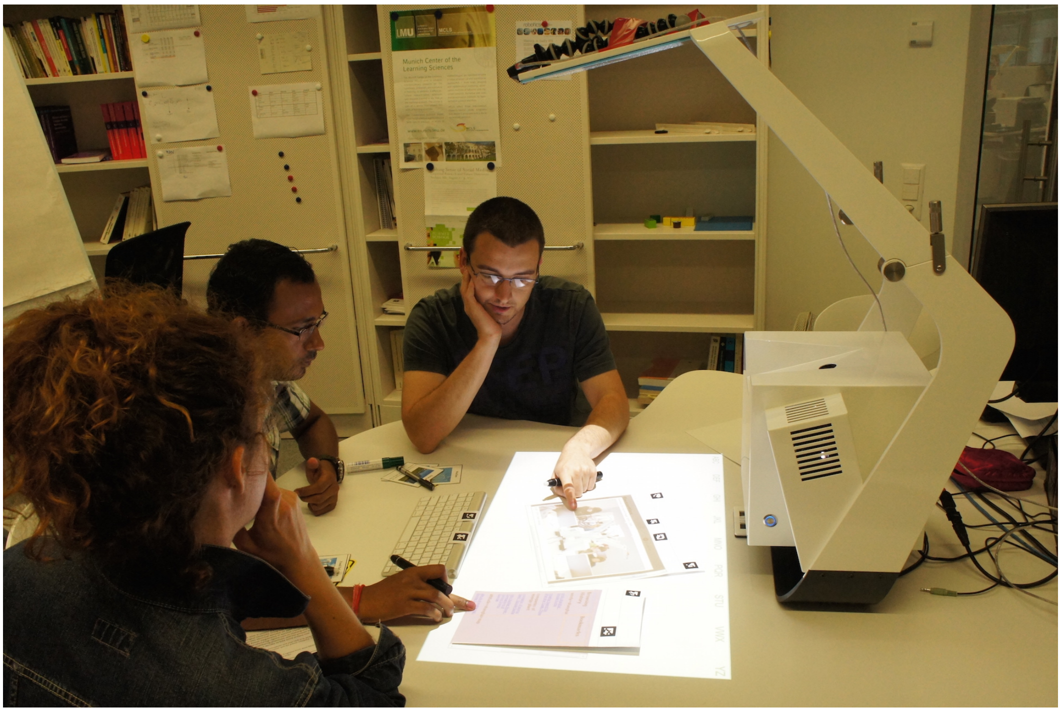


Figure 6.1: Raindrop system setup

hand, it also creates projections for flexible displays that go beyond the table surface. Prior research projects, such as PaperLens (Spindler and Dachselt, 2009) and TinkerSheets(Zufferey et al., 2009), both employed paper sheets as secondary displays in addition to an interactive tabletop surface. We attempted to achieve a similar vision, by employing paper sheets as flexible Web browsers, with contents being projected by a projector. The paper Web browsers can simply be "plugged into" the projection area for searching and viewing information, and "plugged out of" the scene to free the space for the tabletop display.

The RaindropSearch system setup is illustrated in Figure 6.1. It is built upon the TinkerLamp, which is an augmented tabletop environment (Zufferey et al., 2008). The hardware system is composed of a camera, a projector (1280 × 768 pixels) and a mirror. The mirror is facing the table at approximately one meter height to enlarge the projection surface with an increased projector-to-screen distance. The projected area is of 73 by 45 centimeters. The lamp is able to track fiducial markers thanks to a fiducial-tracking library, the Chilitags (Bonnard et al., 2013), which was developed at our lab. In addition to the original TinkerLamp setup, the RaindropSearch system also includes an infrared camera fixed on top of the mirror. This camera is used to track infrared pens, which work as input devices for the paper browsers. An infrared pen has a press-sensitive tip. By pressing it, the tip emits infrared light that can be seen by the infrared camera.

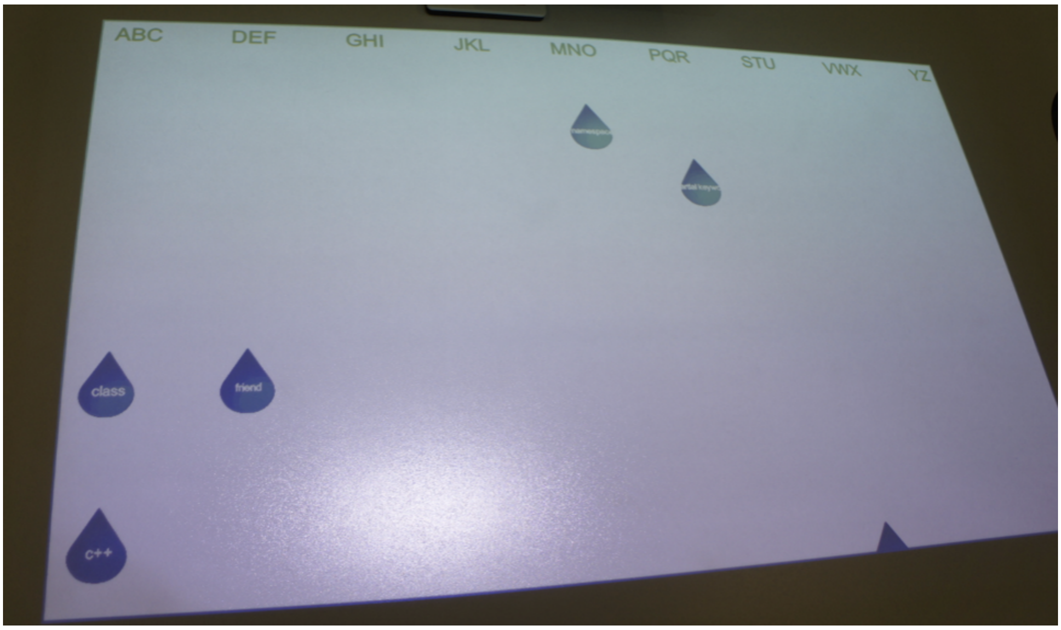


Figure 6.2: Raindrop ambient display

The Raindrop Ambient Display

The tabletop display is designed to silently mirror a group discussion with a visualization of keywords captured in the conversation. However, it is nontrivial to define what spoken words should be considered as keywords. Since nouns usually appear as subjects or objects in a sentence, these words often define the spoken context. In fact, nouns are often used for screening keywords (Shah et al., 2003) in research literature. In the RaindropSearch, we simply treated nouns spoken by the participants as keywords, and visualize them for potential searches. This strategy follows the principle of guaranteed result (cf. Section 3.1.2) by adopting a max-coverage tactic, since the system is essentially a search intermediary who knows little about user's in-situ information need.

We assume recent words are of more interests to the users for sparking elastic information needs, so the design of the display is strongly oriented towards the visualization of the words' "recentness". Initially we implemented two visualizations. The first was a spiral route along which all the captured keywords traveled, with the sizes of words decreasing with time. The second was a pipe hole which emitted words in bubbles that also shrink with time. These two visualizations were simple and intuitive. However, both were messed up as more words were captured: With lots of words in different sizes being scattered around, a user barely recognized any of them in the visualization. To alleviate this problem, we came up with an idea for arranging the spoken words on the display, which gave birth to the RaindropSearch system.

The display adopts a "raindrop" metaphor. Real-time conversational nouns are enclosed in rain drops and fall down from the upper border (See Figure 6.2) at constant speed. These

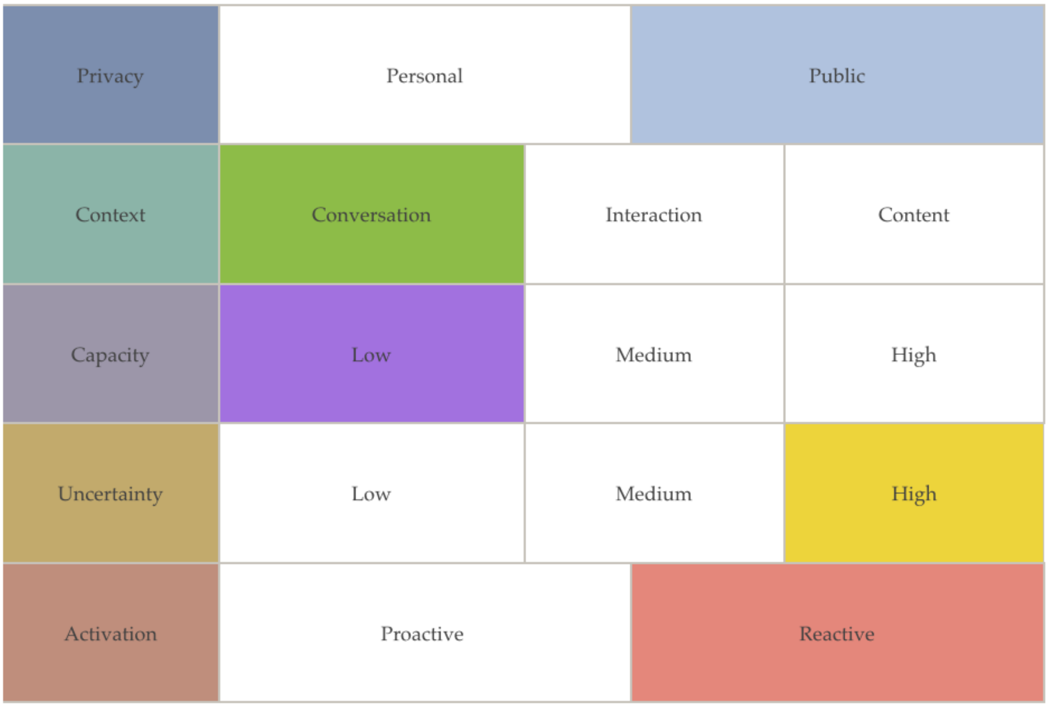


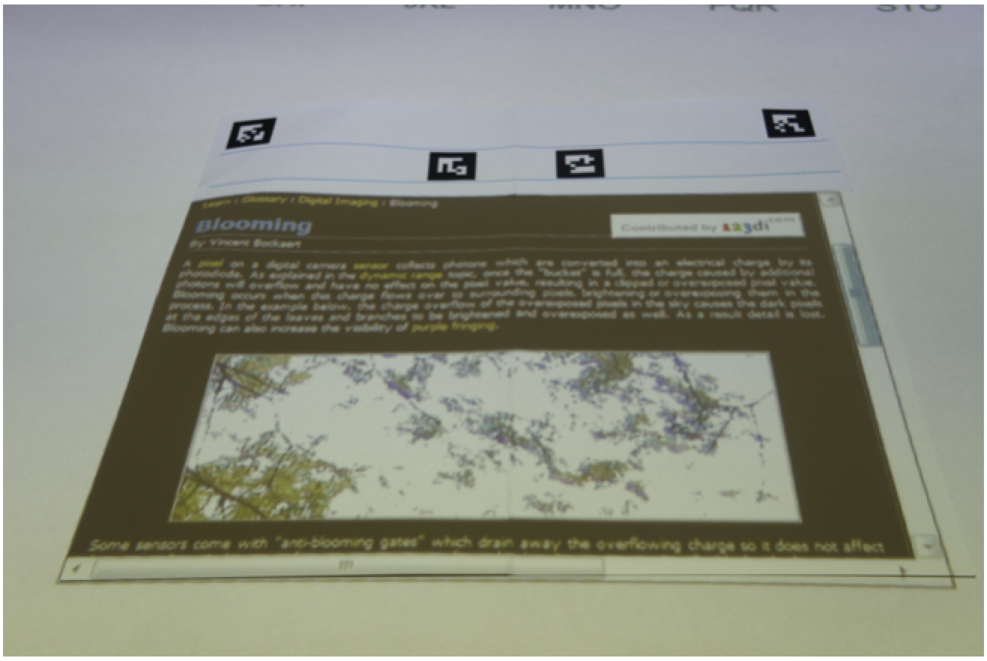
Figure 6.3: Raindrop ambient display in the design space

rain drops are arranged according to alphabetic order of the words contained. A rain drop is highlighted in darker color, when the word it contains is spoken more than once during its life cycle. Before a word reaches the bottom and disappears, it can be used for composing search queries.

The ambient component of the system finds itself in the color cells of the design space (cf. Section 5.2.2) shown in (Figure 6.3). The system reacts to the group conversation and visualizes publicly the spoken words in the form of "raindrops". These visualizations are contextual information scents, but they carry low information capacity and high uncertainty: it simply displays all the spoken nouns, which are barely predicable.

The Paper Interface

The system is designed for small groups composed of at most three users due to space constraints. Each user may have a designated paper interface for personal Web searching and browsing. A paper interface has an A4 size area for displaying projected Web content. Each paper interface also has an identification zone extended to the longer edge, with 4 unique fiducial markers printed on either of its two sides. These fiducial markers are used for identification and tracking purposes. The positions of the markers are designed with ergonomics. The upper part of a paper sheet is less likely to be occluded by human body during interactions, so the marker zone is extended to its top edge. The two topmost markers are placed on corners,



(a) Recto side of a paper browser

(b) Verso side of a paper browser

Figure 6.4: Recto and verso sides of the paper interface

and the region in between is left for displaying interaction feedback, such as commands issued by gestures or input devices. The other two markers are close to the lower center. The empty area between the lower markers and the shorter edges are reserved for swiping gestures with the infrared pen, so as not to occlude the markers while performing the gestures.

The recto side of a paper is for Web browsing (Figure 6.4(a)), whereas the verso side logs personal browsing history and bookmarks. Each visited Webpage is automatically logged in the history. Since each Web search may lead to multiple visits of different Webpages, the history is organized by the query terms under which the page-visits are made. Whenever a user clicks a link in the history, the paper interface loads the Website on the same side, and the other side automatically turns into history logs. Therefore, the notion of recto and verso are not absolute. They are subject to change during user interactions. One advantage of doing so is that a user does not have to remember which side is which.

Multiple paper browsers share the same Web history for increasing awareness. The flexible paper displays extend the public tabletop display with personal spaces. They can be easily ”plugged” into and out of the tabletop, thus separating the ambient and the interaction space without altering the existing environment.

6.1.3 Interactions

Interactions with the RaindropSearch system may take place either on the ambient display or the paper browser display. This section describes various interaction possibilities the system has offered.

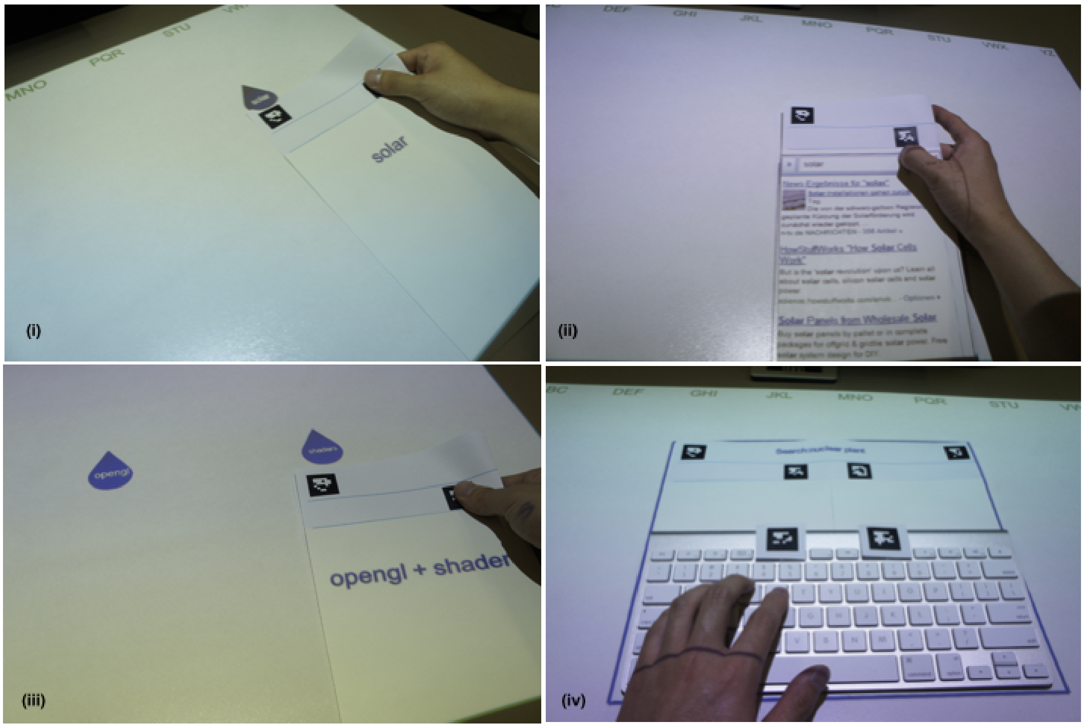
Searching with Paper Interfaces

When a user comes across on the tabletop display a word, of which a Web search is perceived as helpful for the task, s/he can hold a folded paper interface to intercept it with thumb occluding the lower marker (Figure 6.5(a)(i)). Multiple words can be selected with the same gesture to compose a more complex query (Figure 6.5(a)(iii)). Upon releasing the thumb from the occluded marker, a Google search result page is displayed (Figure 6.5(a)(ii)).

Though not of our research focus, we realize that inelastic search needs may arise unavoidably in a group discussion. The system hence supports conventional keyboard search as well. A wireless keyboard tracked with fiducial markers are used in this situation. Placing the keyboard onto a paper browser would connect them, so that a user can query Google with his or her own keywords (Figure 6.5(a)(iv)). Only one keyboard is available for the group, because the support of collaborative search is out of our focus.

Browsing with Paper Interfaces

The paper interface, thanks to its material, is foldable. We have shown previously how a folded half-sized paper browser is used for searching words on the ambient display. A folded paper browser has increased rigidity, so it is easier to grab and hold. Another advantage is its reduced



(a) Searching with Paper Interfaces

(b) Browsing with Paper Interfaces

Figure 6.5: Interacting with RaindropSearch

size. In a limited projection space, resizing a paper browser to half size whenever needed, is expected to be a solution to avoid occlusion. When a user folds a full-sized paper sheet (Figure 6.5(b)(i)) into half (Figure 6.5(b)(ii)), the borders and positions of the projections are adjusted accordingly.

A paper browser's folding states are differentiated by examining which markers are visible to the Tinkerlamp camera. The 8 markers on both sides of the paper interface allow more folding gestures than simply folding into half. Thinking of reading a book, a page can be marked by bending its corner. The bookmark function of our paper interface adopts this metaphor: Its top corners can be bended to add the current Webpage to the bookmark list (Figure 6.5(b)(iii)). With an infrared pen, forward and backward navigation can be achieved with a swipe gesture on the reserved interaction zone.

Expected Usage and Benefits of the RaindropSearch

When working on a collaborative task, the participants are expected to focus on interacting and discussing with each other. The Raindrop ambient display constantly updates information scents in the form of animated conversational words. Accordingly, users are at the implicit interaction phase by default. The contextual information scents "emitted" by the display allow subtle interactions. The information scents are expected to service collaborative work in two aspects, as illustrated by the following two hypothetical scenarios.

First, we expect information needs of the users to be sparked by certain words serendipitously encountered on the display. The users then intercept these words for searching information. As a result, the user enters the personal interaction phase. The benefit of the display lies in the facilitation of recognition of information needs with an externalized representation of group discussion.

Second, we expect users to be aware of certain information needs out of some spoken words in the recent past. The user then locates the corresponding words on the ambient display and compose queries for searching information. Then, s/he enters the personal interaction phase. In this case, the expected added value of the ambient display lies in its convenience for information access. If the information needs are elastic, users prefer searching with minimal effects, according the principle of least effort. Otherwise the users have the option of employing the keyboard to build well-articulated search queries.

6.1.4 Piloting RaindropSearch

We conducted a pilot study with colleagues for completing a collaborative task. No experimental conditions were manipulated, because the goal was to better approximate the usefulness of RaindropSearch. We were interested in checking out if the expectations of system usage and benefit were met.

Figure 6.6: Power technology flashcards

Task and Participants

We recruited 12 colleagues (three females) from the lab as our subjects. They aged between 25 and 49 years old, with background in either social sciences or computer science. The subjects were divided into 4 groups, each containing 3 members. Each group was asked to complete a collaborative decision-making task. The task was about the energy crisis in Shanghai region in China. The goal of each group working session was to decide on the types of power plants to be built for solving the energy-shortage problem in Shanghai. The choices of power technologies included nuclear, wind, solar, fossil fuel, hydroelectric, tidal and sea wave. On average, the subjects rated their general knowledge of power technologies below the mid-point ($\mu = 2.24, \sigma = .95$) on a 5-point Likert scale (1 = least knowledge, 5 = most knowledge). It was assumed both elastic and inelastic information needs would arise during their performance of the task.

Procedure

Each study session started with a brief introduction of the collaborative task. Then, each group was given a pen, a map of Shanghai region, as well as a set of flash cards, each of which contained a picture illustrating one type of power technology and a short text explaining its capacity. As mentioned previously, the main focus of the study was not about examining how inelastic information needs were addressed. For example, they might need to know what a nuclear plant is and how it looks like. These cards gave basic information about the power technologies so that they would not have to search basic facts about some unfamiliar terms at the beginning. Then, we explained the features of the RaindropSearch system, which they had 5 minutes to get familiar with. During the study, each group had 30 minutes to discuss the issue and make a decision. The delivery of the task was an oral presentation for explaining their decisions to the experimenter. At the end of the study, the subjects were asked to fill in an open-ended questionnaire (cf. Appendix B.1) to give feedback about the system.

The experimenter stayed in the same room where the study took place, and employed an Wizard-of-OZ approach to simulate speech recognition. Noun words heard from the conversation were entered manually into the RaindropSearch system. The experimenter also produced observation notes and conducted semi-structured interviews with the subjects at the end of each session.

6.1.5 Evaluating RaindropSearch with the Activity Checklist

The Activity Checklist (Kaptelinin et al., 1999) reviewed in Section 2.2.1 provides a tool to support the design processes of interactive systems. Analyzing systems with the checklist is especially effective in early phases of development. Considering the RaindropSearch is our first exploratory prototype, the Activity Checklist is "the" tool for evaluating the system.

The checklist largely relies on the basic principles of the Activity Theory. It involves four categories of questions to design and evaluate a system within a space of context: means/ends, environment, learning and development. Each of these categories describes from a different aspect how the technology supports target operations, actions and activities. Details about the principles and the four categories are discussed in Chapter 2. Kaptelinin et al. (1999) also provide a set of sample questions derived from the list in each category. According to the authors, not all the sample questions are required to be asked, and evaluators may also formulate new questions based on the checklist. In the evaluation of the RaindropSearch, a selection of the questions were answered. In this section, we discuss in greater details about the means/ends category, because it is particularly relevant for us to understand the design problems of the system.

Means and Ends

Means and ends category corresponds to the principle of "hierachical structure of activity". Specifically, questions in this category are concerned with how the RaindropSearch facilitates or constrains users' goal of making group decisions. Note that this goal is associated with the action (Activity Theory's terminology) of group discussion, which may further consist of a sub-action of searching information for fulfilling the sub-goal of addressing the information needs during the performance of the main action.

(1) Is there any functionality of the RaindropSearch system which is not actually used? If yes, which actions were intended to be supported with this functionality? How do users perform these actions?

RaindropSearch provides means to support the action of group discussion through the facilitation of its sub-action, i.e. information searching. The system offers two search functionalities, searching with keyboard and searching with conversational information scents in the form of "word raindrops" on the ambient display. According to the information journey model (cf. Section 3.1.1), the contextual information scents only attempt to operationalize the stage of recognition of information needs, but the other stages of information seeking may still work at the action level.

Information searched in the activity

In all recruited groups, subjects searched information with means provided by the system throughout the meetings. They usually searched for geographical and demographical information about Shanghai region, such as population, climate, natural resources etc. These information was beyond the information available to them in the map or in the flashcards. The participants used information as such for evaluating the feasibility of specific technologies. Other searches were devoted to understand the implementation details of various power technologies as well as their pros and cons. Most of the aforementioned information was essential for the group to making decision. Of course, the pros and cons of different technologies could be discussed in the group without looking for an immediate answers. From the help-seeking point of view, obtaining answers directly from the Internet is a *executive help-seeking* behavior (cf. Section 3.2.1). However, the subjects were not domain experts and they were recruited for completing the task. Executive help seeking in this case was more favorable for the group participants.

Search functionalities used in the activity

The majority of the information needs associated with the search actions summarized before were inelastic, since they were essential information required for making decisions. Our subjects showed preference of using keyboard for entering search queries most of the time.

When we say a functionality is not actually used, we mean either it is seldom used or no longer used after initial trials. Both applied to the contextual information scents expressed as conversational words, which were assumed to be used for composing search queries. In fact, in the first few minutes of each study session, almost all subjects were deliberately trying to use the conversational words for making searches, perhaps due to novelty effect. However, the subjects treated the words simply as an additional input modality from speech. This phenomenon was both observed by the experimenter and reported by the participants in the questionnaire. One subject commented "*We never looked at the words that were coming out* *automatically. We spoke intentionally the words to search*". Another subject also said "*When we want to search for something, I tried to pronounce the word, and it gave me (the word on the display).*" Even the deliberate use of conversational words diminished after a short while and the words were seldom used again. This phenomena may attribute to two main reasons. First, novelty effects faded away rapidly; Second, most subjects left their paper browsers under the projection space after initial searches. As a result, the words were occluded, so that the information scents were not visible any more.

In fact, elastic information needs also arose during the study. It is reported that subjects in one of the groups would have searched two specific words, "monsoon" and "uranium", if they had seen the words right on time. The two words were spoken by two participants during the discussion about tidal and nuclear technology respectively. The subjects said they glimpsed at the tabletop right after they spoke the words for a very little while without seeing the words, then they decided to leave without it. One user reported the difficulty in splitting attention: "*It is difficult to keep attention on the table and on the problem*". It is likely that the words were not yet shown at the time when the subjects were glancing at the table. These words were essential for making decisions. Rather, they were potentially helpful for the subject to articulated ideas. None of the two words were finally searched with keyboards. The information needs were elastic, but the RaindropSearch unfortunately missed the chance to offer support.

(2) What are the basic limitations of RaindropSearch?

We summarize the limitations of the system from two aspects: the use of system devices such as infrared pen and the foldable paper interface; the design of contextual information scents in terms animated conversational words. Some limitations are owing to hardware constraints, whereas others are due to design flaws.

System devices

The design of foldable paper browsers received very positive feedback from the subjects, who mostly praised the novelty and ease of use. Only two subjects left negative comments. One complained about the limited projection space on the table for multiple people to work together. The other mentioned the projection was sometimes blinking. Both problems are due to hardware constraints. It should be noted that the feedback about paper interfaces were mostly based on the subjects' perception of them as output devices (for viewing and interacting with Web pages) rather than input devices (composing search queries with raindrops on the ambient display). As discussed previously, the subjects left their paper browsers on the table for most of the time, occluding the words, which unfortunately prevented the use of paper browsers as input devices. This is actually yet another complaint about the limited projection space.

Most of the negative feedback regarding system devices were from the infrared pens. Subjects were asked to use an infrared pen as if it were a normal pen. The pen tip emitted infrared light so that it could be tracked. Actually, the emitted light was often occluded by a user's own fingers. As a result, gestures such as swiping for backward navigation were sometimes not recognized.

The design of contextual information scents

Section 5.2.1 presents 5 principles for the design of contextual information scents: calm, context-aware, redundant, trigger-rich and multi-phase interactive. The raindrops of conversational words are definitely context-aware and contain redundancy. It also allows further interactions for composing and searching information. The other two principles, calmness and trigger-richness, are required to be examined.

In fact, none of the participants reported distractions of the display, even the system constantly reacted to conversation and updated the raindrop animations during their meetings. This seemed to confirm the calm design of the information scents. However, an ideal calm technology should subtly convey information that attracts users. We would rather describe the "calmness" of the RaindropSearch display as "omittance", since the words were almost neglected for most of the time.

Were the conversational words perceived as good candidates for forming search queries? This is about the trigger-richness of the information scents. 7 out of 12 users held negative views about it. One subject commented in the questionnaire, "*The (conversational) words are too general to do the search*". Another subject agreed that the words shown on the table were sometimes useful for making searches,"*but we didn't use them. We prefer to use our own words*". Clearly these users had well articulated query terms in mind and believed searching these terms would be a better strategy.

In brief, the design of the information scents adopted by RaindropSearch has limitations in its expressiveness. As a result, the display was ignored and difficult to trigger searches in real collaborative meetings.

Environment

The environment category can be better described as social and physical aspects of the environment (Kaptelinin et al., 1999). The questions in this category focused on the "integration of target technology with requirements, tools, resources, and social rules of the environment" (Kaptelinin et al., 1999).

(1) Is the RaindropSearch system considered an important part of work activities?

The target activity took place in a social setting, where the subjects articulated their own ideas and argued against each other in a group to reach an agreement. When an information need arose, subjects either addressed it with Web searches or by discussion. As a result, searching with the RaindropSearch was not always necessary.

(2) Are characteristics of RaindropSearch consistent with the nature of the environment?

The nature of the collaborative environment can be described as a group of people sitting around a table to discuss face-to-face. The RaindropSearch setup employs a horizontal tabletop display, which was designed to anchor user interactions. However, the system was meant to support information searching, a sub-action subordinated to the action of group discussion. The majority of user interactions did not have to involve the RaindropSearch, and the subjects' attentions were mostly rested on each other. This has made the characteristics of RaindropSearch setup inconsistent with the environment. When the subjects started discussing, the horizontal ambient display was located on the edge of their peripheral visions, making it difficult for the participants to perceive and recognize the information scents. Additionally, the words were scattered around and constantly moving, leading to more obstacles for recognition.

Learning and Cognition

The learning and cognition category corresponds to the "internal vs. external components of activity and support of their mutual transformation with target technology". We discuss how the system has influenced the coordination and reflection of the collaboration work.

(1) Does the RaindropSearch support coordination of individual and group activities through externalization?

The display augments collaborative activities by mirroring the on-going conversation on the table. Since the words are displayed nearly in real-time, and remain on the table for a short while, the display can be seen as an augmented externalization of the group activity, which serves as a short-term external memory for the group. In our hypothetically scenario presented before, such external memory might anchor conversation and coordinate the collaborations among the group participants. This unfortunately did not happen because the display was almost ignored by the participants.

The coordination of group activities was more or less achieved by the employment of individual paper browsers, which allowed the subjects to coordinate their roles in searching and discussing.

(2) Does the RaindropSearch provide representations of users' activities, which can help in goal setting and self-evaluation?

The conversational words displayed on the tabletop display represent the users' discussion activity, whereas the shared Web visiting history on each the paper browser represents the users' searching activity. The conversational words, as discussed before, were mostly ignored by the users, but the history logs were occasionally used by the users to reflect what information had been searched.

Development

The development category involves the developmental transformations of other four categories. Many checklist questions in this category are concerned with the developmental changes over time, e.g. user's attitudes or behaviors. Our pilot study was not conducted in a longitudinal manner, so we do not consider the questions regarding the long-term effects of the system. Instead, we focus on understanding the consequences of employing our technology as well as checking if the TileSearch worked as we expected.

(1) What are the consequences of implementing the RaindropSearch on target actions?

In collaborative decision-making, participants articulate ideas and argue against each other to reach a decision. The most direct consequence of RaindropSearch is its support of information searching to facilitate task performance. Without it, the participants would have suffered from the lack of information. From this perspective, the system achieved its goal, but only through consciously formulating query terms and typing them with the keyboard. Elastic information needs were not supported by the information scents.

(2) Did expected benefits actually take place?

We illustrated two hypothetical scenarios where displaying contextual information scents in the form of conversational words might be useful. Serendipitous encountering did not occur, and searching with conversational words was also not perceived as requiring less effort than simply typing query terms into a search engine.

6.1.6 Lessons Learned

In the preceding section, we presented a formative evaluation of the RaindropSearch system with the Activity Checklist. This section aims at summarizing the lessons learned from the evaluation in the following aspects:

- **Calmness and trigger-richness**. Conversational words carry low information capacity when used as contextual information scents. The words were assumed to offer more freedom for composing search queries, but the subjects in the pilot study preferred to formulate queries by themselves. Moreover, the low capacity of the information scents were overly "calm", so that our subjects ignored them. Perhaps the design focus should be shifted from the facilitation of recognizing query terms to the recognition of potentially interesting information.
- **Separated spaces for different interaction phases**. The current implementation of RaindropSearch separates the space for the ambient and subtle interaction phases and the personal interaction space. Though the separated spaces were conceptually detached, they were attached physically (i.e. share the same projection) in the current setup of the RaindropSearch. As a result, the ambient display was occluded.

- **Habituation**. In usual collaborative work settings, participants are used to ordinary tools such as laptops, pen and paper. RaindropSearch offered an interactive tabletop with "ambient intelligence", together with novel devices such as the foldable paper interfaces and the infrared pen. Although the interactions were designed to be "natural", subjects did not perceive the fluidity of interaction when performing tasks. Therefore, people need time to habituate the novel interactions offered in the system.

6.2 Explorations : TileSearch

Similar to RaindropSearch, the TileSearch also attempted to derive information scents from conversational context. As a follow-up system, the design of TileSearch considered the lessons learned from its predecessor. This section presents the design and evaluation of this system with a collaborative brainstorming task.

6.2.1 Research Questions

Unlike the RaindropSearch, which was prototyped as an initial exploration of contextual information scents, the TileSearch was designed with specific research questions in mind. In order to increase the information capacity so as to improve trigger-richness, the TileSearch performs automatic searches based on the conversational words in the group discussion. That is to say, instead of presenting raw conversational words, TileSearch presents the search results directly as contextual information scents to trigger serendipitous encounters. The change raised two main research questions for us, each will be answered by both qualitative and quantitative evidences.

- **Does presenting conversational information scents in the form of automatic search results actually (1) induce serendipity (2) facilitate group discussion (3) useful for idea generation and validation?** With this research question, we aimed to understand the consequences of displaying the contextual information scents. We anticipated that the information scents potentially carried three different roles. First, they may capture group users' attention and induce serendipitous encountering. Second, the serendipitously encountered information may serve as group facilitator. Last, the system may help a group generate more ideas or validate ideas. The last point was partly inspired by a similar study of the Idea Expander (Wang et al., 2010). As reviewed before, the Idea Expander was an online brainstorming facilitator, which was able to turn conversational words into pictorial stimuli. Wang et al. found that individuals generated more ideas when working with the system than without it. However, Idea Expander extracted conversational context from online chatting messages and the pictorial stimuli were queried from a database of artificially labeled images. It is not clear if the finding is generalizable for a system that retrieve information directly from unlabeled Web resources for the support of face-to-face brainstorming.

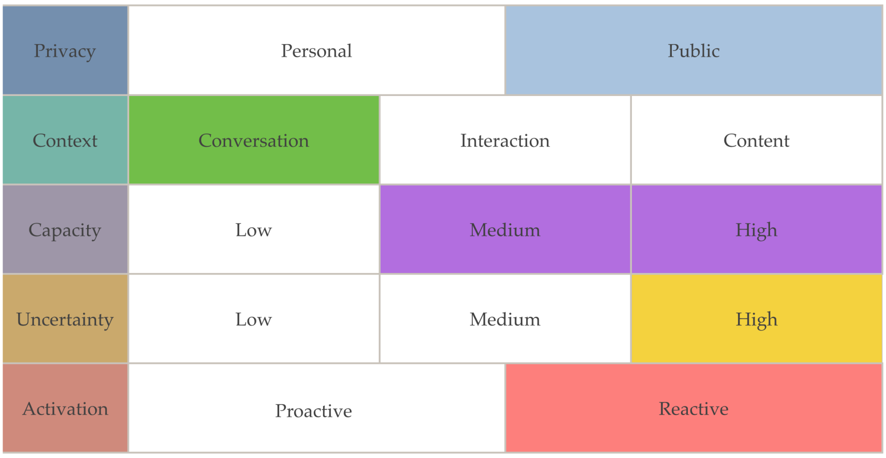


Figure 6.7: TileSearch in the design space

- **Are pictorial stimuli perceived as more useful than textual stimuli to serve the group in the aforementioned aspects?** This question is concerned with the types of search results to be presented to the groups as contextual information scents, which easily provokes one to think of two candidates, images and text. Presumably images are better stimuli since they carry richer information with more abstract representations, which allow immediate comprehension.

6.2.2 The TileSearch interface

Similar to the RaindropSearch system setup, TileSearch also employed a TinkerLamp to render a projection surface for interactions. Paper browsers were abolished in the design to avoid clutters and occlusions. Infrared pens were called off too. As a result, all user behaviors were anchored on the projection surface with more natural finger interactions or hand touch interactions. This would ideally induce more habituated collaborative experiences. The TileSearch user interface consists of two types of views, the *search view* and the *Web view*. The former view presents search results as contextual information scents (implicit and subtle interaction phases), whereas the latter view offers Web viewing experiences (personal interaction phase).

The New Camera Projection System

The TinkerLamp hardware was also employed in the TileSearch setup, but only for rendering a projection. Multi-touch interactions were enabled by a Microsoft Kinect mounted on top of the mirror, replacing the infrared camera in the RaindropSearch. The Kinect has a distinct set of cameras with depth-sensing capability, which introduces the possibility of building a more flexible touch screen. The depth camera distinguishes between objects based on their distances from the sensor using triangulation and trigonometry. The OpenNI SDK [3] was used

[3] http://structure.io/openni

to retrieve depth information from the sensor, and the OpenCV framework [4] was used to detect fingers from the camera image frames. The TileSearch was able to detect and track multiple finger touches from multiple participants, thus enabling more intuitive user interactions.

Search View

The search view has two variants, it either presents pictorial stimuli with image search results (Figure 6.8(a)) or textual stimuli with Wikipedia search results (Figure 6.8(b)). The images were displayed in a 4 by 3 grid, whereas the Wikipedia snapshots were shown in 3 by 2. Fewer Wikipedia results were displayed due to hardware constraints: Texts were no longer legible on the surface, if more than 6 results were shown in parallel.

The system concatenated every N consecutive conversational nouns as a search query into Microsoft Bing [5]. The top M image results or Wikipedia results were then displayed as contextual information scents. Microsoft Bing search engine was used because it offered an API that was more attractive in terms of cost and integratability with the system at the time when the TileSearch was developed. Each new search was made at least 5 seconds after the previous results were shown, and the timing was controlled programmatically. When new results arrived, the display was updated, from left to right and top to down. After a few tests, we found when $N = M = 3$, the updates on the display more or less kept pace with the conversation. The system then works as follows: at the beginning, 3 consecutive nouns form a query term, and then this query term was fed into Bing. Then the first 3 empty tiles in the grid are updated with the top 3 returned results. The next returned search results update the 3 empty tiles that follow, so on and so forth. Once all the tiles have been used, new updates override the oldest tiles, according to a first-in-first-out policy.

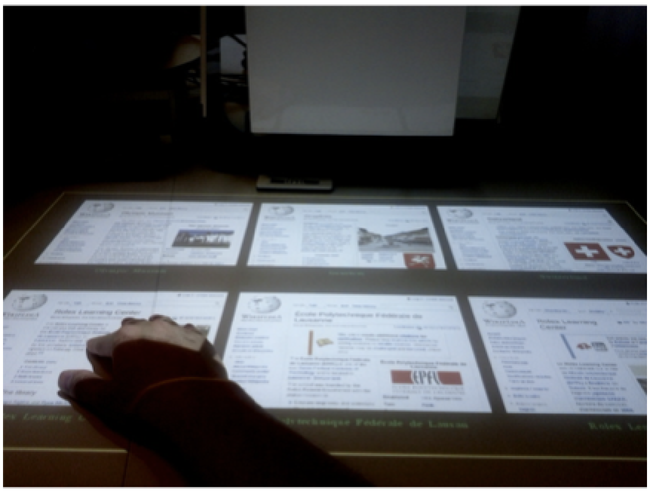
The design of the information scents finds itself in the design space as illustrated in Figure 6.7, where the TileSearch distincts from its predecessor only in the information capacity dimension, i.e. images and Wikipedia pages contain medium and high information capacity respectively.

Web View

Once a specific tile in the search view is selected by hand, the TileSearch interface switches to the Web view (Figure 6.8(c)), which contains a Web browser widget automatically loads the selected result. The widget is fixed on the scene, i.e. it is not zoomable or rotatable, but the Webpage content can be scaled and scrolled with multi-touch gestures. The browser supports forward and backward navigation by tapping the corresponding arrow buttons. There is another rectangular arrow button on the short edge of the interface to allow manually switching between the search view and the Web view. Note that searching with a keyboard as

[4] http://opencv.org/
[5] http://www.bing.com/

(a) Pictorial Stimuli View with Image Search Results



(b) Textual Stimuli View with Wikipedia Search Results

(c) Browser View with Touch Interaction Support

Figure 6.8: TileSearch tabletop interface

in the RaindropSearch was not supported by the TileSearch, since the focus was on the design of the information scents for this study.

6.2.3 Experiment

We conducted a controlled experiment with subjects in the lab working on a collaborative brainstorming task. The conversational words were captured in a Wizard-of-OZ style. Our experiment aimed to answer the research questions in Section 6.2.1.

Tasks and Participants

We recruited 12 subjects (4 females) from our colleagues. They aged between 23 and 50 years old, with background in Engineering (8), Social Science (2) and others (2). The participants were divided into 4 groups, each containing 3 members. The experiment had a within-subject design. Each group had to complete two brainstorming tasks, one with contextual information scents in the form of image search results and the other in Wikipedia results. The FUTURE CAR task required the groups to brainstorm futuristic features a car would have in 20 years; the FUTURE HOME task was about envisioning how intelligent our home would become in 20 years. In order to minimize the carry-over effect in a within-subject design, the order of tasks and information scent conditions were counterbalanced.

Procedure

According to Kunifuji et al. (2007), a brainstorming session can be organized in two phases, i.e. a divergent thinking phase and a convergent thinking phase. In the divergent thinking phase, brainstormers simply produce a large quantity of ideas with no judgment in terms of quality. In the convergent phase, they are allowed to validate the collected ideas based on some criteria. Our brainstorming experiments follow these two phases.

Before the start of each study session, the recruited groups were given a brief introduction of the two tasks as well as instructions about the two brainstorming phases (cf. Appendix C.1). They had 5 minutes to get familiar with the interactions offered by the TileSearch system. Each participant was then given a pen and a set of sticky notes for writing down their ideas during the brainstorming. They first brainstormed one task topic with either the image or Wikipedia interface, and the other condition would follow afterwards. The orders of the tasks and designs exposed to the subjects, as indicated earlier, were counterbalanced. Subjects were given 15 minutes to work on each task. In the divergent phase, they had 7 minutes to develop ideas. Participants who first proposed a distinct idea should write it down on a piece of sticky note and stick it on the table. The end of a divergent thinking phase was signaled by the experimenter, then the group had 8 minutes in the convergent thinking phase where they must justify and shortlist the ideas according to whether or not their visions were likely to be realized within 20 years. The same procedure was repeated for the second task.

The TileSearch interface was projected on the table since the start of the experiment. Our subjects were not obliged to use it, so that we could observe how the system would trigger serendipitous interactions.

Data Collection

In the study, data was collected mainly with questionnaires with interaction logs. The experimenter also took observation notes and conducted semi-structured interviews after the study sessions.

Questionnaires

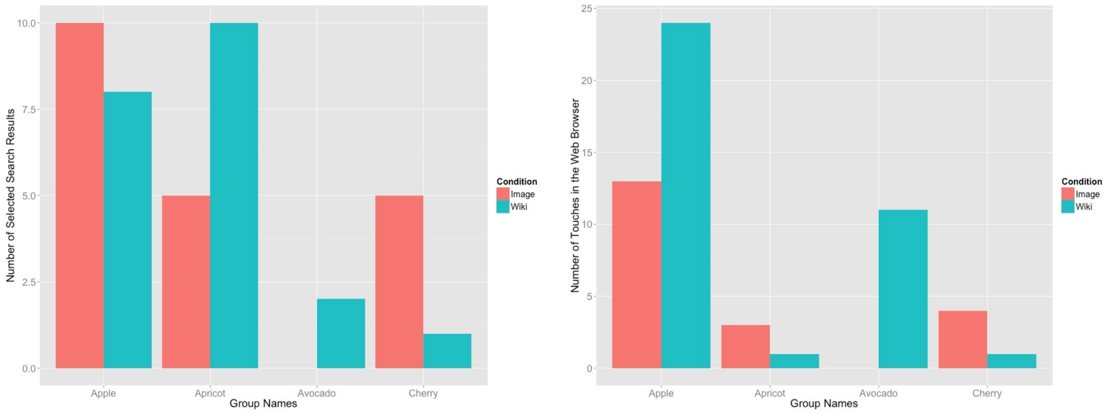
Before the experiment, each subject was asked to fill in a pre-experiment questionnaire (cf. Appendix C.2), in which we asked about demographics, personalities as well as experiences with tabletop systems and searching in meetings etc. After the first task was completed, each group was asked to fill in an intermediate questionnaire (cf. Appendix C.3), which gathered information regarding the number of ideas contributed, subjective perception of collaborations, satisfaction with the information scents etc. Finally, they filled in a post-experiment questionnaire (cf. Appendix C.4) at the end of the second task. In addition to the questions we asked in the intermediate questionnaire, in the final one we also asked them to rate the interactions offered by the system, and to compare the two design conditions.

The subjective perception of group collaborations and the information scents were evaluated on the following aspects:

- **Perceived closeness of work**. We asked each participant to rate how closely s/he worked with other group members to accomplish the task
- **Perceived dominance of work**. We asked each participant how they agree the brainstorming was dominated by someone.
- **Perceived effectiveness of communications**. We asked each participant to rate how effectively the members of the group communicated with each other.
- **Perceived usefulness in the divergent phase**. We asked each participant to rate the usefulness of the system for the divergent phase of the task.
- **Perceived usefulness in the convergent phase**. We asked each participant to rate the usefulness of the system for the convergent phase of the task.

Interaction Logs

Both user-initiated and machine-initiated behaviors were logged, including the Web addresses of the information scents presented on the tabletop, the Web addresses that were selected for further interactions, and the number of touch interactions inside the Web browser.



(a) Selection of Search Results

(b) Interactions on the Web Browser

Figure 6.9: Frequency of TileSearch interactions

6.2.4 Result

Based on the data collected from both the interaction logs, questionnaires, observation notes and interviews, this section reports the evaluation results of the TileSearch system by five themes: (1) TileSearch as serendipity inducer; (2) TileSearch as group facilitation; (3) Tile-Search as idea inspirer; (4) TileSearch as idea validator; and (5) effects of pictorial and textual information scents.

TileSearch as Serendipity Inducer

Over the 8 brainstorming sessions, the TileSearch made on average 72.4 Web searches ($\sigma = 9.5$) in Bing, and 152.1 results ($\sigma = 37.1$) were shown on the table as contextual information scents. The first question that came up before we dig into more detailed analysis is: were the presented search results actually used by the participants? This question is concerned with the role of the TileSearch as serendipity inducer.

As presented in the system description, users use TileSearch in two aspects, i.e. selecting search results and browsing the corresponding Webpages. On the users' side, the frequency of these two kinds of interactions were illustrated in Figure 6.9. We found only one group, when exposed with the system in the image condition, did not use the system. All other groups had used the system, but with large variability in terms of frequency. On average, the automatic search results shown on the tabletop display were selected 5.1 times ($\sigma = 3.9$) during each task, which further led to 7.1 touch interactions ($\sigma = 8.3$) in the Web browser. The standard deviations for both measures are relatively large compared to the mean value, indicating that the general usage was somewhat opportunistic.

In addition to the overview of system usage presented before, we are especially interested in the number of selected search results in the two brainstorming phases, because it measures

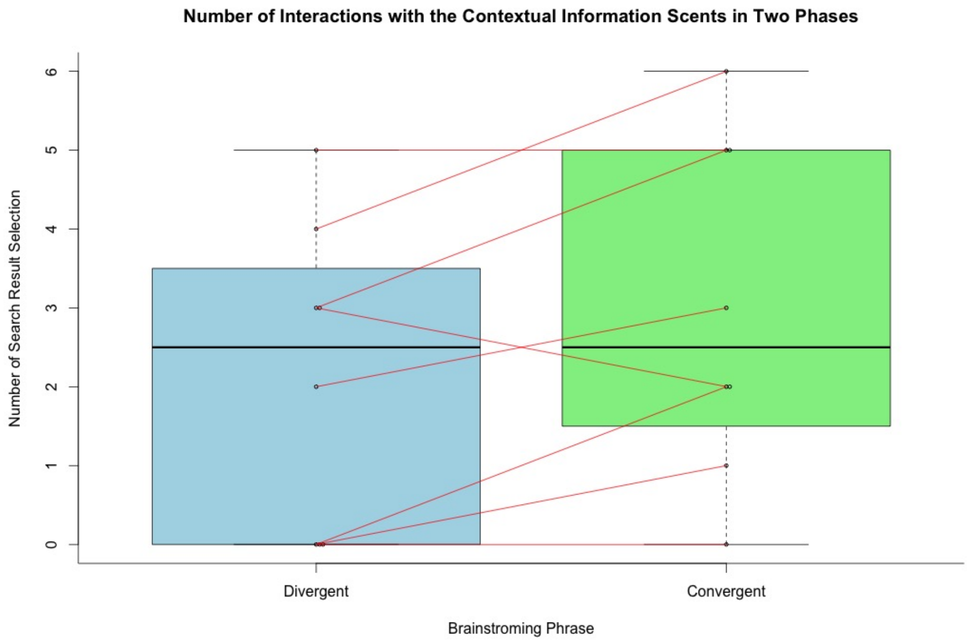


Figure 6.10: TileSearch interactions in divergent and convergent phases. For each of the eight brainstorming sessions, the red ling segment connects the number of selected contextual information scents during the divergent and convergent thinking phases respectively.

how good the information scents are as serendipity inducer. As shown in Figure 6.10, in most sessions, the number of interactions with the contextual information scents in the convergent thinking phase was higher than that in the divergent thinking phase. With paired t test, we found the difference is significant ($t(7)=2.2$, $p=.06$, Cohen's $d = 0.78$) at $\alpha = .1$ due to small sample size. One possible explanation is that the participants explicitly need information for judging their ideas in the convergent thinking phase, so they had prepared mind for encountering information.

TileSearch as Group Facilitator

The examination of the TileSearch's role as group facilitator is based on three subjective ratings collected from the questionnaires, i.e. perceived closeness of work, perceived dominance (imbalance) of work and perceived effectiveness of communication. Our research explores the relationship between these ratings and the interaction frequencies. However, the ratings were collected on an individual basis whereas the interaction frequency was logged in the unit of group. We therefore take the average of the individual subjective ratings per group and examine the relationship between the group average ratings and the group interaction frequency.

Three mixed-effect multiple regression models, each with one subjective rating as outcome

Table 6.1: Effects of TileSearch interaction on perceived group performance

Group Performance Variable	Number of Selected Search Results	Number of Browser Interactions	R-squared (marginal/conditional)
perceived closeness of work	$\beta = -0.01, p = 0.62$	$\beta = -0.01, p = 0.62$	0.04/0.1
perceived dominance of work	$\beta = -0.28, p < 0.05$ *	$\beta = -0.07, p = 0.11$	0.35/0.99
perceived effectiveness of communication	$\beta = -0.10, p < 0.01$ *	$\beta = -0.09, p < 0.0005$ *	0.96/0.97

variable, the frequency of search result selections and browser interactions as two predictors, the group, condition and task as random effects, were built. The results are illustrated in Table 6.1.

Generally speaking, the TileSearch did not significantly affect the closeness of work. We find that the number of interactions with the system has a significant negative relationship with the effectiveness of communication. This result perhaps indicates that the participants attempted to use TileSearch more when they perceived their communication to be less effective. Another interesting finding is the number of interactions also negatively relate to perceived dominance of work, indicating that more interactions with the TileSearch probably reduced the imbalance of collaboration. We may explain this finding as follows: When more search results were selected and examined, each participant had more chances to express their own opinions towards the information, which potentially led to less dominance.

TileSearch as Idea Inspirer in the Divergent Thinking Phase

As presented before, interactions with TileSearch during the divergent thinking phase were infrequent. Sometimes the information scents sparked curiosity which was pertinent to the task, a subject commented: "*I came cross an interesting image, so I just wanted to see more information, but it actually has nothing to do with the task*". According to our observation, some users did attempted to look for information on the display when they ran out of ideas. For example, we observed a subject in one group said: "*...what else ideas....ah!...let's see what the display suggests...* ", but such behaviors were not frequent. However, did the subjects' occasional interactions with the information scents have an effect on inspiring ideas?

In fact, when the number of ideas generated per group per task is predicted, it is found that the number of actions on the Web browser is a significant predictor with negative effect ($\beta = -.45, p = .01$). The effect of the number of selected search results is positive, but not significant ($\beta = .60, p = .16$). The statistics were computed from a mixed-effect multiple regression model where the group, condition and task were modeled as random effects. The overall model fit in terms of R-squared is .90, and the marginal R-squared contributed by the fixed effects is .36. The negative significant effect of Web interaction frequency indicates that more engagement with Webpages is associated with fewer generated ideas. A possible explanation could be the subjects Webpage browsing took time, which could otherwise be spent on creative thinking. On a 7-point Likert scale, the subjects gave below average ratings regarding the system's function as idea inspirer for group ($\mu = 2.33, \sigma = 1.27$) and its overall usefulness in the divergent thinking phase ($\mu = 2.63, \sigma = 1.50$).

The above results are based on the total number of ideas per groups. We are also interested to examine the factors that affect individual contributions. In fact, users also held negative views towards the system as idea inspirer for individuals ($\mu = 2.21, \sigma = 1.38$). Discussions in the previous section have shown the TileSearch had significant effect in facilitating some aspects of group collaboration, e.g. the effective communication.

Do these group performance variables correlate with individual idea contributions? We built a mixed-effect multiple-regression model to predict the number of ideas contributed per person. All the previous discussed subjective group factors were modeled as fixed terms. The subject nested in group, condition and task were modeled as random effects. With backward elimination, only the closeness of work ($\beta = -1.54, p < .001$) and effectiveness of communication ($\beta = .87, p < .05$) were left in the model as significant predictors. The overall model fit in terms of R-squared is .52, and the marginal R-squared contributed by the fixed effects is .41. Note that the closeness of work is a negative predictor, indicating the closer an individual worked together with others, the fewer ideas s/he would contribute. This result echoes an old finding in brainstorming research: Taylor et al. (1958) found that group participation might inhibit creative thinking. On the other hand, more effective communication is correlated with more contributed ideas. Recall the finding in the preceding section that more interactions with the TileSearch correlate with reduced average effectiveness of communication in a group. This may potentially explain why interacting with the information scents was not considered as useful by the subjects in the divergent thinking phase. It should also be noted that the subjective ratings about group performance were not reported specifically for the divergent thinking phase, but for the overall brainstorming process, so the ratings may be influenced by the convergent thinking phase as well.

TileSearch as Idea Validator in the Convergent Thinking Phase

In the given task, the criteria for idea validation in the convergent thinking phase is that the ideas should have the potential to be realized in 20 years. This process often led to reducing ideas with collaborative effort. On average 3.6 ideas ($\sigma = 1.9$) per task per group were excluded in this phase. When the number of dropped ideas is predicted, it is found that the number of search result selections is a significant predictor ($\beta = -.67, p < .05$), but the number of actions on the Web browser is not ($\beta = -.13, p = .37$). The statistics were computed from a mixed-effect multiple regression model where the group, condition and task were modeled as random effects. The overall model fit in terms of R-squared is .62, and the marginal R-squared contributed by the fixed effects is .57. Note that the significant predictor has a negative effect. That is to say, the more search results on the display were selected for examination, the less ideas were excluded. However, it is difficult to interpret the system's effectiveness from these significant effects alone, since interacting with the TileSearch may either confirm or reject an idea. That is to say, with more interaction, the effect may be good because the ideas were supported rather than ruled out, but it can also be bad due to the system's ineffectiveness in validating ideas. From users' subjective feedback, the perceived usefulness of the system in the divergent thinking phase was low ($\mu = 2.63, \sigma = 1.28$), perhaps indicating the system's inability as idea validator. In fact, the subjects had very clear search goals in order to validate an idea. When they found the search results did not match their expectations, negative attitudes towards the system were developed.

Effects of Pictorial and Textual Information Scents

In this section we compare the effects of pictorial (image search results) and textual (Wikipedia snapshots) information scents as serendipity inducer, group facilitator, idea inspirer and validator. The variables we employ to access these effects are the same as those presented in the corresponding sections. We built mixed-effect ANOVA models, where the users nested in group and the tasks were modeled as random effects, to compare the effect of the image and Wikipedia condition on the aforementioned variables. The significances are shown in Table 6.2.

Table 6.2: Comparison of effects of image and Wikipedia information scents

Role of InfoTiles	**Variables**	**Image**	**Wikipedia**	**p-value** ($\alpha = .1$)
As serendipity inducer	number of search result selection in divergent phase	2.25	2.00	0.70
	number of search result selection in convergent phase	2.75	3.25	0.29
As group facilitator	perceived closeness of work	5.67	6.08	**0.05***
	perceived dominance of work	4.5	3.58	**0.06***
	perceived effectiveness of communication	6.17	5.91	0.30
As idea inspirer and validator	perceived usefulness in divergent phase	2.67	2.58	0.86
	perceived usefulness in convergent phase	2.42	2.83	0.29
	number of individual ideas	4.17	3.41	0.32
	number of group ideas	12.5	10.25	0.34
	number of dropped ideas	3.8	3.5	0.83
	number of finalized group ideas	8.75	6.75	0.12

Our first hypothesis is that the image display attracts more serendipitous encounters in the divergent phase and less in the convergent phase, since the images are immediately comprehensible to trigger connections whereas the Wikipedia articles carry more information for idea validation. It turned out that interaction frequency in both conditions were low, and we found no statistical evidence to support the hypothesis.

It is interesting to find that the Wikipedia condition induced significantly higher closeness and less dominance of work, compared to the image condition. One possible explanation is that the Wikipedia pages contain elicited information about well-defined topics, which was better for anchoring group discussions. In contrast, image results were mostly taken from random Websites, which did not always offer credible information. In terms of perceived effectiveness of communication, both conditions did not show significant differences. This might be partly due to the overall feeling of low relevance of the presented results, which many subjects had reported.

Regarding the effects for idea inspiration, unfortunately both conditions received low ratings from the subjects, and the number of ideas generated did not differ significantly by the two conditions.

Summary of Negative Feedback

We summarize the negative feedback collected in the questionnaires regarding interacting with the TileSearch. The problems can be classified into three categories:

- **Time limits** were reported to have constrained the chances for the subjects to use the system. Two subjects explicitly reported that they had no "time" to look at the display, intentionally or not, for "developing ideas" during the brainstorming. Within limited time, they concentrated on interacting with other group members. In contrast, the Idea Exapander study (Wang et al., 2010) did not report time limits for brainstorming, which could partially explain the reported effectiveness of the system.
- **Distractions** of the display were reported by three subjects. In contrast, no one reported the RaindropSearch to be distracting, though it employed more disturbing animations. This was because the RaindropSearch was completely ignored by participants. For the TileSearch system, the information scents actually fostered serendipitous encountering, but the problem is that the subjects were distracted by the fact that most information contributed little to their task performance.
- **Low relevance** was reported by four subjects. This was especially the case when the system was used in the Wikipedia condition during the convergent thinking phase. When certain information was strongly required, the subjects were observed to speak out the formulated keywords and expected specific results to appear on the table. Unfortunately the displayed results did not always meet their expectations due to the "noise" in conversation and the randomness introduced in the query formulation process.

In reality, these factors were intertwined with each other. The system always attempted to offer information scents pertinent to the "current conversation". In contrast, the users did not always need information. As a result, most of the presented information were perceived as lowly relevant distractions. Coupled with the time-limit issue, it is no wonder that users felt difficult to elicit potentially useful information.

6.2.5 Lessons Learned

In the preceding section, we presented a summative evaluation of the TileSearch system with several performance benchmarks to gauge its effect in inducing serendipity, facilitating group discussion, inspiring and validating ideas in collaborative brainstorming activities. This section aims at summarizing the lessons learned from the evaluation results.

- **Contextual information scents in the form of search results**. Compared to conversational words that need to be built as queries, displaying search results triggered by conversation has shown to be less calm but more trigger-rich. This is manifested by more frequent interactions with the information scents in the TileSearch. However, constantly updating the information scents in response to conversations seemed to unavoidably induce distractions. Interacting with TileSearch are shown to have improved imbalance of work, but it negatively affects the number of ideas generated in the divergent thinking phase. In the convergent thinking phase, the users preferred to have search tools that allow customizable search, because the information needs were more inelastic.

- **Pictorial versus textual information scents**. Our subjects did not express strong preferences towards either design. There is also a limitation in the design. The Wikipedia snapshots were not always completely textual, and they sometimes contained images as well. However, the Wikipedia articles offer more credible information than random sites on the Web, which probably explained why the TileSearch with Wikipedia search results resulted in more closeness and less dominance of work.

- **Increase relevance**. Low relevance of the presented information was reported by most participants. Relevance is especially difficult to maintain if the system has to constantly react to group conversations. However, if the system is not very responsive, then it would probably not be able to deliver timely information. Perhaps a better solution is either employing a better strategy to filter keywords, or deriving context from other more reliable sources.

In summary, augmenting collaborative activities with information scents generated by conversation is a non-trivial matter. Our challenges are not limited to the difficulty in speech recognition in natural conversations, but also include intelligently filtering key words in the conversations and using these words for retrieving relevant results. The two prototypes presented in this chapter were our first attempts and a few insights were drawn from the studies. In addition, we recruited only members from our lab as subjects. In the next chapter, we present the design of information scents generated from the user interactions with a collaboration tool, which is expected to yield more focused context than conversations. This possibly offers better opportunity for retrieving more relevant results for supporting elastic information needs arisen in the collaborative activity.

7 Exploring Interaction Context

The preceding chapter introduces two initial explorations of contextual information scents generated from conversation. The RaindropSearch explored the design of information scents as conversational words, and the TileSearch explored information scents in the form of pictorial and textual search results. The second exploration was promising, at least in the aspect of reducing imbalanced work. One of the main lessons we learned from the study is the need to improve the relevance of information scents. This can potentially be achieved with more "intelligent" context elicitation, in terms of word filtering or query formulation etc. In addition, more reliable activity context rather than conversation may also be utilized. This chapter presents the MeetHub Search, which explored the design of information scents generated from interaction context, i.e. user interactions within a groupware. Section 7.1 presents the system design and technical setup. Section 7.2 discusses an exploratory user study to understand the possibilities and constraints of the contextual information scents.

7.1 Exploration: MeetHub Search

The MeetHub Search aimed at exploring different possibilities in the design space for creating information scents from interaction context in collaborative activities. It was integrated as a component in a groupware called MeetHub, which was developed under joint efforts between me and two other colleagues. The goal of the research project was to understand the dynamics in computer supported group collaboration, including various aspects of the software's interactivity, users' awareness of information needs and time, etc. This section starts with a brief introduction to the functionalities that the MeetHub offered as well as contributions made by each individual contributor. This is followed by a detailed explanation of the MeetHub Search component, which is the focus of this chapter.

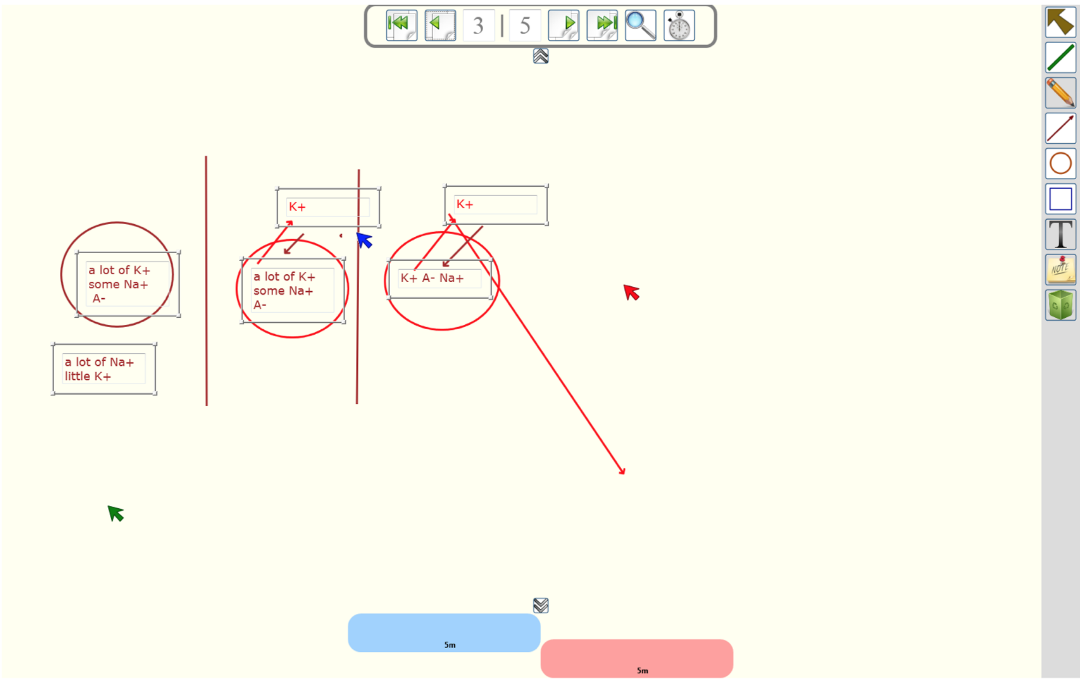
Figure 7.1: MeetHub system setup

7.1.1 MeetHub: Setup and Components

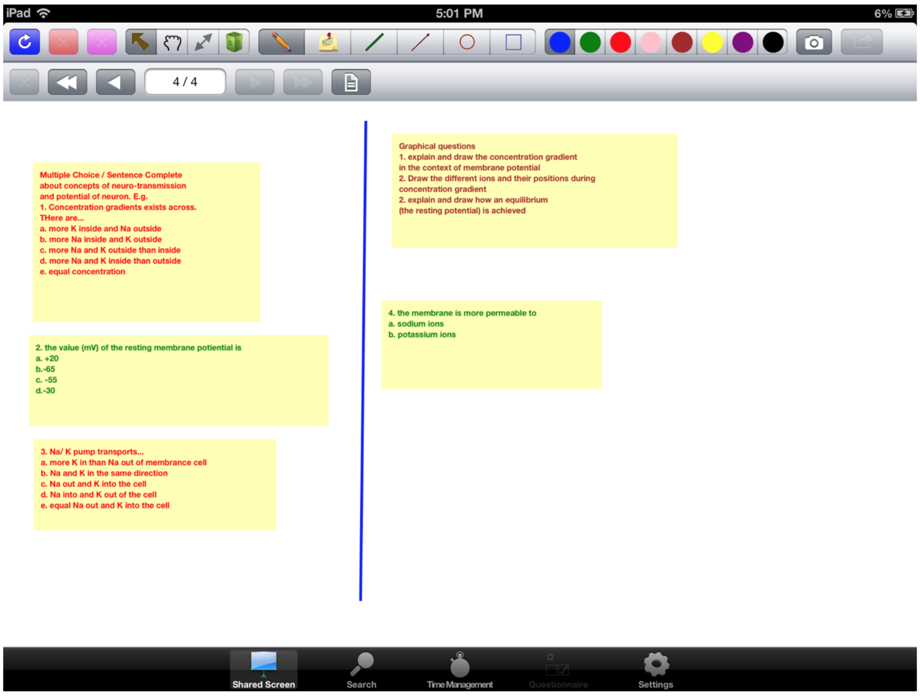
The MeetHub is a groupware that includes interaction elements in both public and private spaces during group work. It also allows smooth transitions between these spaces.

System Setup

A typical MeetHub system can be configured with a meeting table, a PC with connected with multiple keyboards and mice and a few mobile devices. Figure 7.1 illustrates our experimental setup: A special amoeba-shaped table was designed to host small discussion groups of up to five participants. A wall-sized display rendered by a ceiling-mounted projector serviced as the public display of the shared workspace. Five pairs of keyboards and mice are connected to the MeetHub system as input devices for the shared workspace. In addition, iPads with the mobile version of the MeetHub interface were available for the group users. More details about the features and interaction possibilities of the system are described in the upcoming sections.



(a) Shared workspace on the wall-mounted display (PC interface)



(b) Shared workspace on the tablet (iPad interface)

Figure 7.2: MeetHub shared workspace

Shared Workspace

The shared workspace, which anchors most of the group interactions, was created by one of my fellow colleagues. The workspace is shared among all client devices connected to the PC server. Figure 7.2 illustrates an example of the shared workspace representations on the wall display (PC interface) and on the tablet (tablet interface) respectively. The PC interface was implemented with WPF and .NET technologies for Windows computers, and the mobile client was implemented in Objective C for Apple iOS devices. The shared workspace is featured with the following characteristics:

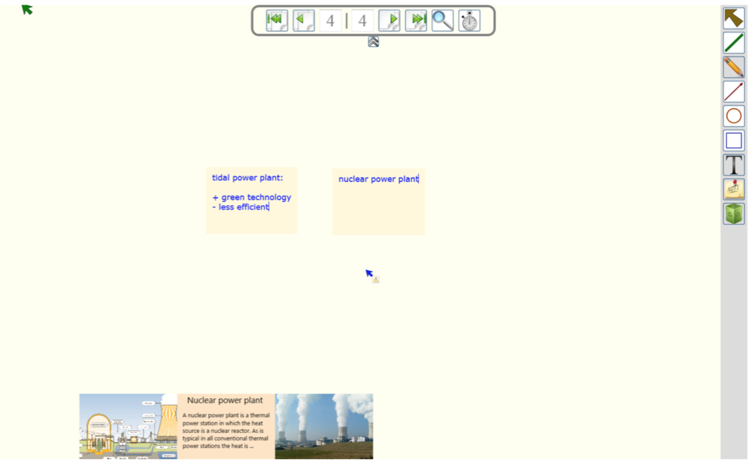
- **Tool Collection**. A palette of tools lie along the right-hand edge of the workspace on the PC interface and the top edge of the tablet interface, respectively. It offers a collection of tools for drawing and annotating. User can issue actions with mouse and keyboard (PC interface) or with touch gestures (tablet interface). For example, a user can create a post-it note or a text box and type on it. Freehand tools, shape tools are available for making rectangles, circles and lines. Selection and trash tools allow selecting an object to move or delete respectively. Additionally, page management tools located on the top of both interface versions let users create new workspaces.
- **Identification of Input Devices**. Each trio of a mouse, a keyboard and an iPad share a unique ID (hard-coded), which is represented by the color associated with the mouse cursor. Objects,regardless of text or shape, created by any of the aforementioned devices with the same ID, are shown in the same color.
- **Synchronized Content**. The content in the workspace, regardless of its representations in the PC and in the iPad, is always synchronized among all the devices via polling. Whenever a user issues a change in the workspace, the change is delegated through the PC server to poll all other devices, which then updates the interfaces accordingly.

Time Awareness Component

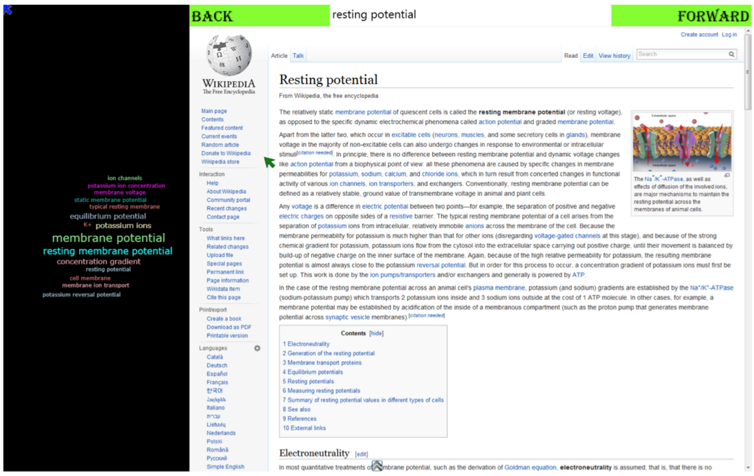
MeetHub also employs a time tool which was developed by another fellow colleague to study time awareness in meetings. With the time tool, users can plan different phases before the start of a meeting. Time management functionality was implemented only on the iPad interface with a dedicated panel for setting meeting durations and phases, whereas time awareness was provided on the PC interface. At the bottom of Figure 7.1(a), the blue/red widget is an example of time notification for awareness. It pops out and blinks for signaling the end of a meeting phase.

7.1.2 MeetHub Search

The research and development of the search component, i.e. the MeetHub Search, were my major contributions to the joint project. The lower two corners on the MeetHub PC interface



(a) “Marquee” as information scents for the shared workspace



(b) “WordCloud” as information scents for the PC Web browser

(c) “Query List” as information scents for the iPad Web browser

Figure 7.3: Contextual information scents in MeetHub

are "hot corners". Moving a cursor to the lower left corner brings out a Web browser, like the one in Figure 7.3(b). Users can type their own keywords in the "search box" and retrieve results from Google. The Web browser is not featured with an address bar, so Web browsing must be started with a Google search. Moving a cursor to the lower right corner brings back the shared workspace. On the iPad interface, there is a dedicated panel reserved for Web search. Again, users can only type in query terms rather than full Web addresses (see Figure 7.3(c)). The search panel can be seen as a personal interaction space, since individual user can search independently on his/her own iPad. A user can share a link to the public Web browser on the PC interface by pressing the button on the top right of the search interface.

In addition, the MeetHub Search provides three types of contextual information scents guiding users to potentially useful information. They are presented in different interaction processes with the MeetHub.

"Marquee" in the shared workspace (PC interface)

In the preceding chapter, we presented the TileSearch system which used image or Wikipedia search results as contextual information scents, which were returned by querying terms captured from conversational words. The study did not confirm which was better in many aspects. The Marquee in the MeetHub system did not abandon either type of design. Instead, each Marquee presents a combination of the two kinds as information scents. Automatic search results were returned by searching query terms built from various "text objects" in the MeetHub system. Figure 7.3(a) illustrates an example of a Marquee, which consists of two image and one Wikipedia thumbnails scrolling from right to left. We tested several time durations for the scrolling animation, from 10 seconds to 1 minute. Finally we chose 25 seconds so that the Marquees were neither too fast nor too slow for users to recognize. A user can click a specific thumbnail during its "travel" to view more details in the Web browser.

How did the system build query terms? In fact, building search queries from context is usually done in two phases. First, keywords are extracted from contexts (i.e. keywords extraction phase). Second, the keywords are combined with specific rules to form queries (i.e. query building phase). For the keyword extraction phase, both the RaindropSearch and the TileSearch in Chapter 6 extracted nouns as keywords. In the studies of the two systems, our subjects expressed dissatisfactions with the quality and relevance of the resulting information scents, which might partially attribute to the simple keyword extraction approach. As discussed in the previous chapter, keyword extraction is a nontrivial task, which can require sophisticated statistical models and natural language processing technology. Despite all these, the effectiveness may also vary from context to context. As a result, there is no gold standard for it. Jean-Louis et al. (2014) presented a comparative study on several online semantic annotators available on the Internet, and the AlchemyAPI[1] performed well in many tests in the study. It provides a RESTful API (Web service) that can be easily integrated in existing applica-

[1] http://www.alchemyapi.com/products/alcheanguakeyword-extrion

tions. Therefore, the MeetHub Search employed the AlchemyAPI to find and rank keywords from the text on the shared workspace. Not only individual words but also phrases can be captured as keywords. As an example, if the sentence "Every neuron has an electrical voltage on both sides of the membrane that is called the membrane potential" is fed into the Alchemy API, the extracted keywords, from high to low relevance, are "membrane potential","electrical voltage", "neuron" and "sides".

In order to generate information scents from interaction context, we must first identify the user-generated objects that may contain contextual information, in the form of text. In fact, there are three types of such objects in the MeetHub: (1) text typed into the text-editing widgets such as text boxes or notes in the shared workspace (2) query terms typed directly in the Web browser search box (3) text in the opened Webpages. The Webpages are not created by the users, but they are presented in the system as a result of user interactions. In addition, intensity and transactivity of the interactions may also be considered as interaction context. We implemented 3 approaches to build queries from the keywords captured in the aforementioned text objects. Searching with the queries then yield results that lead to the construction of Marquees.

Typing-triggered Approach (TA) The typing-triggered approach (TA) only considers texts typed in the text box and note widgets as interaction context. The system detects delimiters such as "?","!",";",".", as well as the "return" keystroke. These symbols usually signal the end of a semantically meaningful text segment, which is then fed into the Alchemy Web service for keyword extraction. Time limit is also considered. If none of the aforementioned delimiters were hit by users and the text being edited is idle for 15 seconds, then the previously typed content is also fed for keyword extraction. The keyword extraction is not performed on the complete text written in a text box or note widget. Instead, each time the system compares the changes between the current text and its previous state when last keyword extraction was performed. Only the changed text is used.

Usually one or more keywords or phrases are extracted, but there are also occasions where no keywords are extracted, especially when the text segment is composed of only stop words. Each keywords forms a query of its own and is fed into the search engine to retrieve results. The top 2 image results and the first Wikipedia results construct a Marquee.

Combinational Approach (CA)

The advantage of the TA is the extraction of instant interaction context from user-generated text, but the result may suffer from the absence of global context. In the combinational approach (CA), search queries combine the keywords extracted by the TA with a list of N (maximum = 5) pre-selected keywords that defines the task context to be performed by users: Whenever a keyword or phrase is extracted, it is mixed with each combination of $\binom{N}{1}$ and $\binom{N}{2}$ of the pre-selected words to form a search query. As a result, each keyword generates a maximum of $\binom{N}{1} + \binom{N}{2}$ marquees, if all the Web searches return valid results.

Weighted-selection Approach (WA)

One disadvantage of the CA approach could be the generation of potentially redundant query terms by an excess combination of keywords and global contextual words. The latter may not be always useful for building search queries at all moments. The weighted-selection approach (WA) collects keywords generated from three types of text objects in the MeetHub system, together with the global contextual words. The WA assigns weights to keywords and builds query terms with highly weighted words.

Specifically, the pre-selected global contextual words are assigned with the largest initial weight: 4.0. The following types of terms are weighted as 3.5 because they are resulted from interactions with shared awareness: (1) search queries explicitly specified by users in the public Web browser on the wall display; (2) terms in the Wordcloud that are clicked by users (will be described later); (3) queries that are used to compose a "consumed" Marquee. The terms associated with the following situations are weighted as 3.0 because they are resulted from private interactions: (1) a user searched a term on his/her own iPad; (2) a user clicked a suggested term in the Querylist (will be described later); (3) a user shares a link from the iPad to the wall display. In addition, keywords extracted from the text-editing widgets (like in the TA) have initial weights of 2.

The weight of a keyword does not stay unchanged. Every time a new Marquee is generated, the weights of all words decrease by 0.5. If a word's weight reduces to zero, then it is removed. Additionally, intensity and transactivity of keyword contributions are considered. Intensity refers to the recurrence of a certain word or phrase. Transactivity means a keyword contributed by one user is repeated by another user. If a term is captured again from the same user, then its weight increase by half. In case of transactivity, the weight doubles. At the beginning, the system composes search queries with a single word with the highest weight and generates a Marquee accordingly. After the Marquee scrolls out of scene, the next search takes 2 words with largest weights. The process continues by taking one more word each time to compose a query, until (1) no results can be returned from the search engine; (2) the number of selection exceeds the number of positively weighted words available. Under these two circumstances, the number of selection is reset to one, and the above described process repeats.

Expected Usage and Benefits

When group users type into the text-editing widgets in the shared workspace, we assume their goal is to note down certain things or to express ideas. Usually there is no immediate explicit need for searching information at the time of typing. The marquees aim to capture users' text interactions and present automatic search results as contextual information scents, which are expected to cue for latent information needs.

"Wordcloud" in the Web browser (PC interface)

In the MeetHub Search, we design a Wordcloud (Figure 7.3) alongside the Web browser on the wall display. When a user views a Web page, the Wordcloud displays a maximum of 15 most relevant keywords it contains with distinct colors. The keyword extraction is done via the AlchemyAPI, which returns a list of words or phrases with corresponding relevance scores. The sizes of the keywords are proportional to their relevant scores.

The Wordcloud serves as contextual information scents in a similar way as words in the RaindropSearch presented in Chapter 6: The words are immediately searchable as query terms. A user can click a specific keyword and the Web browser will navigate to the corresponding Google search result page.

Expected Usage and Benefits

When group users view a particular page, they may want to quickly grasp its main topic. The Wordcloud may serve for this purpose. The "cloud" layout offers instantly comprehensible situational information about the Webpage being viewed. In the meanwhile, certain key concepts in the page may spark additional information needs. The benefit is the convenience for searching potentially useful information with a mouse click.

"Querylist" in the Web browser (iPad interface)

Due to space constraints, the shared workspace in the iPad interface is not featured with Marquees. However, the search results carried by each Marquee, as well as the corresponding query terms are display in the Querylist in the Web search panel alongside the Web browser. As Figure 7.3 illustrates, the query term (i.e. Lady Gaga) is displayed as a list item. Clicking on the item leads to its expansion, which allows a user to find the images and Wikipedia articles that were previously displayed in the marquees. The Querylist creates contextual information scents as "search memories", so that a user can be aware of all the searches conducted by the MeetHub system in the past.

Expected Usage and Benefits

The Querylist is only visible on iPads. Suppose a user is focusing on interacting with the iPad, his/her attention is definitely away from the wall display. As a result, the user is not aware of what contextual information scents have been presented or what others have been doing in the meanwhile. The Querylist increases the awareness of the system's proactive behaviors and others' activity, which would probably be helpful for collaboration as well as for self-reflection.

7.1.3 MeetHub Search in the Design Space

With three types of contextual information scents and three search query building approaches, the MeetHub Search is concerned with multiple cells in each design space dimension (Figure

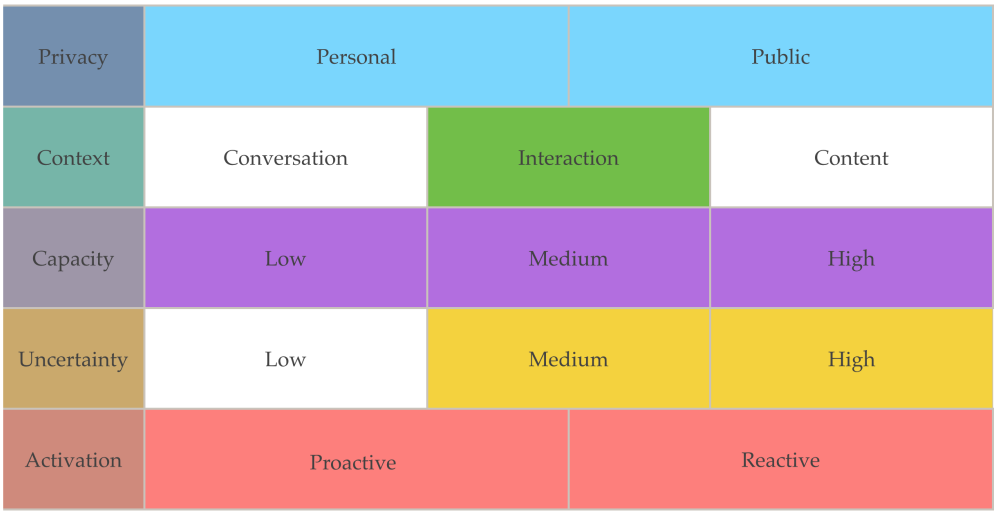


Figure 7.4: MeetHub Search in the design space

7.4). The Wordcloud works in reaction to Webpages loaded in the Web browser. It publicly displays only words and phrases, which carry low information capacity and medium uncertainty. The Marquees carry both medium (images) and high (texts in Wikipedia) capacity of information and work only on the public display. Information presented with both the TA and the CA induce relatively high uncertainty. The former is based on every changed text segment. The latter takes global context into account, but the combination algorithm induces randomness. The WA adopts a weighting algorithm so that the information uncertainty reduces to a medium level. For the *Activation* dimension, the marquees generated with the TA only reacts to changed text. In contrast, the CA proactively pushes new information scents based on combinations of task-relevant words, and the WA does it as long as positively weighted words remain.

7.2 MeetHub Search: Research Questions

As described previously, the MeetHub Search offers various information scents based on interaction context. Generally the research is more exploratory, with the following research questions to be addressed:

(1) What are the practical consequences of presenting the varying information scents generated from interaction context to collaboration groups?

In the design of the MeetHub Search, we attempted to make contextual information scents "ubiquitous" in the groupware environment. The users are expected to "smell" the scents, no matter they are collaborating on the public workspace, searching in the public browser, or working individually on the iPads. As a result, we expect the users to be influenced by the provided contextual information scents. We aim at studying how influential the technologies are for collaborative work.

(2) Does varying query building approaches make a difference?

The previous projects with conversation as context suffered from lowly relevant information scents, partially due to the simple keyword extraction and query building approach. In contrast, the user interaction context in the groupware is more reliable and we came up with different query-building variants that attempt to increase relevance. Will users perceived the difference? We need a study to explore this issue.

(3) Does assigning a dedicated information searcher make a difference?

In the previous studies, our group participants accomplished tasks mainly through discussions. As a result, their attentions were mostly not focused on the table, where the contextual information scents were displayed. In contrast, the MeetHub served as group's shared workspace, which in principle attracts most of the users' attentions. Still, the users may not gaze at the contextual information scents. We are interested to see if assigning an information searcher's role to a group participant to consciously monitor the information scents would influence the serendipitous encountering.

7.3 MeetHub Search: Study

In order to answer the research questions posed in the preceding section, we design and conducted a user study with recruited group participants working on different tasks.

7.3.1 Participants

We recruited 25 participants (5 females) aged between 21 and 28 years ($\mu = 23.6, \sigma = 1.5$) from our university for the user study of the MeetHub. The participants were students enrolled in a Master-level CSCW course offered by our lab, where taking part in the study was part of the course curriculum. The participants were partitioned into six freely formed groups, five of which had 4 members and one group had 5 students. Two groups consisted of all male students, and the other four groups were mixed. The user study was longitudinal, lasting for 4 weeks. Considering many features offered by the MeetHub system, the first week was reserved for the subjects to get familiar with the system as well as with each other. The groups were asked to complete a different task in each of the weeks thereafter.

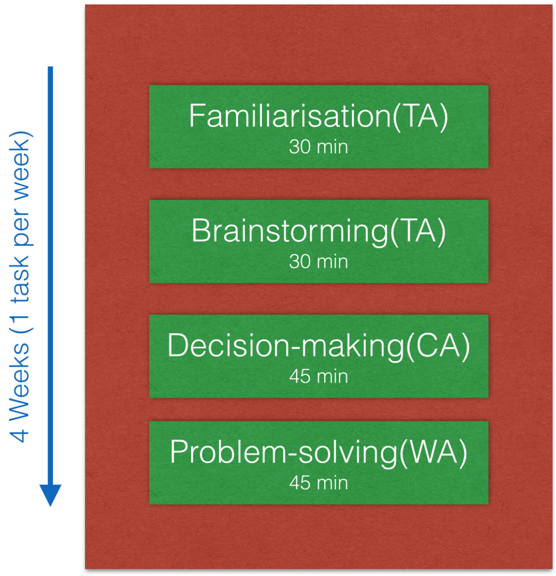


Figure 7.5: Task design for the MeetHub study

7.3.2 Tasks

Unlike the studies of the RaindropSearch or the TileSearch, which were only experimented with either a decision making or a brainstorming task, we planned to approximate the usefulness of the MeetHub systems with different collaborative tasks:

- **Brainstorming Task** (cf. Appendix D.1). We used a well-know task in brainstorming research, i.e. "Extra Thumb" (Taylor et al., 1958) for our study. This task required participants to brainstorm the pros and cons, if all humans were born with an extra thumb in the following year.

- **Decision-making Task** (cf. Appendix D.2). A similar task as the one in our study for the RaindropSearch was used in the study. The task was about making an energy plan for a given Chinese province, which was experiencing power shortage. This time the task was more complicated. The groups were first given basic statistics about the energy challenge, based on which they must make an estimate of power shortage in five-year time. Afterwards, they were required to make a ten-year energy plan for the given province to resolve the problem, regarding the types, numbers and locations of power technologies. They must also consider economic and environmental factors.

- **Problem-solving Task** (cf. Appendix D.3). The task we used was originally designed by Sangin et al. (2011). This task required group participants to first comprehend an instructional text about the "resting potential" in neurotransmision. Then they had to work together to draw a schematic representation for the generation process of resting membrane potential. In addition, the group participants played the role of teaching assistants for a neuroscience course, and they were required to prepare an assignment for their students.

The familiarization task, together with the three tasks described above, each was completed in a week. The order and duration of the tasks and the corresponding query-building approach employed in the MeetHub system are illustrated in Figure 7.5.

7.3.3 Condition

The only condition we manipulated in the study was the specification of roles in groups. Half of the groups worked without specified roles (NOROLE condition), though roles might emerge during the collaborative work. For the other half of the groups, the participants were assigned specific roles before the start of group work (ROLE condition). The assigned roles include the following: *group leader*, *time manager*, *content organizer*, and *information searcher*.

The group leader was responsible for facilitating group discussion, resolving conflicts and debriefing. The time manager's job was to keep track of the elapsed time with the time management tool so as to increase the groups' awareness of their work progress. The content organizer was responsible for organizing the visual objects created in the shared workspace and creating new pages if necessary to avoid cluttering. Finally, the information searcher was in charge of coordinating information search, as well as consciously monitoring the contextual information scents in the MeetHub system. Note that the information searcher did not have to be the only member to conduct searches. The roles only specified what they must do, but did not preventing others from doing the job. The roles were assigned in the very first week of the study, and all the subjects continued to play the same role during the rest of the weeks.

7.3.4 Procedures

In the familiarization week, all the subjects were asked to choose a seat around the amoeba-shaped table where they would sit throughout the whole study process. They were also asked to filled in a pre-experiment questionnaire regarding demographics and personality. This questionnaire was completed only once during the study period. In the following 3 weeks of formal experiments, the study procedure can be described below:

1. Before the start of each week's experiment, the groups were provided with necessary descriptions and relevant supporting materials of the task. Then, they had time to read and comprehend the task, as well as to ask questions to the experimenter.

2. Next, they were asked to discuss among the group to make the agenda of their meeting. As an outcome of the discussion, they were required to specify the planned number and durations of meeting phases with the time management tool.

3. Then, the meeting started. All tools and functionalities of the MeetHub system were available for use. One experimenter usually sat behind the group in the room to observe the study sessions.

4. When time was up, each subject was asked to fill in a post-experiment questionnaire regarding their experiences and feedback about the MeetHub system. The same questionnaire was repeated each week.

7.3.5 Data Collection

As mentioned in the study procedure, we had a pre-experiment questionnaire as well as weekly post-experiment questionnaires (cf. Appendix D.4) during the study. In addition, each group session was video recorded, and user interactions were automatically logged. All the collected data were for joint use by me and two other colleagues. For the analyses in this chapter, the video recordings are not concerned.

Questionnaires

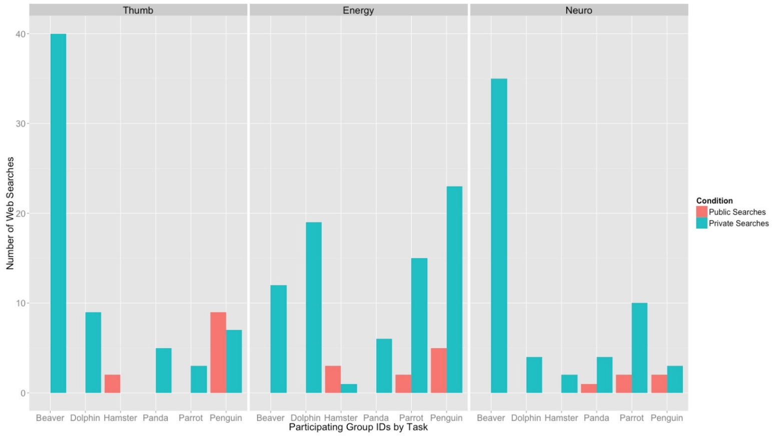
The questions we asked in the questionnaires were about different aspects of the MeetHub system. This chapter discusses those related to the perception and interaction of the contextual information scents:

- **Attention of focus**. We asked each participant which display, i.e. the wall display or the iPad, received more attention of the respondent.
- **Distraction of the Marquees**. We asked each participant to what extent the subjects agreed that the Marquees were NOT disturbing.
- **Helpfulness of the Marquess**. We asked each participant to what extent the subjects agreed that the marquees were helpful for accomplishing the task.
- **Usefulness of the Wordcloud**. We asked each participant to what extent the subjects agreed that the Wordcloud was useful.

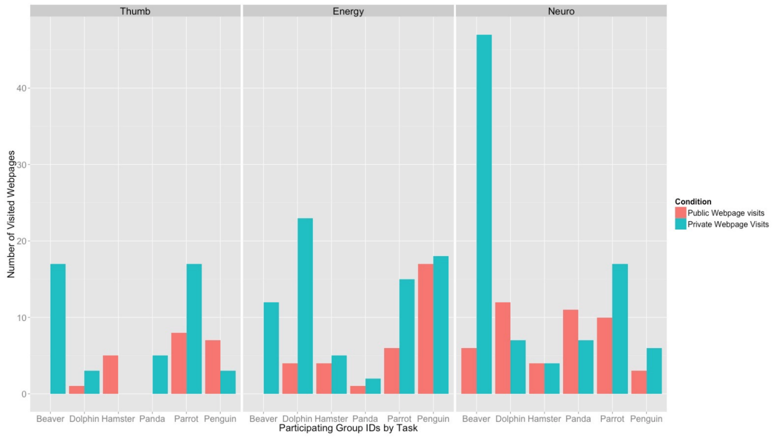
The latter four questions were all 5-point Likert scale questions (1 to 5 representing strongly disagree to strongly agree).

Interaction Logs

For the interaction logs, we are interested the number of searches conducted by each group, as well as the number of interactions with the contextual information scents in the MeetHub system.



(a) Public Web Searches vs. Private Web Searches



(b) Public Webpage visits vs. Private Webpage Visits

Figure 7.6: Searching with dedicated search tools in the MeetHub

7.3.6 Results and Findings

The MeetHub system offered many tools across different devices. My two other colleagues had special interests in studying group interactions and time management respectively. Detailed results in these aspects have been published in (Roman, 2013; Verma, 2015). This section exclusively reports results regarding the MeetHub Search component. The results are presented in two main themes: (1) searching and viewing information, and (2) interacting with information scents. We examine the effects of search devices, and the presence of search roles. For the information scents, we also compare the effectiveness of different query-building approaches.

Searching and Viewing Information

The MeetHub offers two types of dedicated search tools. Users could either search together with the Web browser on the wall display, or search individually on their own iPads. Additionally, Webpages being viewed privately on an iPad can be shared onto the public display. In the weekly questionnaire, the participants were asked to indicate on which device their attentions were mostly focused during group work. 12.2% (6 unique subjects) reported to have mainly focused on the iPad, whereas the majority (87.8%) concentrated on the wall display. However, their search activities exerted a different pattern.

During the 3-week formal study, each group conducted on average 1.44 (σ = 2.36) public searches on the wall display and 11 (σ = 11.51) private searches on the iPads, which further led to 5.5 (σ = 4.72) and 11.55 (σ = 11.16) Webpage views respectively. Figure 7.6(a) and Figure 7.6(b) illustrate the number of searches and Webpage visits of each group for each task. Not all the groups conducted both public and private searches, but search activities on the iPads are visibly more intense. Two three-way within-subject ANOVA (groups are modeled as random effects) are conducted to compare the effects of both search devices and the role conditions on the number of Web searches and Webpage visits respectively. Users conducted significantly more private searches ($F(1,27)=13.35, p=.001$) and visited more Webpages ($F(1,27)=4.92$, $p<.05$) on the iPads than on the wall display. The presence of pre-defined roles ($p=.71$ and $p=.49$) and different tasks ($p=.84$ and $p=.25$) do not show significance in both outcome variables .

The findings above show that, more often than not, the group participants searched information independently. Assigning the role of information searcher to a specific participant in a group did not seem to exert significant influence on the groups' search behaviors. Additionally, we have no statistically significant evidence to prove the groups searched more often in one task than another, indicating different tasks might require similar intensity of Web searches to support group work.

According to Figure 7.6, some groups seem to have been obsessed with searching information on the Web. For example, the Beaver group conducted in total more than 40 individual searches on the Thumb task (30 minutes), and more than 30 searches on the Neuro task (45 minutes). These groups ended up with spending much effort on searching information individually, which is generally not desired for the group work.

Interacting with Information Scents

As described earlier in this chapter, three different types of information scents were presented in the MeetHub system. We also altered the query-building approaches for retrieving relevant results from the Web with three variants, each for one task. On the wall display a group selected on average 1.28 (σ = 1.67) marquees and 0.33 (σ = .97) words in the Wordcloud. On the iPad interface, 1.06 (σ = 2.80) query terms were selected in the Querylist. Once a query term was selected, for 66.7% (σ = 28.9%) of the time, the user also further selected the images or Wikipedia results in the expanded list.

Frequency of Interaction

According to Figure 7.7, the Wordcloud was used in the "Thumb" and "Energy" tasks (6 times

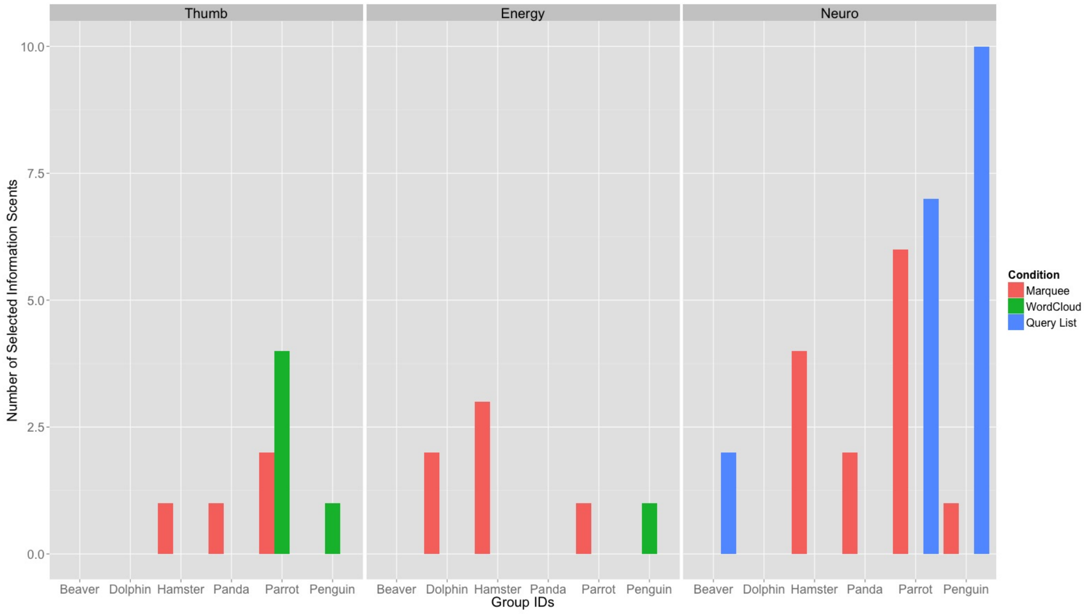


Figure 7.7: Interacting with information scents in the MeetHub

in total), the Querylist was used in the "Neuron" task (19 times in total), and only the Marquees were used in all of the three tasks (23 times in total). In addition, the use patterns shown in the figure is quite sparse. So, was interacting with the contextual information scents offered in the MeetHub opportunistic? If we take a look at the Wordcloud, the answer might be yes, since it was used only 6 times in 18 study sessions over 3-week time. The small number of use also makes it difficult to make any statistical analysis meaningful. In contrast, the overall interaction frequencies of the Marquees and the Querylist in the Neuron task was visibly higher than in the previous weeks, which may not be a coincidence.

Each task was associated with a different query-building approach, with which the search results were returned for composing the Marquees and Querylist. The TA, CA and WA[2] generated 59.17 (σ = 29.68), 134.17 (σ = 13.16) and 75.5 (σ = 7.78) query terms. When the interaction frequencies of the Marquees and items in the Querylist were concerned, we built a within-subject two-way ANOVA with the query-building approach, the role condition as independent variables. The query-building approach was shown to be a significant factor (F(2,10)=4.07, p=.05). Pair-wise comparisons further revealed significant differences in the interaction frequency between the WA and the SA (t(10)=-2.56, p<.05), the WA and the CA (t(10)=-2.37, p<.05). There was no significance difference between the SA and the CA (t(10)=-.18, p=.85). Interactions with the information scents were in fact less frequent in the ROLE group ($\mu = 1.78, \sigma = 3.56$) compared to the NOROLE group ($\mu = 2.89, \sigma = 4.01$), though the difference was not significant (p=.54).

Perceived Distraction

[2] the statistics for the WA were estimated from only 2 samples due to logging errors

As discussed in Section 5.2.1, one important principle for the design of contextual information scents is that the design must not be distractive. The MeetHub users focused primarily on the wall display, where the marquees scrolled from to time. Therefore, a certain level of distractions for the marquees are foreseeable. In the post-experiment questionnaires, we asked the users to give subjective ratings regarding the distraction of the Marquees. The subjects overall held neutral views about it ($\mu = 3.05, \sigma = 1.17$) throughout the whole study. Note that the 3 query-building approaches employed different activation schemes for the marquees, which may induce different level of distractions. The perceived distraction for the marquees generated with TA, CA and WA were ($\mu = 3.2, \sigma = 1.15$), ($\mu = 2.75, \sigma = 1.15$), and ($\mu = 3.2, \sigma = 1.19$) respectively (higher ratings are associated with less distractions). With one-way within-subject ANOVA, we found marginally significant effect of the query-building approach on the perceived distraction (F(2,47)=2.97, p=.06). Pair-wise tests showed the Marquees generated with the WA and TA were significantly less disturbing than the ones from the CA (t(47)=2.12, p<.05), the difference between WA and TA is however not significant (t(47)=0,p=1).

Perceived Usefulness

The Wordcloud was interacted only 6 times in 3 weeks. Accordingly, the average perceived usefulness of it was low ($\mu = 2.79, \sigma = 1.02$). The perceived helpfulness of the Marquees generated with the TA, CA and WA were ($\mu = 2, \sigma = 1.12$), ($\mu = 2.21, \sigma = .93$) and ($\mu = 3.04, \sigma = .84$), respectively. With one-way within-subject ANOVA, the effect of query-building approach was significant (F(2,47,p<.0001)). Pair-wise tests suggest that the marquees generated with the WA were significantly more helpful than with either TA (t(47)=4.7, p<.0001) or CA (t(47)=3.65, p<.001). The difference between the TA and CA was not significant (t(47)=1, p=.32).

7.3.7 Discussions

The MeetHub Search was designed to investigate how contextual information scents generated from user interactions with a groupware can enhance collaborative activities. In this section we further discuss the results presented before.

Presence of designated information searcher

The groups participating in the study were divided into NOROLE and ROLE conditions. We expected the groups in the ROLE condition to interact more with the contextual information scents, because the information searchers were asked to consciously monitor the contextual information scents, such as the Marquees. However, we did not have sufficient evidence to show significant difference in the number of interactions between the "ROLE" and "NOROLE" conditions. In other words, the ROLE condition was not shown to induce more serendipity. A possible reason could be the groups did not constantly need information, so that an information searcher did not always have a prepared mind for serendipitous encountering. Initially the information searcher might carefully carry out the duty of consciously monitoring the information scents. As more and more contextual information scents were not useful, the information searchers were likely to just superficially look at them rather than to make connections in mind.

"Chicken-and-egg" problem

The "chicken-and-egg" problem is often described as a philosophical dilemma as "which came first, the chicken or the egg". We expected the contextual information scents of the MeetHub to enhance collaboration, but this is in analogy to a "chicken-and-egg" problem: In order to generate contextual information scents based on interactions, users must first interact with the groupware. Otherwise the information context does not exist. Take the Marquees as an example, we expected the users to first articulate ideas by typing them into the shared workspace. As we observed in the study, the groups usually discussed verbally or searched individually before interacting with the groupware. For most of the time, they started typing in the workspace when they had to input the discussed results into the system. In other words, most of the "learning" processes through discussion and argumentation were not captured by shared workspace. The Marquees, especially those generated with the TA, represented only the learning context in the past, which may violate the principle of timeliness (cf. Section 5.2.1).

Effectiveness of the information scents

In the discussion of the "Chicken-and-egg" problem, we argued that the contextual information scents were not generally effective to enhance collaborative work. However, we found that the WA was perceived as significantly more helpful than the other approaches. This is yet an interesting finding. The Marquees associated with the WA induced more serendipitous encounters, because it took into account the global task context as well as individual queries, which were never produced on the shared workspace. These information was used by the system to retrieve search results either before or right after user interactions (assured timeliness). This finding especially highlights the importance of including keywords from the task descriptions as well as explicit search terms to build query-terms for the generation of contextual information scents. These words usually contain more unknown truth that require the users to explore, while the words in the shared workspace are almost certain.

Limitations

The design of the MeetHub Search as well as the study of it has several limitations. First, the choice of combining two images and one Wikipedia result rather than other combinations for the design of a Marquee could not be justified. Second, the scrolling animation of the Marquees might also be disturbing. We could have used the still visualization as in the TileSearch, but there was not enough screen real-estate to display still image or Wikipedia tiles.

Third, each query-building approach was associated with a different task, which may have influenced the comparison results.

7.3.8 Conclusion

This chapter presents only one design prototype of contextual information scents generated from user interactions with a groupware. However, we explored different designs within a single system. As in the RaindropSearch discussed in Chapter 6, presenting contextual information scents as keywords in the MeetHub Search, again failed to induce serendipity, though the keywords were better extracted semantically from Webpages and were presented with a Wordcloud layout. Therefore, we learned a lesson that presenting contextual information scents as individual keywords may not be effective.

A more generalizable lesson from the MeetHub study is that the interaction context in a groupware may stuck in a "chicken-and-egg" dilemma, which goes against the *timeliness* design principle. Note that the two explorations with contextual information scents from group conversation in Chapter 6 mainly violate the *relevance* design principle. Therefore, future systems must be designed to best satisfy these two principles at the same time. But how can we achieve it? A better prediction of search queries in advance seems to be a solution. In the next section, we present a prototype that attempted to meet these criteria in a collaborative MOOC viewing scenario.

8 Exploring Content Context

The previous projects mostly incorporated user-generated context, such as conversation and interaction to create contextual information scents for augmenting collaborative learning activities. As discussed in the previous chapter, it is difficult to tackle the challenges of relevance and timeliness at the same time. In addition, searching information with general-purpose search engines also induces randomness. This chapter further investigates how the content of learning materials can be exploited for investigating the design of contextual information scents. We introduce a new, but authentic learning scenario, collaborative MOOC learning, where students sit together to view and discuss MOOC lectures. This scenario distincts from the previous collaborative learning activities in two aspects: (1) apart from group discussion, participants spend most of their time for passively receiving knowledge from lecture videos, as they attend lectures in the classroom (2) participants learn through fixed curriculum and learning materials (the learning context), allowing contextual information scents to be elicited and prepared by experts in advance. The second point in particular makes it possible to ensure timeliness and relevance at the same time.

In this chapter we present the BOOC Player, an application that synchronizes textbook content with MOOC. The application leverages a tablet display split into two views to present lecture videos and the corresponding textbook content simultaneously. The display of synchronized textbook is intended to serve as peripheral contextual help for collaborative video viewing activities [1].

8.1 Exploration: BOOC Player

Massive Open Online Courses (MOOCs) have been in recent years growing in popularity. Popular platforms such as Coursera[2] and edX[3] typically replicate classroom pedagogy online,

[1]The content in this chapter has been published or presented in different research venues in various forms. Publications [1,2,4,6] and presentation [14] in publication list (last page in the thesis) are concerned.

[2]http://www.coursera.org

[3]http://www.edx.org

featuring with various learning components such as lecture videos, online quizzes, tutorials, discussion forums and Wikis. These learning components on one hand have made self-guided individual learning possible. On the other hand, the massively distributed nature of MOOC learning has posed many new challenges for education researchers (Yuan and Powell, 2013). One direct negative consequence is the elimination of intimacy between instructors and learners. As a result, learning feedback cannot be obtained directly from the instructors (Kop et al., 2011). Instead, automated algorithm-driven processes as well as peer assessment are employed to grade one's work. MOOC learners often have to seek remote support from their fellow students with discussion forums to achieve "the learner is the teacher is the learner" (Siemens, 2006). Unfortunately the social forum interactions are often temporary and asynchronous. The ties between the learners are loose and timely support are not guaranteed. Furthermore, the diversity of learners (Kizilcec et al., 2013) makes it especially difficult to maintain the activeness and quality of the loose collaborations within a forum. In fact, lecture video viewing is the central activity in MOOC (Breslow et al., 2013; Seaton et al., 2014), and prior research has shown that only around 5% of the students actively participated in the forum (Huang et al., 2014). This gives sufficient grounds for forming MOOC study groups, which mixes the activity of lecture video viewing and collaborative learning.

Revisiting the MOOC components listed before, we find that textbooks, which are typical learning materials in traditional classroom pedagogy, are missing. Textbooks may be available as references in the Wiki, but they are not essential for learning through MOOCs. The role of the textbooks is probably displaced by the lecture videos, which can be seen as "video books" that allow students to refer to at any time with video navigations. However, this does not mean textbooks are not useful in MOOC learning, because they are known to contain more structured and complementary learning content. In a MOOC taught with a companion textbook, Belanger et al. (2013) found that the students often spontaneously identified related content in the textbook, and then shared and discussed them in the forum. This finding does not only exhibit that textbooks are still functioning as supporting materials in MOOCs, but also implies supervised book-to-video references made by instructors are potentially useful for the students.

The above discussion triggers our reflections: What if we elicit content from textbooks as contextual information scents to augment collaborative MOOC learning activities? Considering MOOC videos are pre-recorded, additional efforts can be made to link and display relevant textbook pages alongside the lecture videos being played. The textbook pages serve as information scents that are always pertinent to the video content, and both relevance and timeliness principles are met at the same time. This section presents related works as well as the design of technology, i.e. the BOOC Player.

8.1.1 Related Work

In Chapter 2, we briefly reviewed collaborative learning, but in this section we recapture the main concept of it with a special focus on MOOCs. This is followed by a literature review in collaborative viewing and contextual help, which are of high relevance to the collaborative MOOC learning scenario that the BOOC Player is targeted for.

Collaborative Learning for MOOCs

As we are thinking of MOOC study group, one may naturally associate it with online study groups. In fact, online study groups have been well studied in literature. As an example, Curtis and Lawson (2001) studied groups in this format in a small course of 24 students, who were required to work on course assignments in self-selected groups via a dedicatedly designed web-based application or emails, students reported to have suffered from asynchronous discussion and collaboration with strangers of diverse background. Similarly, Smith et al. (2011) found that learners reported to have experienced more frustrations in online groups than in their face-to-face counterpart. The frustrations may attribute to the differences in study goals, imbalanced participation and the quality of individual contribution, as summarized by Capdeferro and Romero (2012). Despite of these possible frustrations, online study groups still have potentials for stimulating collaborations. As another MOOC initiative, NovoED[4] creates a social incentive system to tackle the challenges of online MOOC groups: small group collaborations are enforced and are implemented via Google Doc and Hangouts. Individual contribution in a group is peer-rated so as to encourage participation and contribution. Most of the courses offered on NovoED are entrepreneurship courses whose curriculum consists of group projects, where online collaborations are expected to take place. In other words, collectively creating knowledge in group projects is an explicit requirement for most NovoED courses. In comparison, courses in other domains, such as in technology and mathematics, focus on mastery learning with knowledge duplication (Siemens, 2012), and they do not necessarily have tasks designed for groups.

In contrast to online collaboration groups, collocated study groups are common practices in schools and universities, regardless of the requirements of group-based projects. Students often form spontaneous study groups to learn a course together, and such spontaneously formed groups are shown to be effective in achieving better outcome in terms of grades than individual learning (Tang, 1993). However, concerns are more given to the seemingly impracticalness of collocated study groups in MOOCs. In fact, as MOOCs have reached large scale, geographical clusters of students are likely to emerge. This trend can be seen from the Coursera Meetup[5], where students that are geographically close to each other have the opportunity to study together. We have observed that local meetings are actually being organized spontaneously, but they are mostly unstructured. Meetup in its current form does not provide

[4] https://novoed.com/
[5] http://www.meetup.com/Coursera/

suggestions on how to study together. In addition to the Meetup, universities naturally gather learners. It is highly probable that a student is following the same MOOC with his/her fellow students. Furthermore, in many universities, participating MOOCs becomes part of the course curriculum (Martin, 2012). This is achieved with the flipped-classroom model (Tucker, 2012). The proliferation of flipped-classroom teaching has provided opportunities to on-campus students to form face-to-face MOOC study groups as they often do for traditional courses.

Collaborative Video Viewing

As MOOC learning is centred on lecture videos, a more general group-based learning approach have the potential to arise on the activity of lecture video viewing. There exists an extensive body of research in the field of collaborative video watching in research literature. One of the earliest study may date back to the 1970s, when Gibbons et al. (1977) coined a term *Tutored Video Instruction (TVI)* to denote the scenario where remote students watch video lectures in small collocated groups with a tutor. With TVI, both students and tutors were able to pause video lectures initiate discussions when problems and questions arose. Gibbons et al. found that in terms of average grade obtained by the students, TVI students outperformed students who watched live video lectures in the classroom and those who watched offline video lectures. Surprisingly, TVI students also outperformed on-campus students who attended the lecture in the classroom. In another study, Stone (1990) found that even in the situations where tutors were not present, i.e. simply watching lecture videos in a collocated group was still advantageous.

In the late 1990s, a group of Sun and Microsoft researchers (Sipusic et al., 1999; Smith et al., 1999) extended the original TVI methodology to distance learning, where the webcams and microphones were used to mimic the collocated version of TVI. The authors coined a new term *Distributed Tutor Video Instruction (DTVI)* to distinguish it from the original collocated TVI. They found that the advantages of group watching video lectures were repeated in the distributed condition as well. A follow-up research on DTVI without tutors also confirm the same conclusion (Cadiz et al., 2000). In this work, a new term *Collaborative Video Viewing (CVV)* was coined to represent the scenario where TVI is conducted without tutors. Similar to DTVI as compared with TVI, DCVV was used to represent the distributed condition. In addition, Cadiz et al. (2000) compared CVV with DCVV on learning and interaction behaviors. Their results exhibited that the co-located groups were significantly more comfortable with pausing videos so that they discussed for longer duration both in total and per pause as compared to the distributed groups. More discussions are considered as beneficial for TVI groups, as Weisz et al. (2007) showed in their research that discussing while video watching was perceived to be an engaging and enriching social experience by the participants.

From the above review of prior research on CVV, the advantage of this model is notable. Early research on TVI required the presence of a tutor, which is not a realistic solution for MOOC. DCVV can be potentially feasible, but current MOOCs do not technically offer synchronous online collaborative video watching experiences. In comparison, CVV can be achieved for students that are close to each other.

Contextual Help

The BOOC Player links textbook pages to lecture video content, so alternative or complementary explanations are provided to learners in case they are required. This is similar to the notion of *contextual help* (Capobianco and Carbonell, 2001; Carenini and Moore, 1993), which is described as delivering specific information that users may need at the right moment, when they are carrying out a task with a computer. Such contextual help can be of operational guidance, which is intended to assist users in using the functions of a computer program. It can also be of task guidance, which is designed to help users complete a task (Heift, 2006).

In research literature, many research projects about contextual help have been devoted to operational guidance. ToolClips(Grossman and Fitzmaurice, 2010) embeds video tutorials as contextual assistance for tool functionality understanding in a software application; InterTwine (Fourney et al., 2014) creates information scents by linking Web browsers with software features to assist users in finding help information on the Web. Other projects focused on the design of contextual help to support learning procedural knowledge of a software. These projects endeavored to create links between graphical interface to be learned and video tutorials. FollowUs (Lafreniere et al., 2013) demonstrates that software learning can be enhanced by multiple demonstrations of tutorial videos from other community members. Pause-and-Play (Pongnumkul et al., 2011) is similar to the BOOC Player, it employs a method to detect task-performing events in the video and link them with user actions in the target application as the user tried to imitate the procedure. This method avoids manually switching between the user context and the online tutorial. Contextual help for supporting learning tasks has also been explored. The E-tutor (Heift, 2006) is a language tutoring system that instructs learners to complete a language-learning task. It automatically generates error-specific feedback, grammar hints as well as additional help from a dictionary, in case a learner has failed in the task.

8.1.2 Design of the BOOC Player

The BOOC Player was implemented as an iOS application for Apple iPad devices. In this section we describe the user interface, the interactions it offers as well as how it is positioned in the design space o fcontextual information scent.

User Interface of the BOOC Player

As shown in Figure 8.1, the system has 4 different views. The *Catalogue* view lists the titles of the available lecture videos. Tapping an item in the list automatically navigates to the *MOOC Video* view and plays the corresponding video. The *Quiz* view allows a user to view the online quizzes of the MOOC. Finally, the *Admin Control* view is only for experimenters to manage

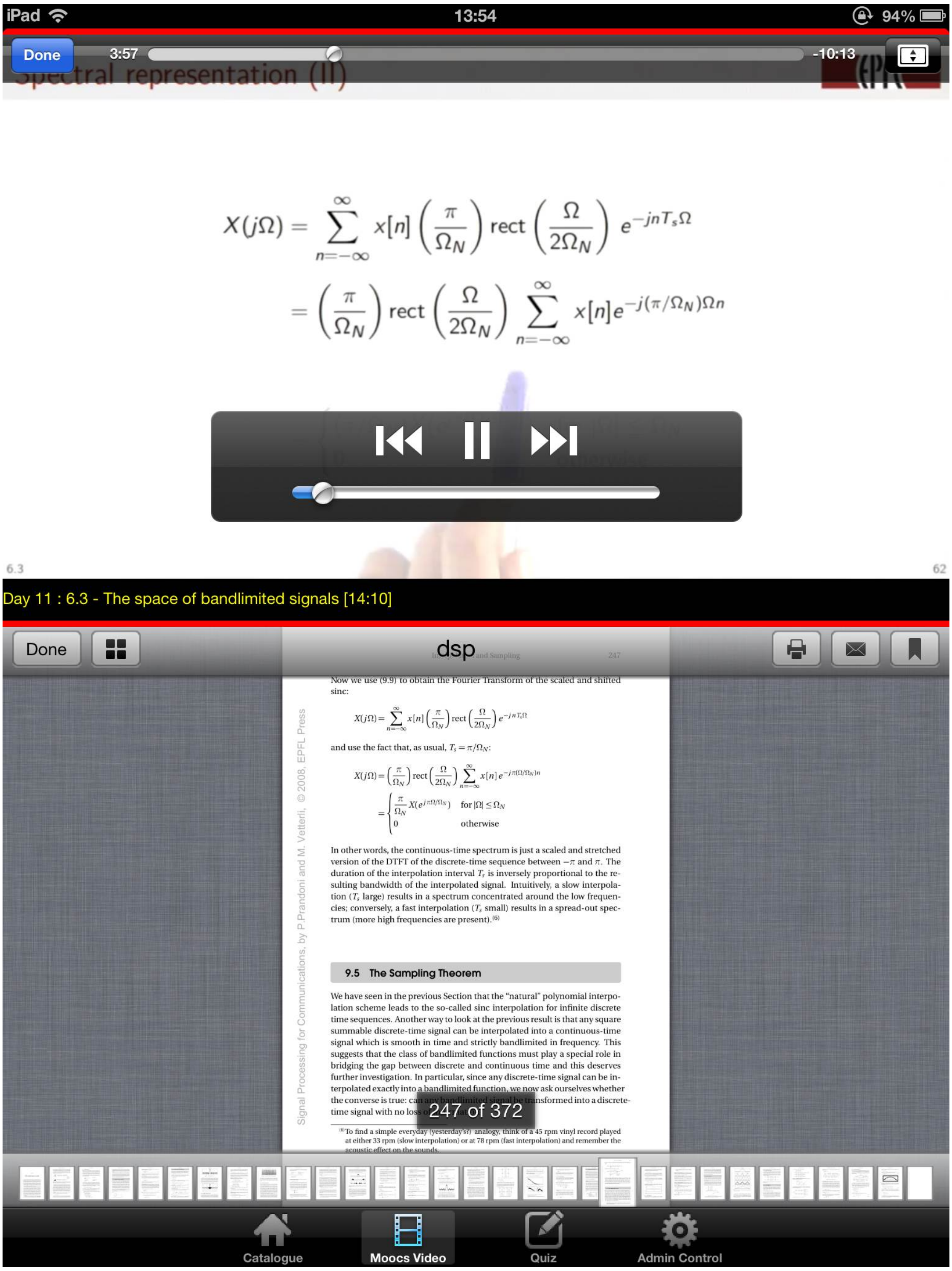


Figure 8.1: The BOOC Player

experiments, therefore is not visible to users.

In fact, we refer to the *MOOC Video* view, which is divided into two parts, as the BOOC Player. In its main screen, a video controller plays lecture videos, and a PDF controller displays digital textbooks and manages page navigations. Each controller has a status indicator designed as a straight colored line. When both indicators are shown red (as seen in Figure 8.1), it indicates the contents are synchronized. Otherwise, it is a sign that they are not synchronized. Interactions for toggling the synchronization state will be explained later. By "synchronized", we mean the most relevant book page of a textbook at the time is displayed as the lecture video progresses. The videos and textbooks are dually mapped, so changes in one controller are likely to be reflected in the other. Conversely, desynchronized state means that changes on one side will not affect the other.

Interacting with the BOOC Player

The video controller offers full functionality of a video player. A user can play, pause and jump in the video back-and-forth by scrubbing the playhead. The PDF controller allows panning and scrolling a page. It also supports swiping gestures for page navigation. A user can also select a page thumbnail at the bottom of the controller to view the corresponding page content.

The two controllers are synchronized by default. If the current video content being played is beyond the textbook, then the PDF controller is greyed out, indicating that no textbook pages are relevant at the moment. In case multiple pages are related to the same video segment, only the most relevant (supervised by tutors) one is presented in the view. The page numbers of the other relevant pages are shown as yellow text in the middle right of the screen. Students can navigate to those pages at their own effort. Videos and books are dually mapped. Users can also navigate through the digital book to get the corresponding video explanations, if available. In case of multiple mappings, the system pops out a list of other relevant videos for selection. Double clicks on the PDF controller toggle synchronization states. When desynchronized, the PDF controller's status indicator will turn green.

BOOC Player in the Design Space

Though the BOOC Player was implemented as an iPad application, it may also be adapted for public use, especially when it is connected to a projector. We will investigate this aspect in the user study to be presented later. Since the target usage scenario is collaborative MOOC video viewing, much of the learning context is anchored around the video content, based on which the contextual information scents, i.e. the textbook pages are elicited and presented. It is impractical to automatically make reliable mappings between lecture videos and textbooks. In order to ensure quality, the mappings in the BOOC Player were manually made by teaching staffs. Thus, the most relevant complementary learning materials in the book can be timely

Privacy	Personal		Public
Context	Conversation	Interaction	Content
Capacity	Low	Medium	High
Uncertainty	Low	Medium	High
Activation	Proactive		Reactive

Figure 8.2: BOOC Player in the design space

presented to learners with low uncertainty. The contextual information scents in the BOOC Player do not react directly to user behaviors. Rather, they are reactive precisely to the ever-changing video content context. Overall, the system finds itself in the design space as shown in Figure 8.2.

8.2 BOOC Player: Research Questions

Given the advantages for students to study together, we are motivated to replicate the CVV approach in MOOC learning. We consider to investigate self-formed CVV groups without tutors, but with the BOOC Player. As previously mentioned, self-formed study groups are common for studying traditional courses at schools and at universities, but little is known about how this approach can be replicated to MOOC-based learning. Watching videos on a shared display as experimented in prior CVV research is definitely an option for arranging a study group. Considering MOOC stresses personalized learning experience and most MOOC learners have personal computers, it is also of natural practice for students to watch lecture videos on personal devices at their own pace while maintaining the group atmosphere for spontaneous discussions. Our goal is to understand the group dynamics in both types of group arrangements and investigate the effectiveness of the BOOC Player accordingly. The investigation of these issues may also provide pedagogical implications for the flipped classroom teaching, the organization of Meetups as well as distance educational programs for developing countries where digital infrastructure is limited. In all, the following research questions will be addressed in this chapter:

(1) How do students in different group conditions watch videos and discuss with each other?

As students mainly learn through watching lecture videos, we anticipate them to encounter difficulties or problems from time to time. Clearly, they can initiate discussions with group members, pause and think by themselves, turn to the textbook, or re-watch parts of the videos. There may be certain behavioral patterns emerged naturally with the study groups. We are interested in when (e.g. during or after watching videos) and how much the students discuss with other people. The discussion behaviors may in turn influence the video interaction behaviors, especially for the groups who watch videos on separate devices. These students have the option to watch at their own pace, but will they actually watch in this way? What are the key latent factors that mediate the group learning?

(2) In what aspects is the BOOC Player effective for MOOC video viewing and discussion groups?

As mentioned previously, a group student may have several options to deal with difficulties, and turning to the textbook is just one option. In order to find out the appropriateness of the BOOC Player, we are interested to see the advantages of the BOOC Player over print textbooks available to the group. How often will they use a print textbook or a digital book in the BOOC Player? We are interested to understand both paper and digital book interaction patterns and their potential effects in the collaborative learning experience.

8.3 Experimenting BOOC Player in MOOC Study Groups

In order to answer the research questions, we conducted a 5-week longitudinal study in the spring of 2013. Our study was based on two Engineering courses offered by our university at Coursera, namely, Numerical Analysis (NAS) and Digital Signal Processing (DSP). The first 7 weeks of both courses were arranged as flipped-teaching format. During this period, the students were required to watch videos and solve quizzes at home. Classroom sessions were reserved for exercises and advanced tutorials. For the rest weeks, the on-campus courses were offered in traditional classroom.

8.3.1 Participants

We recruited on-campus students from the two courses for our study. In total, 25 students (8 females/ 17 males) from the NAS and 9 students (all males) from the DSP course formed 8 study groups. The distribution of participants are illustrated in Table8.1.

Table 8.1: Participants in the study

Course	Total Groups	Females	Males
Numeric Analysis (NAS)	6	8	17
Digital Signal Processing (DSP)	2	0	9

Each group contained 4-5 participants. The groups were self-formed, so that the students in each group were well acquainted with each other. We let the students to form groups freely because it is common practice for students to discuss with familiar fellow students. All participating students reported to have participated in study groups in the past. However, most of them had no previous experience of learning MOOCs. The study lasted for 5 weeks, from the second course week to the sixth. Each subject was compensated 150 Swiss Francs, together with a print companion textbook of the corresponding course for participating 5 weekly study sessions.

8.3.2 Formats of MOOC Study Groups

As discussed in the research question section, we believe MOOC study groups can be practically configured in two formats: Students in a group can either watch videos on a shared display or on their own computer devices. We use two dimensions, the *display* and *video controller*, to represent the configuration space of MOOC study groups.

Table 8.2: Two dimensional configuration space for MOOC study groups

		Display	
		Centralized	Distributed
Video Controller	Centralized	CC	-
	Distributed	DC	DD

The two aforementioned formats correspond to the CC and DD cell in Table 8.2, respectively. More detailed explanations of the corresponding technical setup in our study are described below:

- **Centralized video control and centralized display (CC)**: We implemented the CC format by connecting an iPad to a beamer, which projects the videos on a wall-mounted display (cf. Figure8.3(a)). Video interactions from multiple participants in a group are through a single tablet with touch actions.
- **Distributed video control and distributed display (DD)**: Each group member is assigned an iPad with the BOOC Player for individual use (cf. Figure8.3(b)). Each student can not control his/her own video. The groups wear ear phones during video watching in order not to disturb with others.

Besides the CC and DD, in Table 8.2 we list a DC group format, where study groups also watch MOOC videos on a shared display. In contrast to the CC, each DC group member has individual control over the video. This is not feasible with traditional computer devices. My colleague extended the functionality of the MeetHub system presented in Chapter 7 to

(a) CC condition: Group members watch MOOC videos on a projected display with a single tablet to control videos

(b) DD condition: Each group member watches and controls MOOC videos on a separate tablet

Figure 8.3: The CC and DD study group formats

allow multiple users to interact with a video player. He conducted an independent study with 4 groups formed by 20 students recruited in the same subject recruitment process. More details about this study are documented in (Verma, 2015). Note that centralized video control with distributed displays, i.e. a CD group format is not available in Table 8.2, because it is meaningless in practice. The study in this chapter only deals with the CC and DD because they can be easily configured and popularized with minimal technical requirements. In the discussion hereafter in this chapter, we use the term "group condition" to refer to the group format rather than an experimental design condition. In other words, the goal of having two group formats in our study is to understand how self-formed groups study MOOCs and use books in both conditions, rather than to compare their effect of the formats on group performance.

8.3.3 Procedure

The 8 groups were evenly distributed across the two conditions, such that each condition had 4 groups (3 groups of NAS, 1 group of DSP). Each group met once a week to study the lecture materials in the corresponding week. The study lasted for 5 out of 7 weeks of the whole flip-teaching period for both courses. The participants were asked not to watch the MOOC videos before coming to the study group sessions. We did not intend to instruct the groups to watch videos or to learn together in a particular way. They were encouraged to behave as naturally as possible. However, we requested each participant to bring the print textbook to the weekly study session, and place it on the table for necessary use. In addition, we manipulated the type of the video player on the tablets as a within-subjects design. The participants watched MOOC videos with a normal video player (without the synchronized PDF) for the first 3 weeks, and the BOOC Player used in the last two week. Such manipulations allow us to observe the potential behavioral changes before and after the introduction of the BOOC Player. As said before, the book-to-video mappings in the BOOC Player were made

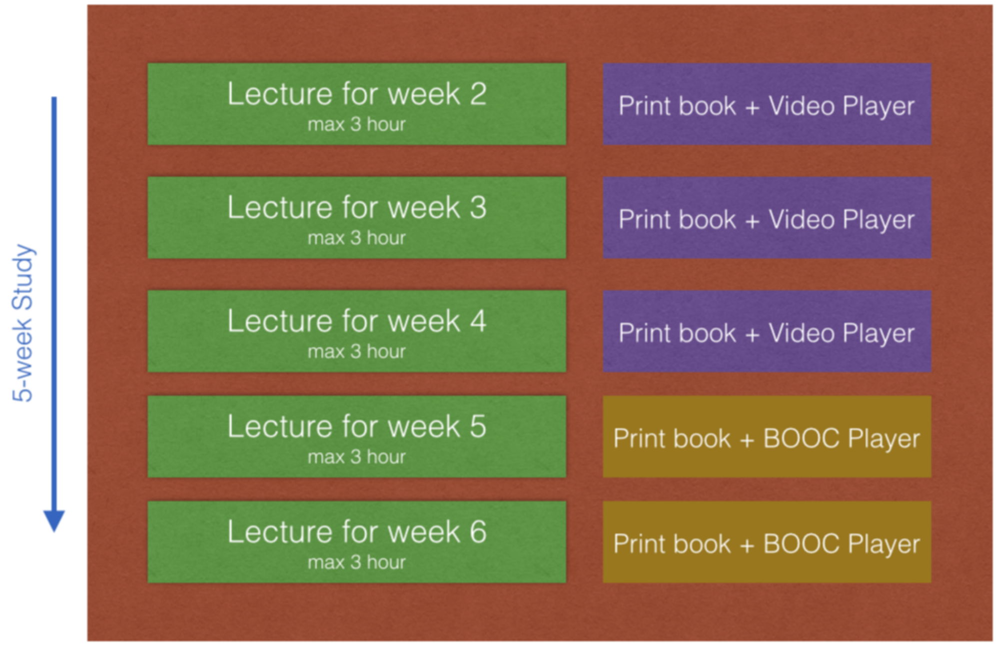


Figure 8.4: Learning material and tools for the experiment

manually with the help of a designated teaching assistant of the corresponding course. Each week's videos (less than 10) required around one hour to map with the textbook. The overview of study design is illustrated in Figure 8.4.

In the first week of the study, we asked each participant to fill in a pre-experiment questionnaire (cf. Appendix E.1) regarding demographics, personalities as well as experiences with study groups and MOOCs. This questionnaire was completed only once during the whole study period. Then, the participants were trained to use the tablet application to be used throughout the study. The training was only given in the first week and in the fourth week, when the BOOC Player was introduced.

Each week, the study procedure can be described below:

1. Before the start of each study session, the groups were provided with a printout of quizzes to be completed for the week. The quizzes were the same as the weekly quiz on the Coursera course website, so the students could also view them in the "Quiz" view of the tablet application. The participants were not obliged to complete the quizzes in group, but they were opt to do so.

2. Next, the group was asked to start the study session. They could watch videos and discuss with others at their own pace. In principle, a study session was not strictly time-bounded. Normally the total video length of a week was between 1 hour and 1.5 hours. The students were given 3 hours to complete a weekly session.

3. When time was up or the groups terminated the study session by themselves, each participant was asked to fill in a post-experiment questionnaire (cf. Appendix E.2) regarding their feedback about the system as well as their learning experiences. The same questionnaire was repeated each week. This was followed by a semi-structured interview for around 15 minutes regarding their collaboration and the use of textbook.

8.3.4 Data Collection

As described before, we had weekly questionnaires and semi-structured interviews to collect students' subjective experiences. In addition, we had video interaction logs obtained from the tablet video player. We also videotaped the study sessions. Two cameras were employed to capture both the front and rear view of the group interactions. In addition, the weekly post-experiment questionnaires provided us with data about the students' subjective experiences.

Questionnaires

The questionnaires covered many facets of group learning. This chapter focuses on the following aspects regarding group collaboration:

- **Video Difficulty**. We asked each participant to rate the overall difficulty of the videos watched in the study session.
- **Discussion Quality**. This question is concerned with the ratings of the participant's perceived quality of discussion during the study session.
- **Equal Contribution**. Each participant was also asked to rate how equally the group members participated in the discussion.

All of the three aspects were assessed on 5-point Likert scale (1 to 5 representing strongly disagree to strongly agree). We use these subjective ratings to inspect the factors that have influenced group collaboration and how video difficulty was addressed. In addition, we also summarize the open questions about the textbook usage, and both print book and the BOOC player were concerned.

Interaction Logs

The tablet video player automatically logs the timestamp when a specific video was played, paused or seeked. This allows us to study the video interaction patterns emerged in MOOC video viewing activities. It allows to log the PDF interactions such as page flipping and selection on the BOOC Player.

Video recordings

Though the interaction logs recorded the interactions on the BOOC player, print textbook interactions as well as collaborative interactions involving multiple subjects with either type of book were not traced. We manually examined the video recordings obtained from the five-week study sessions of all groups and coded these occasions. Coding the use of books provides us with deeper insights about the role and benefit of the BOOC Player. In addition to coding the textbook use, we also coded the amount of speech in each group to understand how group discussions took place.

8.3.5 Revealing Group Dynamics

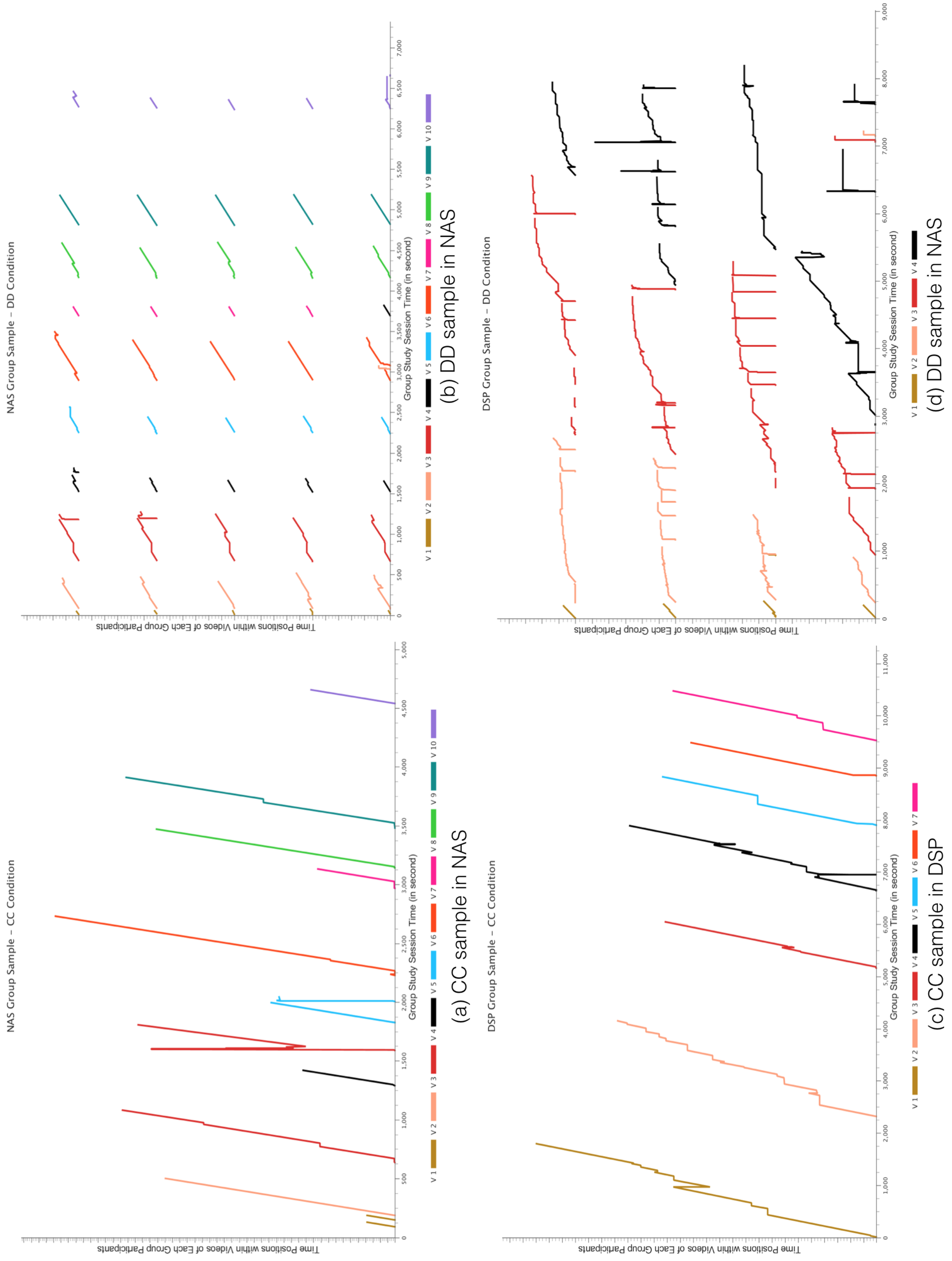
Although our main research interest concerns the use of textbook, especially the contextual information scents rendered by the synchronized PDF in the BOOC Player, it should be noted that collaborative MOOC video watching was the primary activity in the group sessions and the using textbooks came second. In this section we present our findings regarding groups' video interactions and discussions, which would facilitate our understanding of textbook uses that are to be discussed in the next section.

Group Video Navigation Patterns

In order to learn how the study groups interacted with MOOC videos, we visualize their video interaction patterns. Four plots, each illustrating the interaction patterns for a representative sample group in a different condition, are shown in Figure 8.5.

The horizontal axis of each plot represents the timeline of a study group session, and the vertical axis denotes the timeline of the videos. Both timelines are measured in seconds. The groups watched multiple videos in each week, and each color in the plot represents a different video. Figure 8.5(b) and (d) are the examples of the DD condition, where the patterns for multiple students are shown in parallel. A straight line-segment with a positive slope indicates that the corresponding video was played without interruptions; a straight horizontal line-segment is a sign of a pause; jitters depict seek forwards and backwards within the video; the gaps between two continuous series are the time periods when students were discussing about the problems or doing quizzes (no videos were being watched at that moment). Students did not take breaks, so the plots portray a complete picture of the activities during a group study session.

Generally speaking, we can see the students usually watch the videos in order. As soon as a group finished watching a video, they often had short discussions about the just-watched video or the associated quizzes before starting the next video. If we look at a video navigation pattern for a single student in the DD condition of the NAS course (in Figure 8.5(b)) and compare it with the CC condition (in Figure 8.5(a)), the patterns do not differ distinctly: students in both conditions tended to watch videos one after another and few pauses or seeks occur in videos.



(a) CC sample in NAS

(b) DD sample in NAS

(c) CC sample in DSP

(d) DD sample in NAS

Figure 8.5: Sample video navigation patterns of the study groups. Each plot corresponds to the navigation patterns for a weekly session. The horizontal axis represents the session time (in seconds). The vertical axis corresponds to the position in the video (in seconds). Each video is associated with a specific color.

In fact, the pattern for the DD group of the DSP course (Figure 8.5(d)) is the one that stands out of the four patterns. Students in the group visibly interacted with MOOC videos a lot more than those from the NAS course. It should be noted that Figure 8.5(d) is not an exception, and the patterns in other weeks are largely similar. In fact, we received much feedback that our recruited DSP students thought the DSP course was hard to follow, whereas the NAS students generally perceived the NAS course to be easy. In addition, the professor of the DSP course did not require campus students to solve the quizzes posted on the MOOC website, and the content of the course was more advanced than the NAS. These factors constitute a strong course dependent effect that would break the homogeneity of the analysis. Considering only 2 DSP groups were recruited, we have less statistical evidence to prove a finding. Therefore, all the statistical analyses reported in this Chapter are solely based on NAS groups.

Video Interactivity

The plots presented previously visually illustrate how group students interacted with MOOC videos. In this section we attempt to quantify these patterns. As we know, video interactions consist of various types of actions, i.e. play, pause, seek forward/backward, each contributing to the total video watching time. Therefore, we define *time-spent-on-video index (TSOVI)* to gauge the level of interactivity. TSOVI refers to the ratio between the amount of time spent on watching videos in a week and the total length of video contents that are watched (not necessarily full videos). Possible values are theoretically any numbers that are above or equal to 1.0. Both pausing and rewinding videos result in an increase of this value: an index of 1.0 indicates that all the watched videos were played exactly once without being paused or re-watched, otherwise the students would have spent additional time and the TSOVI must exceed 1.0. In addition to the TSOVI, we also computed *pause frequency index* (PFI), which is the number of pauses per video minute in a study session.

The average TSOVI values for the CC and DD groups are 1.19 ($\sigma = 0.16$) and 1.22 ($\sigma = 0.13$) respectively; Meanwhile the PFI values are 0.19 ($\sigma = 0.11$) and 0.21 ($\sigma = 0.10$). It should be noted that the PFI and TSOVI are highly correlated ($r=.61$, $p<.0005$), meaning that the pauses strongly influence the time spent on videos of a group. In fact, both the TSOVI and PFI values are not large, indicating the students did not interact much with the videos, at least for the NAS course. Further, with a within-subject ANOVA, we did not find significant difference of group condition effects on either TSOVI ($p=.57$) or PFI ($p=.67$). Many social effects can explain the above non-significant result. For the CC groups, the single video control might make students hesitate to pause due to social pressures such as when and who should make the pause. On the other hand, the students in the DD groups perhaps pause generally less in order to stay synchronized with each other, as we will discuss in the next section.

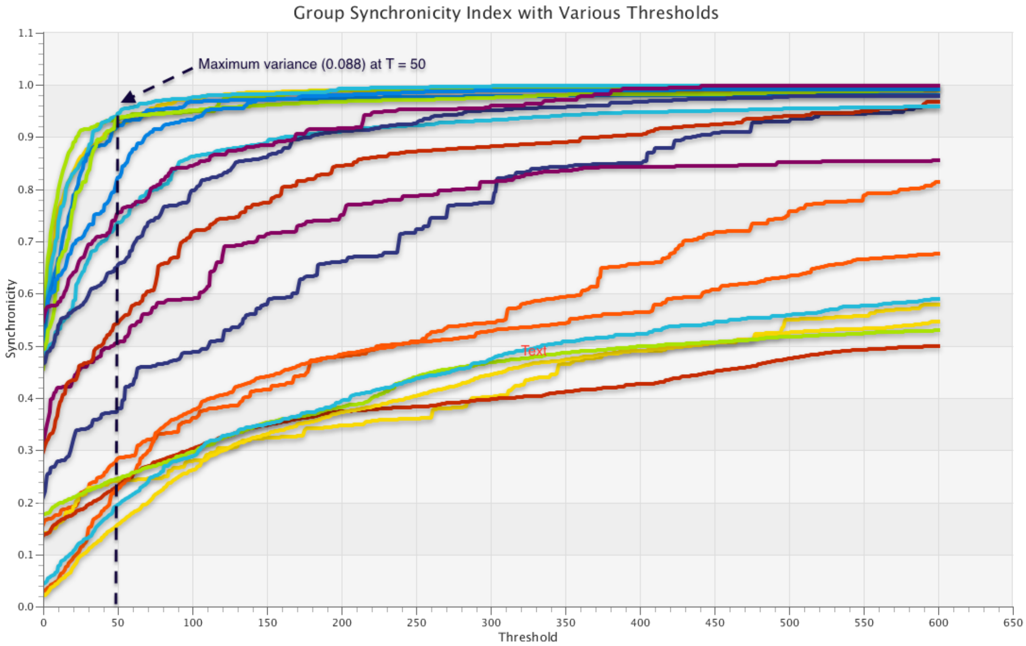


Figure 8.6: Computing group synchronicity index with varying thresholds

Synchronicity

The TSOVI and PFI quantify group's video interactions. For the DD groups, the students had the freedom of watching MOOCs at their own pace, but the setting also created some social incentive for them to stay "synchronized". In this section, we introduce another facet of group behaviors in the DD groups, i.e. synchronicity, which is denoted as how synchronously individual students in a group watched videos together. The synchronicity between two students in a group is obtained by computing the ratio between the total synchronous time and the length of the study session. Synchronous time means that the two students are either simultaneously watching the same video content or not watching any videos (e.g. they might be having a discussion). A threshold value **T** (measured in second) was introduced to determine the synchronous status. For each second of a study session we look **T** seconds ahead and behind to see if the two students used to be or would be watching the same video content within **T**. In other words, we are checking if one student catches up with the other in **T** seconds. If yes, then they are in synchronous state. We coined the term *individual synchronicity index* (ISI) to measure the average synchronicity between all pairs involving the same student. Each student in a group has a different individual synchronicity, which signifies how the student synchronized with other students in the same group. Another term *group synchronicity index* (GSI) is used to denote the average of all individual synchronicity values in a group.

How was **T** determined? Different **T** values lead to different synchronicity values. Figure 8.6 illustrates how synchronicity values for all groups in each week vary with different **T** between 0 and 600 seconds (10 minutes). As we see, the larger the **T** is, the larger the synchronicity

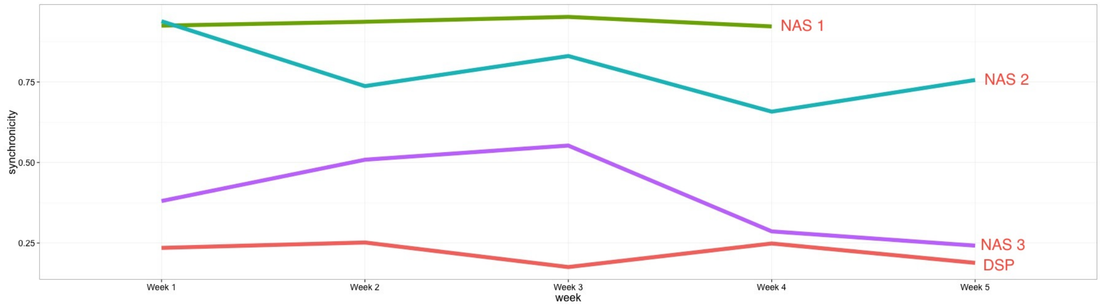


Figure 8.7: Changes in group-synchronicity index over weeks

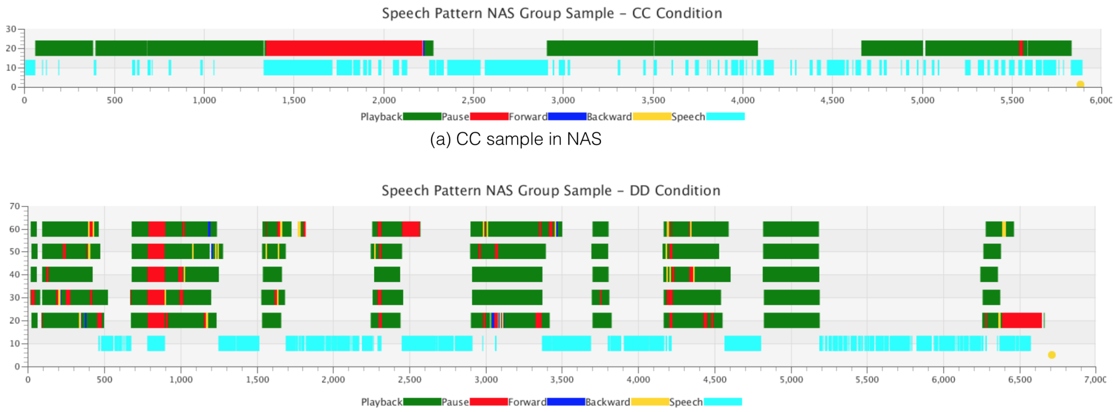
value. But the value may converge to 1.0 with very large Ts. A **T** value close to 0 only has theoretical meaning, because in reality we do not expect different people to watch the same video frame at exact same time. We chose **T** = 50, because this is where the variance among all possible synchronicity values of different sessions reaches maximum (0.088). It indicates that choosing this **T** value would maximize the differences among all the groups. In other words, the synchronicity patterns are most distinct under **T** = 50. The threshold of 50 seconds also makes sense in real world: a teacher usually explains the same concept within this period, so it is reasonable to say that students are synchronized on the same ground.

Synchronicity over Time

The GSI over the five weeks for each DD groups (including the DSP group) are shown in Figure 8.7. The fifth week's data for the NAS group 1 was missing due to a technical problem during the experiment. This figure shows that the GSIs are roughly stable over time. In addition, a clear gap among the synchronicity series is seen in the middle range of the vertical axis, which separates highly synchronized groups (NAS 1 and NAS 2 groups with GSI>0.65) from lowly synchronized ones (NAS 3 and DSP groups with GSI<0.6).

Variation in Synchronicity

While the group (i.e. average) synchronicities are roughly stable over weeks, individual synchronicity may vary with other factors. With mixed linear regression analysis, the TSOVI showed a significant negative effect ($\beta = -0.19$, 95% CI = [-0.301, -0.078], p<0.005, model R^2=0.9). This negative correlation is interesting to us, because it indicates that more video engagement time creates fewer opportunities for students in DD groups to keep synchronization. More time on videos implies that more interactions (e.g. pausing and replaying) occurred, which apparently makes it difficult for students to stay synchronized. Highly synchronous groups, according to our semi-structured interviews, reported that they usually noted down the problems during watching the videos, and discuss the problems right after everyone finished watching. The groups were self-regulated, and many students deliberately started and finished video watching more or less simultaneously.



(a) CC sample in NAS

(a) DD sample in NAS

Figure 8.8: Sample speech patterns of the study groups

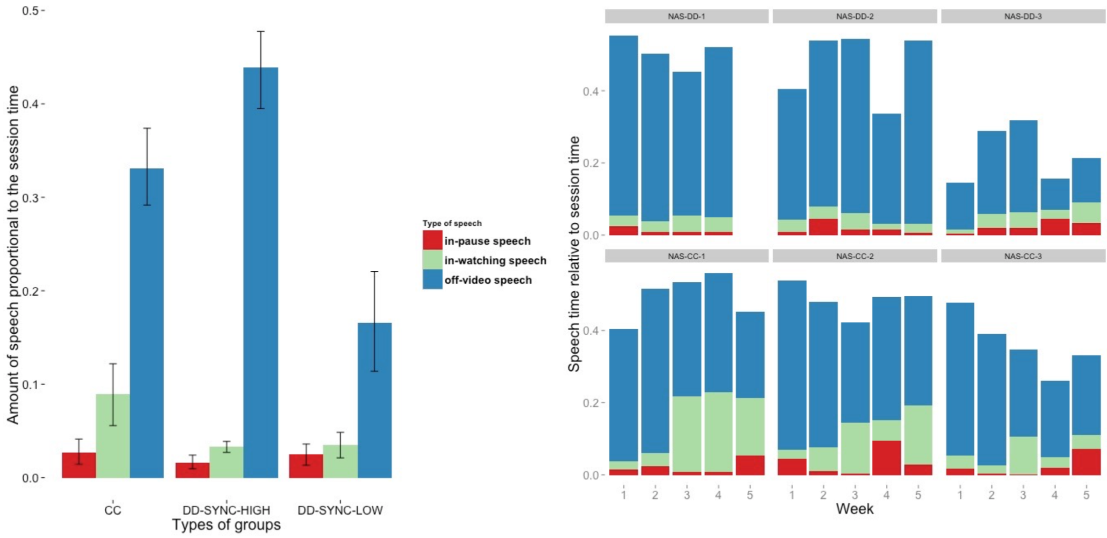
Amount of Speech

Previously we discussed the results regarding group video interactions, and identified study groups in the DD condition watched MOOC videos with diverse synchronicities. We also separate the DD groups into two categories according to their GSI. In this section, we shift our focus to group discussions by investigating the amount of speech in DD groups, with CC groups as baseline for comparison.

Speech Patterns

Figure 8.8 plots the amount of group speech alongside the video interactions for two representative group sessions in respective CC and DD conditions, so that we can have a visual perception of how discussions were distributed through a study session. As expected, both the CC and DD groups had talked most during video pauses or in the gap between two consecutive videos, but the CC group also sparsely talked during watching videos. In the discussion hereafter, we refer respectively to the aforementioned periods as *in-pause speech*, *off-video speech* and *in-watching speech*. The amount of each type of speech adds up to the total amount of speech in a study session. We measure the amount of speech at the group level, without differentiating whom it is from. The speech time is then divided by the total length of the corresponding session for normalization. The means of different types of normalized speech are plotted in Figure 8.9(a) with confidence intervals, and detailed patterns for each group over the five weeks are presented in Figure 8.9(b). In these graphs we plot the speech data from the NAS DD groups only. The separation between highly and lowly synchronized group is consistent to that described in the previous section (NAS 1 and NAS 2: DD-SYNC-HIGH, NAS 3: DD-SYNC-LOW).

We first compare the common characteristics of different types of speech in Figure 8.9(a). The off-video speech contributed the largest to the total amount of speech. This type of speech happened after a video was finished, when students jointly solved quizzes and problems



(a) Amount of speech among different conditions

(b) Amount of speech in each group over five weeks

Figure 8.9: Amount of speech in study groups

encountered in the video, both of which required group discussions. The fact that the amount of in-play speech was always larger than in-pause speech is partly due to the overall higher length of video-play time compare to video- pause time. In addition, CC students could deliver spontaneous speech without pausing the video, and students were likely to discuss in subgroups, when some students were still playing videos. Both factors may also contribute to a higher amount of in-play speech.

Next we investigate the differences in speech across conditions. The CC groups and highly synchronized DD groups resemble each other in large amount of total speech, whereas lowly synchronized DD groups overall talked noticeably less. It indicates that highly synchronized DD condition was similar to the CC condition in terms of total speech, and the difference was that the CC groups talked more during video watching and less after the videos while the highly synchronized DD groups resulted oppositely. This observation can be explained by the group setup: The DD students were wearing headsets during video watching, so that they could not talk easily while watching videos. The loss of discussion during video watching was instead compensated after the videos. Note that even the speech patterns within the same group vary over five weeks, and the variation is shown in Figure 8.9 (b). As for the comparison within DD condition, the interpretation of the bar charts for the lowly synchronized groups needs to be made with caution, since only one group is classified in their category. To get a better picture of the dynamics in the DD condition, we use statistical tools to quantify the patterns, which will come next.

Effect of Synchronicity on Speech

In Figure 5(a) we identified distinct patterns for lowly synchronized DD groups, which suggests

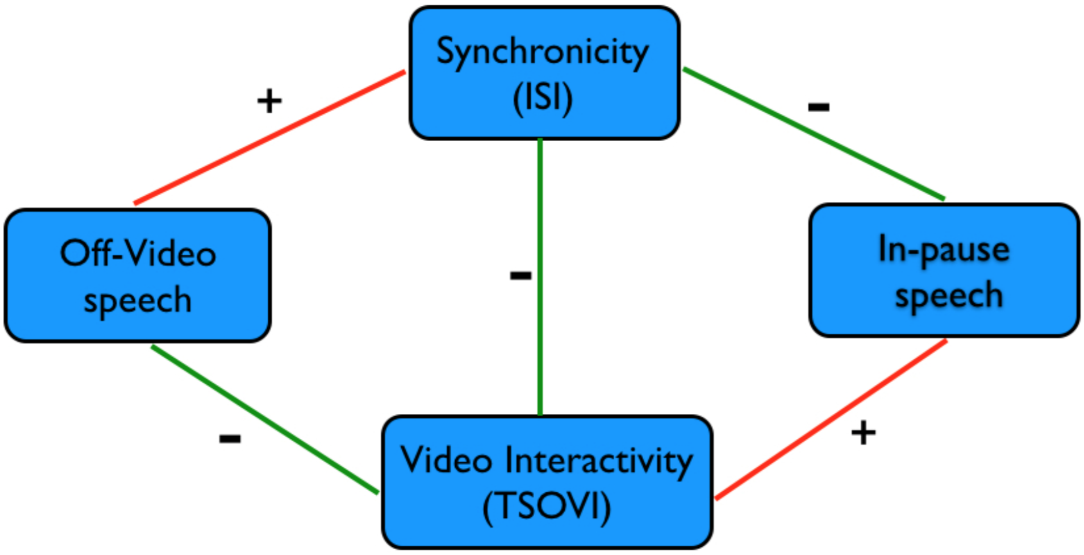


Figure 8.10: Relationship between video interactivity, synchronicity and amount of speech

a potentially systematic effect of synchronicities. With mixed linear regression analysis, we found that synchronicity (ISI) had a positive correlational effect on the amount of off-video speech (β=0.45, 95% CI = [0.319,0.621], p<0.0005, model R^2=0.84) and a negative effect on the amount of in-pause speech (β=-0.06, 95% CI = [-0.076, - 0.011], p<0.05, model R^2=0.21). This result suggests that more synchronized groups spent less time in pauses within videos, but more time after watching the videos. To complement the result above by relating the speech to video interactivity, we found that video interactivity (TSOVI) positively affects in-pause speech (β=0.05, 95% CI = [0.026,0.074], p<0.0005,model R^2=0.28), and negatively affects off-video speech (β=-0.113, 95% CI = [-0.202, -0.022], p<0.05,model R^2=0.82). This result is in line with the negative correlation between ISI and TSOVI we reported before. The interactions between the just-reported correlations are illustrated in Figure 8.10.

Figure 8.10 suggests that the more students in the DD groups engaged in video watching, the less synchronized they were, which in turn increased the amount of off-video speech and decreased the in-pause speech. In other words, highly synchronized students sacrificed video engagement for gaining synchronicity, resulting in more discussions during off-video periods. The amount of off-video speech seems to be balanced with that of in-pause speech: one aspect wanes, the other waxes. This effect is confirmed with marginal significance (p=0.07). However, it is difficult to interpret the result as the students talk more **about the videos** during video break if they talk less in video pause, since they also talk **about the quizzes** during the video break.

8.3.6 Understanding the Use of Textbook

Previous findings about group video interactions and group discussions have deepened our understanding about the primary activities in study group sessions. Up to now, we have not yet talked about the use of textbook. This section focuses on this issue by presenting how the groups in different conditions used textbook, before and after the introduction of the BOOC Player. Hereafter the term "textbook" includes both print textbook and digital textbook in the BOOC Player.

Coding Textbook Use

During the study sessions, students' interactions with the print textbook were not logged. The interaction logs recorded digital interactions on the BOOC Player, but the context of use, especially the occasions when the book anchored group discussions were also not captured. In order to gain deeper insights about textbook use in study groups, I coded 440 book interactions found from video recordings of 30 NAS study sessions (2-3 hours each), and identified the following textbook interaction modes:

Table 8.3: Textbook interaction modes. "Mode" is the type of book interaction. "Explanation" describes each mode in details. "Type of Book" refers to whether the corresponding mode applies to physical or digital book. "Scope of Interaction" indicates the interaction can be performed individually or collectively.

Mode	Explanation	Type of Book	Scope of Interaction
Browse	Turning the print book rapidly to look for relevant pages	Physical	Individual
Glimpse	Glancing at the book to follow the video (mostly) or quiz. Such interactions typically lasted shortly, with eyes quickly jumping back-and-forth between the book and the video / quiz sheets	Physical	Individual
Read	Resting the eyes on the print book page for longer time to read text	Physical	Individual
Turn	Turning pages on the PDF book	Digital	Individual / Collective
Scroll	Scrolling the PDF page to view its different parts	Digital	Individual / Collective
Zoom	Zooming in / out to see details in the PDF book	Digital	Individual / Collective
Talk	Talking to other members with the book, sometimes with pointing gestures to the book	Physical / digital	Collective

The above interaction modes can be identified from the videos with the help of both front-view and rear-view recordings. Among these modes, *Browse*, *Glimpse* and *Read* are only performed **individually** and apply only to the physical print book, according to our video coding scheme. In principle, *Glimpse* and *Read* also apply to the PDF. However, since the video and PDF are displayed in parallel, we cannot tell if a student's eyes are rested on the video or the textbook from the video recordings. The modes of *Turn*, *Scroll* and *Zoom* are exclusively associated with the embedded PDF in the BOOC Player and they can be performed both individually and collectively. By interacting **collectively**, we mean an interaction is performed with conscious awareness of multiple group members, e.g. when a student zooms a PDF page for shared interests in the book. *Talk* is a collective verbal interaction by nature.

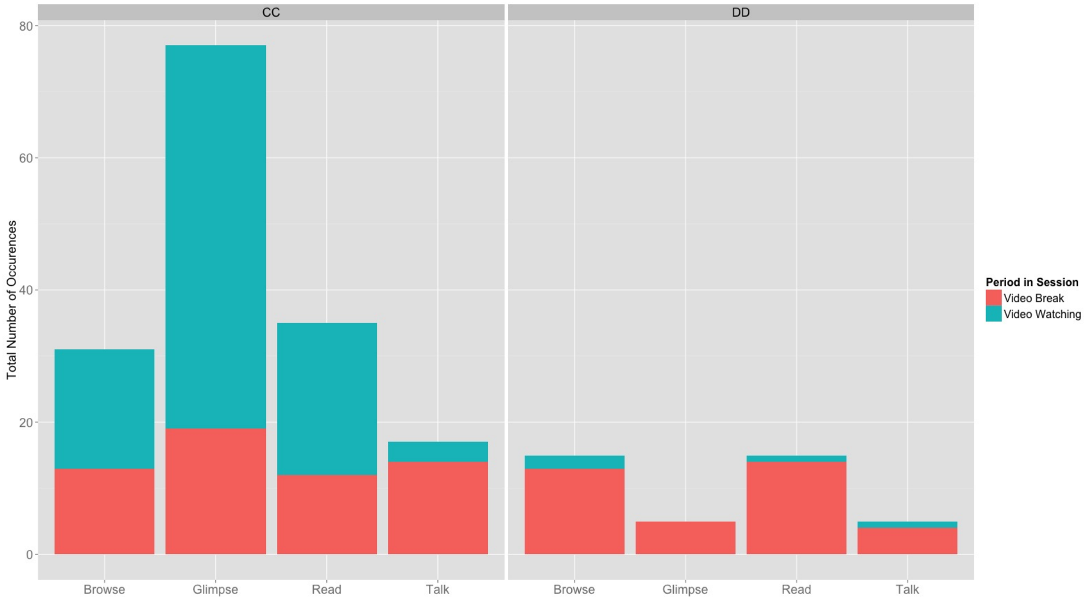


Figure 8.11: Frequencies of the print book interaction modes in pre-BOOC sessions

Textbook Use Before the Introduction of BOOC Player

In the first three weeks, the groups watched lecture videos on normal table video players, and they were required to place their textbooks on the table. The book was not essential for studying the course, so they could use it on a voluntary basis. Were the print textbooks really used and how often were they used? Do the uses of textbook vary in group conditions? We present results regarding these issues. Not every group used the print textbooks in every session, and our goal was not to predict the book use. Rather, our interest lies in book interaction modes.

Distribution of Interaction Modes

As we learned from the group video navigation patterns presented in Section 8.3.5, the study groups watched videos one after another, and they usually reserved a certain period of time before starting the next video for discussing about the just-watched video or the associated quizzes. We refer to this period of time as *video break*, as compared to *video watching*, which is the time when students were engaging in watching videos. Video watching includes both the playing and pausing of videos. Figure 8.11 illustrates the frequency of occurrence of the book interaction modes for the DD and CC groups in the first three weeks. We find that the time period when the book interactions occurred differed significantly between CC and DD (χ^2(200,1)=34.98, p < .0001, ϕ = 0.43). The CC groups had more balanced use of textbook between the *video break* and *video watching* periods, whereas the DD students mostly used the textbooks only during *video break*. This difference was especially notable for the *Glimpse* action. Several occasions were found, where the CC students watched videos with the textbook open and they shift their attention back-and-forth at times between the video and the book. None of the DD students had ever done so. A possible explanation on this matter could be that the DD students were afraid of breaking the video watching synchronicity by interacting the book, since every student was watching MOOC videos in his/her own private space without being aware of others' activities.

Overall there were 20 occasions when the students grabbed the book on the table and looked for relevant content while watching videos (*Browse*). On average 32.5 seconds were spent on each *Browse*. However, not every *Browse* was effective, we found in 3 out of the 20 *Browse* situations, the students failed to find the intended information on the textbook. For the *Talk* mode, we found a student either asked questions to the others with reference to the textbook or read aloud the book content to the whole group. *Talk* was the only collective interaction mode during this period. It accounted for respectively 10.6% and 12.5% of the total number of interactions in CC and DD conditions and seldom (17.7% and 20.0%) happened during video watching.

Reasons for not Using the Print Textbook

We expected group students to turn to the textbook whenever they encountered problems. In fact, they did not alway do so. During the weekly semi-structured interviews, we attempted to understand why some of the students did not use the book. The main reasons include:

1. They were afraid of loosing time in looking for information in the book
2. They did not know exactly what is not clear while watching videos
3. The lectures were easy, so the videos were sufficient for comprehension
4. They preferred to ask in the group first, which usually solved their problems

Feedback (3) is an intrinsic contextual factor. The students had no problems with the MOOC videos, they would not turn to help of any kind. Feedback (4) is an extrinsic contextual factor. The students had an option to discuss with other group members for solving problems and clearing doubts, which is the goal of study groups. According to the Principle of Least Effort (cf. Chapter 3), the students tended to believe group discussions required less effort than textbooks for resolving problems. Feedbacks (1) and (2) actually confirmed the potential needs for peripherally displaying book references, so that the students can quickly judge the usefulness of book content without loosing time to browse and find information first.

Textbook Use After the Introduction of BOOC Player

We deployed the BOOC Player in the last two weeks of the study, but students were still asked to bring the print textbooks during the study sessions. In fact, only one student had used the print book in the final two weeks of the study, because the digital book was less tangible than

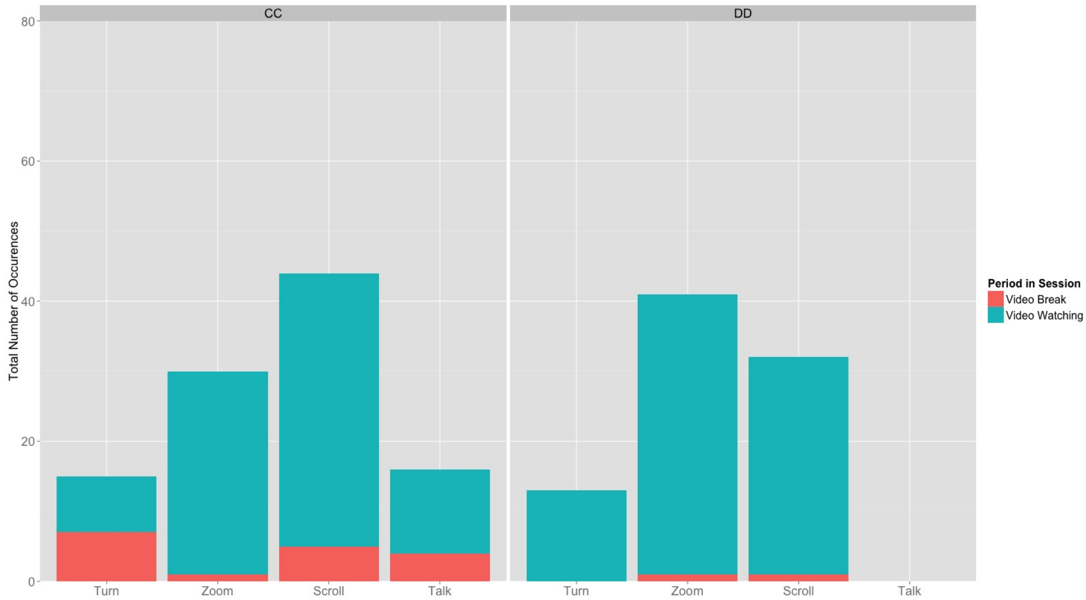


Figure 8.12: Frequencies of PDF interaction modes in BOOC sessions

a paper book. Nevertheless, the PDF mapping in the BOOC Player assisted this student to quickly navigate to the intended pages.

Reported Use Scenarios

To analyze the use of the peripherally displayed synchronized PDF in the BOOC Player, we first summarize the usage scenarios collected from the questionnaire:

1. **Extended Knowledge**: when the students saw another explanation of a concept with detailed theorems and examples;
2. **Alternative Presentation**: when the teacher was talking too fast or the videos were not visibly clear;
3. **Information Confirmation**: when the students had doubts about certain concepts and need to confirm their understanding;
4. **External Help**: when none of the group members knew the answer or when they were arguing about certain concepts.

Among these reported usage scenarios, the advantage of the peripheral design is especially notable in (1), where the students had no explicit needs of help. The synchronized textbook pages "emit" information scents pertinent to the video being viewed so that the students serendipitously encountered certain pieces of information that are helpful.

Distribution of Interaction Modes

Figure 8.12 depicts the frequency of occurrences of book interactions modes with the BOOC Player. It is not surprising to see that in both conditions, the interactions predominantly happened during video watching rather than during video break due to the book-mapping feature. What is more interesting is that the DD students used the book significant more ($F(1,51)= 12.2$, $p < .001$) during video watching, which seldom happened before. A possible explanation could be the synchronized PDF increased the visibility and accessibility of the potentially useful information in the book. As a result, the students were offered better opportunities to address their situational needs without the fear of loosing synchronicity.

For the students in the CC condition, the most notable change after the introduction of the BOOC Player is the increased occurrence of collective textbook interactions. During this period, not only *Talk*, but also *Turn*, *Scroll* and *Zoom* could be collective. The proportion of collective interactions doubled to 20.9% compared to that of 10.6% in the first 3 weeks. Even if we count the *Talk* interactions only, the proportion increased to 15.2%, with 75% happened during video watching, as compared to 17.7% before. A mixed model ANOVA with student nested in groups as random effects shows that the BOOC Player has significantly increased *Talk* interactions during video watching ($F(1,56)= 63.8$, $p < .00001$). The reason behind the increments, we believe, is that the shared display of synchronized book content increased shared attention, so that the students could have more chances to collaborate with the book.

8.3.7 Addressing Videos Difficulty and Facilitating Group Discussion

So far we have explored several facets of group-based MOOC learning behaviors, including how the students interacted with MOOC videos, textbooks as well as other group members. We found that synchronicity is a key concept for the DD groups, which significantly relates to the distribution of group speech and video interactivity. The introduction of the BOOC Player largely increased (1) collective interactions with textbook (the *Talk* mode) for the CC group (2) individual textbook interactions during video watching for the DD group. These findings are themselves interesting, but further investigations are required to understand their impacts: Do the amount of speech, synchronicity and textbook use relate to video difficulty, discussion quality and equality of contribution among group members? This section serves to answer this question.

Video Difficulty

Factors that relate to video difficulty may include the session week, frequency of textbook use, TSOVI, proportion of speech time during video pause and video break, respectively. For DD groups, synchronicity is a potential factor as well. We built mixed linear multi-regression models with the aforementioned variables as covariates and the 5-point video difficulty ratings obtained from the questionnaire as outcome variable, the student nested in group were modeled as random effects. Backward elimination was used to remove non-significant covariates as well as those lead to multicollinearity.

For the DD groups, the proportion of speech time during video pause (β=15.2, 95% CI = [-1.0,29.4], p=.05) and video break (β=5.5, 95% CI = [-1.1,9.3], p<.01) are significantly correlated with video difficulty. The model R squared is 0.69. This result indicates video difficulty in the DD study groups can be reflected by the speech during pauses or breaks. The more they talk during these periods, the more the videos are difficult for them.

For the CC groups, we did not have enough evidence to prove any of the listed potential factors correlates with the perceived video difficulty. It may seem surprising that the amount of speech during video break or video pause does not shown a significant effect. However, in the CC condition, students talked at any time, even during watching videos. Difficult problems did not have to be discussed in a specific time period.

Quality of Discussion and Equality of Contribution

A similar process for constructing mixed linear models was employed to relate behavioral factors to two subjective measures of group discussion respectively.

For the DD groups, we found that the more synchronous a group was on watching MOOC videos, the higher they perceived the discussion quality (β=1.308, 95% CI = [0.263,2.546], p<0.05, model R^2=.45) as well as equal contribution (β=1.438, 95% CI = [0.386,2.499], p<0.05, model R^2=.35). Possible reasons could be more synchronous groups had more chances to discuss during video break, which might have positive influence on both the quality and equality aspect of discussion. This result indicates that synchronization is a desired attribute of DD study group.

For the CC groups, the proportion of speech during video break is shown to correlate positively to the equality of contribution (β=2.24, 95% CI = [-0.34,4.57], p=.07, model R^2=.68). Note that the correlation is significant at $\alpha = 0.1$. The same variable (β=5.58, 95% CI = [2.28,8.50], p<.001) together with the frequency of *Talk* interactions (β=0.24, 95% CI = [0.05,0.42], p<.05) during video watching both significantly correlate with the discussion quality. The R squared for the latter model is 0.48. The CC groups could talk at any time. Why does the proportion of speech during video break relate to the perceived quality of discussion and equality of contribution? A possible explanation could be discussions during video watching were usually specific, and students might still feel distraction. In contrast, talking during video break could be more thorough and all group members could participate without worrying about distractions to watching videos. The significance of *Talk* interactions during video watching on perceived discussion quality is of special interest to us. In the previous section we have shown that the introduction of BOOC Player has significantly increased the *Talk* interactions. Combined with this result, the increment in Talk interaction in turn enhanced discussion quality. A possible reason is that group students can easily refer to highly relevant and reliable complementary content in the book to support the discussion.

8.3.8 Discussion

In this section, we discuss how the results presented in the preceding sections answer the two research questions posed at the beginning of this chapter, which respectively concerns the interaction dynamics of study group and the role of the BOOC Player.

Interaction Dynamics of MOOC Study Groups

Through this study, we find that watching videos synchronously is central to MOOC study groups. In the DD condition, where the students were allowed to watch videos independently, we observed a clear cut in the groups' synchronicity. Some groups chose to stay highly synchronized by "sacrificing" the freedom of individual interactions with the video. These highly synchronized groups discussed significantly more during video breaks as compared to the lowly synchronized groups, and were on par with the condition with a centralized display (CC condition). In other words, less individual video engagement and less in-video discussions associate with higher synchronicity, which in turn relate to a high quality of discussion and more balanced participation perceived by the students. The associations between the synchronity and the subjective ratings do not seem to be direct. A potential confounding factor is the proportion of speech during video breaks, which also positively correlates with the synchronicity. Highly synchronized groups actually had more time to initiate a more thorough discussion after finishing a video. In addition, students who perceived higher difficulty in videos also tend to talk more during the video break, indicating talking after videos was a common way to resolve difficulty. These observations indicate that synchronously watching videos empowers learners with a sense of being in a team. On the other hand, interacting too much with the video promotes individualism as the learners are only concerned with their own learning rather than share and validate their understanding with others; and this idea also goes against the theme of study-groups.

In the CC condition, the study groups were naturally synchronized in watching MOOC videos. In this situation, we did not find the proportion of speech during video breaks significantly correlate with perceived video difficulty. A potential reason is that the students were able to talk to each other at any time when encountering a problem during a study session. Similar to the DD condition, the proportion of speech during video breaks also positively associates with discussion quality and equality of contribution. The potential reasons might be similar as discussed before, i.e. talking after watching videos allows more thorough discussions.

Role of the BOOC Player

When we discussed the role of the TileSearch system in Chapter 6, we mentioned contextual information scents in learning activities may carry 3 different roles: (1) inducing serendipity; (2) facilitating group discussion; and (3) servicing the learning goal. Specific to the information scents rendered by the synchronized textbook in the BOOC Player, the three roles are crystallized as: (1) anchoring occasional interactions with the digital textbook; (2) improving discussion quality or equality of contribution, and (3) helping students with problems.

In the CC condition, the BOOC Player definitely induced serendipity. We found the BOOC Player drastically increased group discussions with references to the textbook during video watching, which in turn significantly enhanced group discussion quality. This finding reveals the effective role of the BOOC Player as group facilitator. As highly relevant, complementary information was constantly displayed in the periphery of the group students during video watching, they could easily refer to specific content to support spontaneous discussions. This way, the BOOK Player increased mutual awareness of the information. However, as mentioned by several subjects, the textbook was not their primary resource for seeking help. Significant correlation was not found between the uses of digital book and perceived video difficulty.

In the DD condition, the students almost never used the print textbook while watching videos in the first three weeks. Some students argued that looking up in the textbook was time consuming. So, they might be afraid of loosing synchronicity with others. The BOOC Player significantly increased the frequency of textbook use during video watching in the sense that it provided situational help to the students without loosing much synchronicity. They could consume the information immediately, rather than looking for it in the first place. However, the role of the BOOC Player for the DD groups remains at the level of inducing serendipity. We did not have significant evidence that showed interactions with the BOOC Player enhanced group discussion quality or equality of contribution. Moreover, as discussed before, the DD groups tended to address video difficulties by discussing during video breaks, not through turning to books for help. Several students reported that they might use the BOOC Player to address difficulties when study MOOCs alone at home.

Limitations and Prospects

The current design of contextual information scents in the BOOC Player has played a positive role as serendipity inducer and group facilitator in specific conditions. The information scents were highly relevant and they were both calmly and timely presented. However, the BOOC Player only displayed information from a single resource and in a single form, i.e. the book pages in the textbook. The display of the textbook took too much screen real-estate on the iPad display, and this issue was complained by several participants. Future systems may consider to incorporate other forms of resources together with the textbook. Small pieces of information can be extracted and presented to the students as contextual information scents, so as not to take up much of the limited space.

In the current BOOC Player implementation, we enabled dual mapping between the PDF and the video. During the experiment, linking from book pages to videos were never intended, since watching videos was their main activity, not reading books. The students were sometimes annoyed due to abrupt video changes when they accidentally swiped the PDF to a page with a different video mapping. We argue dual mapping might be useful when used at home than in time-bounded group study sessions, and this aspect could be explored in future work.

8.3.9 Conclusion

This chapter presents a design prototype of contextual information scents based on video content for MOOC study groups. Textbook pages with supervised mappings rendered high quality contextual learning materials for the students.

Up to this chapter, we have presented four design prototypes that attempted to create information scents with three different types of contexts and experimented them all in collaborative learning scenarios. The BOOC Player performed best amongst all of them so far. It was used more often by the subjects than the previous systems. In specific conditions, the BOOC Player was found to have played an effective role either for inducing serendipity or for enhancing the quality of group discussions. Unfortunately we did not prove its effectiveness in helping students resolve difficulties. However, we should note that for most of the time group discussions were considered as the first choice when students encountered difficulty in this collaborative MOOC learning scenario. In other words, there was a *zone of proximal development* in the study group, and the students usually tended to achieve learning through guidance from a *more knowledgeable other* (MKO) rather than from learning resources, because the former was considered as requiring the least effort (principle of least effort) in this regard.

Compared to collaborative MOOC video viewing, a more representative scenario of MOOC learning for learners is to follow a course online and study alone for most of time. Clearly, a "live" MKO is missing in this scenario. But can technologies play the role of a MKO to help students resolve difficulties? This would require us to understand the learners' online learning behaviors, especially video interaction behaviors. For example, what video interactions of a learner may indicate s/he is experiencing difficulty? In the next chapter, we will dig into this issue.

9 Evaluating Zone of Intervention

The BOOC Player presented in the preceding chapter demonstrated the potential strength of designing contextual information scents with an elicitation of textbook content for collaborative MOOC video viewing activities. However, the study groups have not used the textbook as primary means to address video difficulty. Instead, discussing in groups was a preferred solution, and the synchronized book content in the BOOC Player were only used sometimes to service group discussions. It is impractical to change the students' difficulty tackling strategy, since asking a *more knowledgeable other* in the group was shown to be a clear winner in terms of convenience. In this chapter we continue investigating the MOOC scenario, but with a shifted focus on learning online without the presence of groups. When individual students learn MOOCs, we believe their video interactions create *zone of intervention*, the notion of which was discussed in Section 3.1.1. Originally this term was used in information seeking research, to refer to the "area in which an information user can do with advice and assistance what he or she cannot do alone or can do only with difficulty" (Kuhlthau, 2004). We adapt the definition to refer to the occasions when students have potential needs for help, so that contextual information scents rendered by various forms of learning materials may intervene. (Kuhlthau et al., 2007) posit that technological interventions outside the *zone of intervention* may be unnecessary, and are likely to be perceived as intrusive and overwhelming by the students. The key issue to be explored in this chapter is the identification of such zones in MOOC learning, so that proper interventions (e.g. in terms of contextual information scents) can be imposed timely.

Unlike the previous chapters, this chapter is not concerned with a new design prototype. Instead, we attempt to make inferences from large-scale MOOC learning dataset[1], in order to understand (1) the video interaction features that reflect students' perceived video difficulty; (2) the video interaction patterns emerged in MOOC learning as well as their impact on students' performance. The findings will provide insights for us to evaluate the *zone of intervention* for contextual information scents.

[1]The content in this chapter has been published in different research venues, and publications [3,5] in publication list (last page in the thesis) are concerned.

9.1 Background

MOOC learning experiences offered by popular platforms such as Coursera and edX are enabled by a combination of learning resources, such as lecture videos, quizzes, tutorials, discussion forums and Wikis. Complete pictures of how students learn through the online platforms can be rendered by investigating the use patterns of these resources, which has led to the rise of MOOC analytics in recent years. With tons of learners taking courses, MOOC analytics is making a big leap forward. Research interests have been centered around social engagement in discussion forums (Brinton et al., 2014), video engagement (Kizilcec et al., 2013), performance (Jiang et al., 2014), demographics (Guo and Reinecke, 2014), video interactions (Kim et al., 2014), dropout prediction (Halawa et al., 2014; Sinha et al., 2014b), just to name a few.

Although students may interact with various learning resources, videos remain as the primary media for the delivery of learning content, which has made video viewing the central MOOC learning activity (Breslow et al., 2013; Seaton et al., 2014). Students play, pause, or seek in videos to study at their own pace. This is a natural and self-regulated learning process. As thousands of students interact with MOOC videos, we can plausibly find meaningful patterns that yield a closer look at how students learn through videos. In this section we first make a brief review of research literature on in-video interaction analysis. This is followed by a special focus on recent work about video interactions in MOOCs.

9.1.1 Video Interaction Analysis

Videos players typically offer a limited types of interactions, each of which is associated with a time span. The sequential execution of the actions entail the Markov model a popular approach for video analysis. In early research, such analyses mostly aimed at evaluating the quality of service issues (Dey-Sircar et al., 1994; Li et al., 1996; Shenoy and Vin, 1995). Research that attempted to model video click behaviors came to light since Branch et al. (1999) found that video interaction behaviors, in terms of the time spent on each viewing mode (i.e. play, pause, fast-forward, fast-rewind) can be modeled with lognormal distributions. The authors also proposed a first-order Markov chain model for modeling different types of actions. Later, Syeda-Mahmood and Ponceleon (2001) studied subjective video browsing states with a Hidden Markov approach, with the goal of generating video previews that best represents interesting video segments. All of the above studies were conducted in the time when the control menu of video players were restricted to only continuous interactions, lacking discontinuous interactions that are common in modern video player, such as seeking forward/backward, which allow jumping between different time positions.

Research on clustering video interaction behaviors also started before the MOOC era. In the early 2000s, Mongy et al. (2007) proposed a method to apply K-means clustering with the Kulbach-Leibler distance between the state-transition matrices, but little is discussed about the data collection, the validation of the results and the scalability of the approach.

9.1.2 MOOC Video Interaction Analysis

Compared to the interaction possibilities offered in traditional video players, MOOC video players additional feature with speed controls that allow adjusting the video play rate. Analyzing video behaviors has received more and more research attention recently. One typical type of MOOC video interaction analysis is predicting course dropout in MOOCs. Sinha et al. (2014a) turned video interaction into click sequences, and performed n-gram analysis to predict students' dropout. One of the limitations of the n-gram approach is that it did not consider the duration of each action. Sinha et al. (2014b) also combined video and forum interaction footprint of students, and used a graph-based approach to extract MOOC participation features to predict student attrition. The prediction result was shown to outperform the n-gram approach.

Researchers also analyzed MOOC video interactions in order to gain understanding about specific video interaction behaviors. Kim et al. (2014) found that students are more likely to dropout a video when the videos are not watched for the first time, or the videos are tutorial videos rather than lectures. Longer videos are also shown to be associated with higher in-video dropout. In addition, the authors exclusively studied temporal interaction peaks, which are sudden spikes observed in aggregated video events per video second. The interaction peaks only considered play, pause and seek events, but speed changing interactions were not taken into account.

Another typical MOOC video analysis is clustering. For example, Kizilcec et al. (2013) adopted K-means method based on the students' longitudinal online learning activities to categorize MOOC student's engagement trajectories. The study concerns the number of videos watched as well as the navigation sequence. To our best knowledge, research gaps remain in clustering video behaviors with click-level interactions.

9.2 Research Questions

Our research emphasizes click-level video interaction analysis, which aims to render a closer examination of how a student interacts with each video lecture, e.g. what types of video interactions are employed, when they happen and how intense they are. We assume the video interactions reflect students' learning states, e.g. encountering difficulties, being confident, so on and so forth. An inspection of the video interactions may allow us to infer these latent states, creating zone of intervention. The key research questions to be answered in this chapter are:

(1) How do video interactions of different types and intensities reflect students' perceived difficulty?

MOOC student may encounter problems from time to time during video watching, and we anticipate the students to adjust their video interactions accordingly. For example, they can pause the video to think or search information on the Internet. In case of confusion, difficult parts of the videos can be re-watched. Increasing or decreasing video speed may also serve for specific purposes. Understanding the effects of each type of interaction is a preliminary step to investigate more complex video interactions.

(2) Can we categorize video interactions into groups of similar patterns?

Thousands of MOOC students in each MOOC interact with the same videos. Therefore it is likely that video interactions can be categorized into groups of similar patterns, which describe how the students typically use MOOC videos to achieve their learning goals. We are interested in identifying students' video interaction patterns and investigate how these patterns relate to video difficulty, video revisiting behaviors as well as performance. The analysis would render a more comprehensive evaluation of *zone of intervention* for contextual information scents.

9.3 Investigating MOOC Video Interactions

To answer the two research questions posed previously, we analyze video interaction datasets from two MOOCs offered by our university at Coursera: The *Reactive Programming (RP)* took place in the autumn of 2013, it covered advanced topics in programming with the Scala language; The *Digital Signal Processing (DSP)* is a foundation course for Electrical Engineering students. The MOOC we analyzed in this chapter was a more recent edition of the same course presented in Chapter 8. It was offered in the spring of 2014. Both courses had similar presentation styles, i.e. professors present the lecture with PowerPoint slides, holding a digital pen as both pointer and annotator.

9.3.1 Video Interaction Datasets

The following table summarizes descriptive information of the datasets from the two MOOCs. "Subject" refers to the course subject. "Week" is the total duration of the course. "Videos" is the number of videos posted. "Length" is the average length of all videos. "Quiz" is the total number of quiz sets. "Active" is the number of students who watched at least a video. "Passed" is the number of students who passed the course. "Sessions" is the total number of video play sessions logged in the video clickstream. "Events" is the number of video events in all video sessions.

Table 9.1: Overview of the two MOOCs in our dataset.

Subject	Week	Videos	Length	Quiz	Active	Passed	Sessions	Events
RP	7	36	18:50	6	22,794	5,276	470,994	4,001,992
DSP	10	58	16:20	16	9,086	263	117,959	1,138,558

The RP course attracted three times more active students than the DSP. There were no mid-term or final exams in both courses. Instead, assessments of students were made with weekly quiz sets. Students' grades were computed as the sum of their best quiz scores of all trials in each week. The RP course allowed an unlimited number of quiz submissions, whereas the DSP course permitted five submissions per quiz at maximum. As a result, the RP students made more attempts to achieve better grade, which perhaps further led to higher pass rate in the RP course (23.15 %), compared to only 2.89% in the DSP course.

9.3.2 Data Wrangling Pipleline

The raw video events were logged in the clickstream data in JSON format. Before proceeding to data analysis, we developed a data wrangling tool to reconstruct the watching histories of each student. For each unique video in our dataset, user-based watching histories were created by arranging the events in chronological order. The events were separated per video play session for each student. Next, we aggregate these events in each video play session and compute a set of video features that quantified the students' interactions, which will be explained in detail in later sections.

The processed video events include pauses, seeks and speed changes. In fact, Coursera video players do not only generate pauses when a user clicks the pause button. Automatic pauses are generated when an in-video quiz pops up or when the video progresses to the end. Such automatic pausing events are removed for the analysis in this paper. In addition, students usually watch the lecture videos in uncontrolled environments, so the pauses are found to last for a maximum of several days. We removed the pause events that have a duration of more than 10 minutes, which are rather "breaks" than "pauses". Seeking events are usually created when the user clicks or scrubs the playhead to a new position on the time bar. When scrubbing interactions occur, the logging system automatically generates a number of intermediate seeking events.

Many students left in the middle of the videos, leading to the so-called in-video dropouts. During the period when the two MOOCs took place, it was not guaranteed that each time when a student left a video was successfully logged. We also removed or corrected data entries containing inconsistent timestamps or event types, e.g. the case that two consecutive pauses at different time positions is considered as an logging error.

9.3.3 Methodology

The goal of investigating MOOC video interactions in this dissertation is to evaluate the zone of intervention, which in our opinion requires certain quantitative measures to gauge. Student performance in terms of scores is an example of objective measure, and lowly achieved students perhaps need more help. However, performance scores did not capture the learning process, so they are more of a kind of summative evaluation. In the MOOC context, students watched videos one after another. It is reasonable to assess the perceived difficulty for each video for students, so that the zone of intervention can be detected earlier and more contextual.

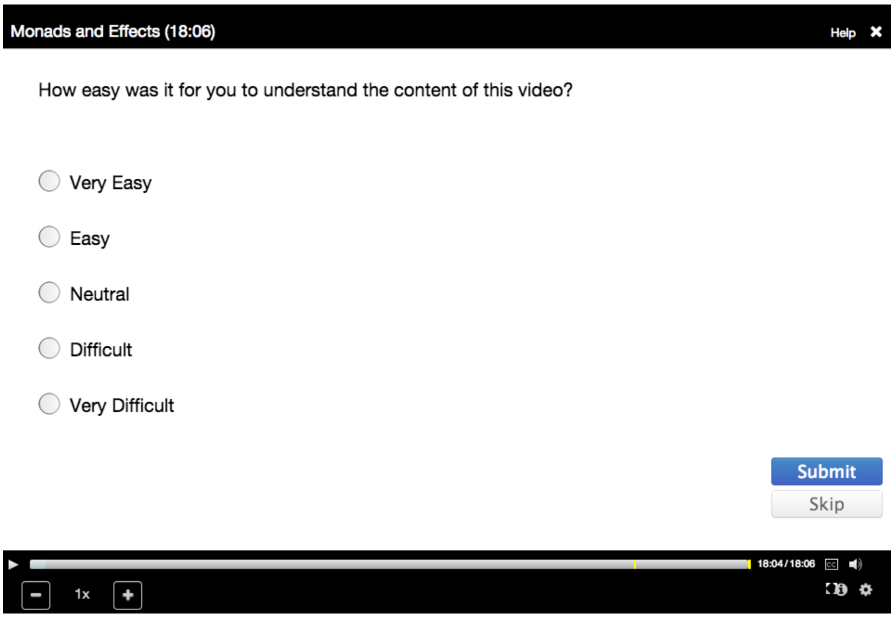


Figure 9.1: Example end-video survey about perceived video difficulty

In order to measure the subjective video difficulty, an in-video survey is placed at the end of each video during the enactment of the courses (cf. Figure 9.1). Only one question was asked: How easy was it for you to understand the content of this video? These surveys are posteriori evaluations that were answered by the learners right after they finished watching the video, providing ground-truth knowledge that allows us to reveal the hidden relationships between the video interaction and the perceived video difficulty. The surveys were not graded, so the students participated voluntarily. The responses were then coded with integer values from 1 to 5 to represent the difficulty ratings from "Very Easy" to "Very Difficult". Students may watch the videos multiple times and leave more than one ratings for the same video. The average video difficulty for first-time and revisiting video sessions are respectively 2.699 and 2.837 for the RP course, and 2.478 and 2.593 for the DSP course. Revisiting video sessions were clearly rated more difficult than first watching sessions. In the analysis of this section, we will only focus on the rated video difficulty of the first-watching sessions. The response rate for the RP course (188,138 sessions) is 79.0%. For the DSP course (28,994 sessions), the rate is 60.8%.

9.3.4 Inspecting Video Interaction Features

In response to the first research question posed in Section 9.2, this section aims to deliver an understanding of the relationship between different types of video interactions and perceived video difficulty. This implies the need to extract video features for each type of interactions. In this section, we first partition the datasets into video interaction profiles, and then analyze video features in each profile separately.

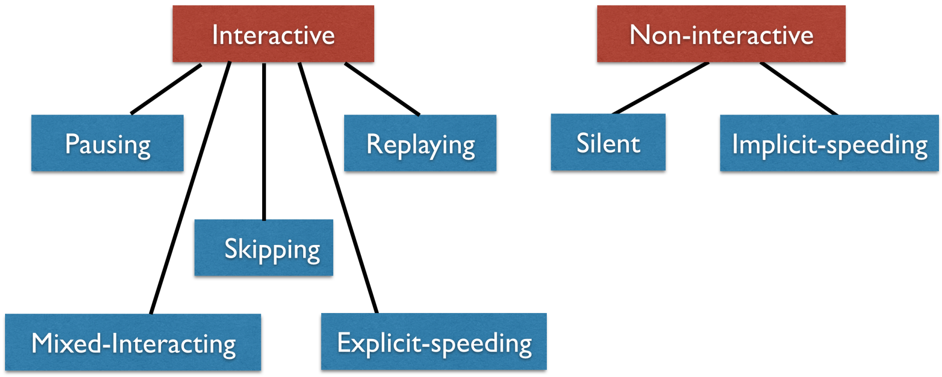


Figure 9.2: Schematic diagram of video interaction profiles

Video Interaction Profiles

Coursera offers four types of video controls, namely, play/pause (toggle), seek forward, seek backward and adjust video speed. In addition, the video players are found to consistently maintain video speed across videos. If a user changes the speed of a video, then the new speed is inherited for all subsequent video sessions. Therefore, a video may be streamed with varying initial playrates at the start of a session. Based on the types of interactions in the video sessions (profiles), we divide the dataset into subsets. For video sessions start with the default video speed (1.0), each of the four video controls is associated with a unique video interaction profile, which we name as *pausing*, *skipping*, *replaying* and *explicit-speeding*[2] respectively. Video sessions without interaction events but start with the default video speed are called *silent*, whereas those start with higher or lower video speeds are called *implicit-speeding*. The rest video sessions are categorized as *mixed-interacting*, as they combine different profiles of video interactions. The partition scheme can be visually illustrated with a schematic diagram in Figure 9.2. *Silent* and *implicit-speeding* video sessions are ***non-interactive***, because they do not contain any video interactions, others are ***interactive***.

An overview distribution of the 7 video interaction profiles for two courses are presented in Table 9.2. Each profile for each course corresponds to two cells in the table, which presents respectively the proportion of video sessions in the dataset and the average perceived difficulty. It is necessary to stress once more that we only consider single type of video interaction when defining the profiles, except *mixed-interacting*. For example, the *replaying* sessions contain only backward seeks but no other interactions.

For both courses, nearly half of the video sessions contain more than a single type of video interactions (*mixed-interacting*); Around one fifth of the video sessions (*silent*) contain no interactions at all; Pause interaction (*pausing*) is most frequently employed by students during video watching. Without statistical significance being concerned, Table 9.2 shows that

[2] In publication [5], explicit-speeding also includes video sessions that do not start with default speed, but in this dissertation, those with video speed other than 1.0 are classified into mixed-interacting.

Table 9.2: Descriptive statistics of interaction profiles. The cells with percentage values represent the proportions that the corresponding profile accounts for in the dataset for the course. Other cells with decimal values are the average perceived difficulty.

Course	Non-Interactive		Interactive				
	Silent	*Implicit-speeding*	*Explicit-speeding*	*Pausing*	*Skipping*	*Replaying*	*Mixed-Interacting*
RP	22.54%	6.22%	1.40%	17.46%	1.79%	3.45%	47.13%
	2.61	2.64	2.15	2.72	2.52	2.73	2.76
DSP	18.90%	3.05%	0.85%	19.12%	3.74%	3.12%	51.22%
	2.51	2.41	2.30	2.43	2.64	2.60	2.50

the different types of interactions seemingly reflect different perceived video difficulty. The *explicit-speeding* profile indicates the least perceived difficulty, whereas *pausing*, *replaying* as well as *mixed-interacting* inform that the students may have experienced higher difficulty.

In the remainder of this section, we extract video interaction features for *implicit-speeding*, *explicit-speeding*, *pausing*, *skipping* and *relaying* video sessions and build regression models to deeply investigate the relationship between each type of video control and perceived video difficulty. Video sessions of each profile contain only one type of video events, so that we can avoid the impact of complex interaction effects. The *mixed-interacting* sessions are thereby not analyzed. Having several observations per user in the dataset allows us to adopt a mixed model, where student is modeled as a random effect. Mixed models are known to be robust to missing values and unbalanced groups. In addition, least-square means (hereafter referred as LS means) mimic the main-effect means but are adjusted for group imbalance. These methods are used throughout the analysis. We will only report the analysis of the RP course due to its larger size, however the results for the DSP course are not dissimilar.

Implicit-Speeding Profile

Coursera video player offers 7 levels of speed ranging from 0.75 to 2.0 with a stepwise change of 0.25 and the video player inherits the video speed from the previous sessions. *Implicit-speeding* video sessions do not contain any video events, but the videos are started at a playrate other than 1.0. Intuitively, video speeds are presumably associated with the students' skilled or personal preferences. However, with very high or low speeds, we find the voices in videos are very much distorted. If a student decides not to switch to normal speed but to stay with the initial distorted one, it is reasonable to believe other factors such as video difficulty may have an influence.

We attempt to model the effect of initial speed by computing the LS means for the video sessions with different initial speeds and show the means with confidence intervals in Figure 9.3 . The two numbers separated by a slash ("/") under each category are respectively the number of survey responses and the total number of video sessions in the corresponding

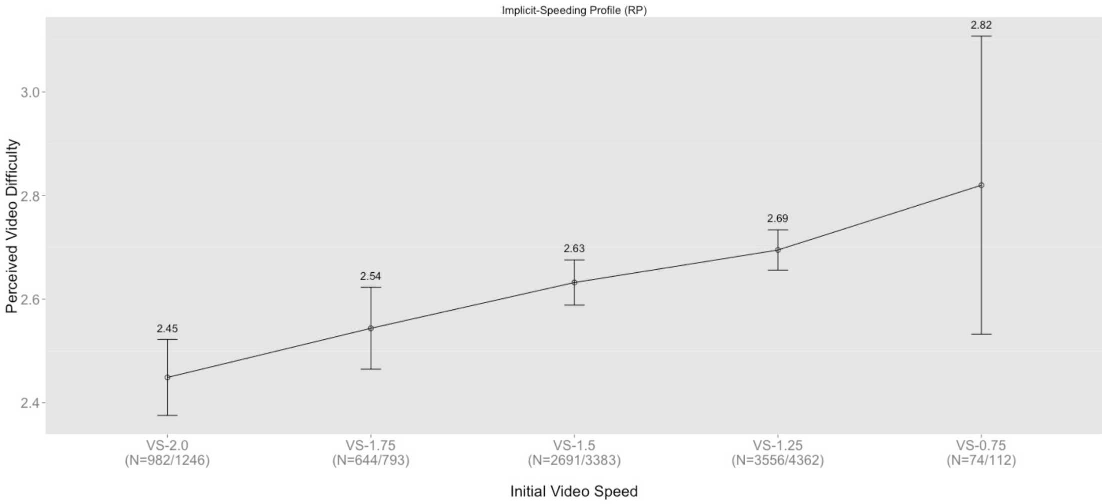


Figure 9.3: LS means of perceived difficulty in video sessions with different initial speed category.

Finding 1: Implicit-speeding shows a negative linear effect on the perceived video difficulty

Figure 2 shows a linear relationship. Considering the levels are numeric, statistically we assess the effects with a mixed linear model, which shows significant negative effects ($\beta = -0.08$, 95% CI = [-0.10, -0.05], p < .0001). That is, an increase of 0.25 video speed results in an average decrease of perceived difficulty by 0.08.

Explicit-Speeding Profile

In explicit-speeding video sessions, students use and only use speed changing controls to adjust video speed while watching. In principle, speed changing behaviors can be measured in the frequency and time dimensions, so we will examine the following features:

1. **Effective number of speed-ups and speed-downs**. This includes two measures that count the number of increase and decrease events respectively. However, simply counting individual events logged in the clickstream may raise two problems. First, a student may simply try out different playrates in a short period in order to find his or her preferred one. Counting intermediate states of video speed is likely to be inflated. Second, speed change is only achieved in multi-steps. For example, decreasing the video speed from 1.5 to 1.0 requires at least two stepwise changes of 0.25, which may also lead to inflation. Therefore, the events that happened within 10 seconds are grouped as a single event, the frequency of which is what we called effective number of speed-ups or speed-downs.

2. **Effective change of speed**. Before introducing this feature, we define the *average video*

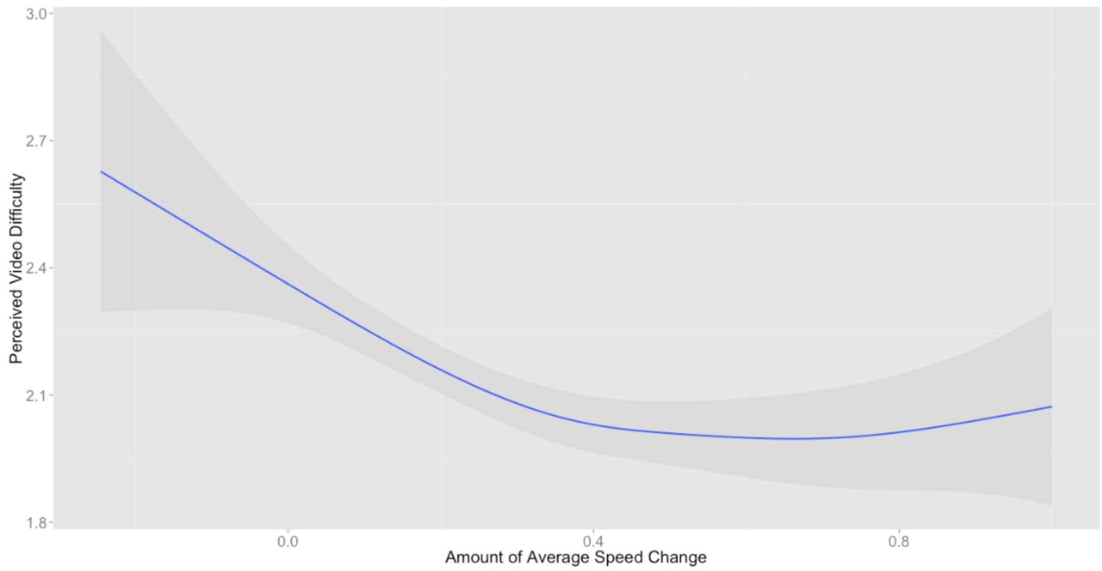


Figure 9.4: GAMM fit for amount of average speed change with confidence interval band

speed feature as the weighted arithmetic mean of video speeds at each video second. The *effective video speed change* is then computed by subtracting the initial video speed from the *average video speed.*

The three features are empirically distributed in a lognormal shape. The values of *effective change of speed* are of ratio type, ranging between -0.25 and 1.0. In 95% of the sessions, either the number of speed-ups or speed-downs is not more than 3. We did not expect their relationships with the perceived video difficulty to be linear, so we fit Generalized Additive Mixed Models (GAMM) for capturing the non-linear relationships. Compared to Generalized Linear Models (GLM), GAMM fits the data points with a spline smoother, which is able to capture non-linear relationship. Our reported statistics include the estimated degrees of freedom (edf) together with the p-value of an F-test that tests whether the smoothed function significantly reduced model deviance. This GAMM modeling technique are used primarily throughout the analysis in this section for features with widespread and highly-skewed distributions.

Finding 2: Speed-down frequency has a positive linear effect, while the amount of average speed increase has a monotonically negative effect till saturation point 0.4

We built a multiple regression GAMM model with the number of speed-ups, number of speed-downs and effective change of speed as explanatory variables and perceived video difficulty as outcome variable. The effective number of speed-down events shows significant effect ($\beta = 0.06$, 95% CI = [0.02, 0.09], $p < .005$), but the speed-up frequency did not ($p=0.73$). This is interesting, because only video sessions that were started with 1.0 speed were included in the analysis. Obviously there are more options for increasing the speed (i.e. 1.25,1.5,1.75, and 2.0) than decreasing (i.e. 0.75). In fact, more speed-down events were only possible if the video

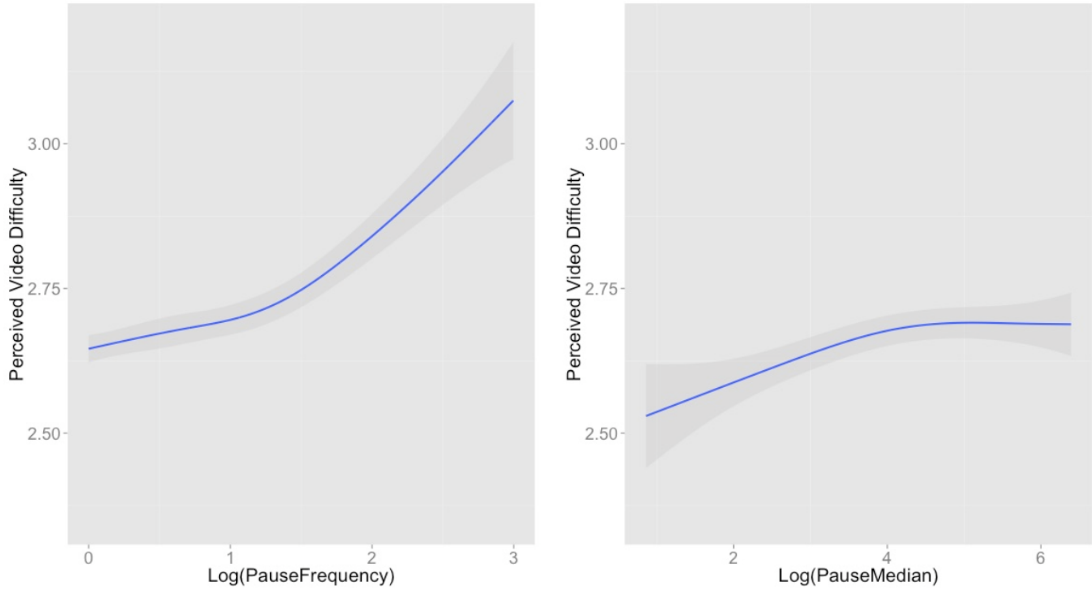


Figure 9.5: GAMM fit for pausing profile with confidence interval band

speed had been raised high. Keeping the effective speed change constant, this model actually suggests that slowing down a video from a previously increased speed significantly reflects higher video difficulty.

The effect of effective speed change on perceived video difficulty is non-linear (edf = 2.683, $p < .0001$), as depicted in Figure 9.4. As expected, the amount of speed change is negatively associated with the perceived video difficulty. This effect is only prominent when the changed amount is less than 0.4, after which the effect starts to saturate. For further increases the effect is weakened.

Pausing Profile

For the *pausing* profile, we hypothesize that the following two features potentially relates to the perceived video difficulty.

- **Median duration of pauses**. The durations of pauses distributed exponentially with long tail, so we then use the median of pause duration to gauge the time dimension of pauses. This statistic is more robust compared to "mean" or "sum" statistics, under the given data distribution.
- **Number of pauses**. As discussed previously, we only took into acount pauses that lasted between 2 seconds and 10 minutes. In fact, numerous pauses shorter than 2 seconds or across several days are observed in the dataset. The extremely short pauses do not make much sense in terms of cognitive processing. Those long ones, on the other hand, may actually indicate breaks rather than pauses. The choices of 2 seconds and 10 minutes

as thresholds are arguably arbitrary, i.e. it is difficult to articulate why 3 seconds or 11 minutes are not chosen, but we have tried slightly different values, and they did not make big differences in terms of results obtained from statistical model, which will be presented later.

We built a multiple GAMM to model perceived video difficulty with both features presented before. In fact, the data distributions of both features are highly skewed with long tail, so logarithm transformations (natural base) are applied on both features.

Finding 3: Pause Frequency matters more than duration

The pause frequency (edf = 3.14, $p < .0001$) and the pause median duration (edf = 2.439, $p < .0001$) both show significant non-linear effects on perceived video difficulty, and the corresponding GAMM fits are illustrated in Figure 9.5. We can see that the effect of pause frequency has visually steeper slope over the pause median duration. Lots of video sessions were found to contain great number of pauses (e.g. more than 10), where the students may constantly encounter problems in the videos. Note that the curve for median pause duration achieves its maximum and starts to stabilize at around 4.1 logarithm unit of the media pauses, which corresponds to roughly 60 seconds. This indicates when pauses are longer than 1 minute, the duration feature loses its predictability for perceived video difficulty.

Skipping Profile

For the *skipping* video sessions, we evaluate the following two features:

- **Number of forward seeks**. The total number of forward seeking events generated by scrubbing the playhead or clicking new positions in the video time bar.
- **Skipped video length**. The skipped video length refers to the amount of video seconds skipped by forward seeks. Closing a video before it ends also results in video content being skipped, but this is not considered in the analysis.

We built a multiple GAMM to model perceived video difficulty with these two features, whose distributions were also highly skewed with long tail, so logarithm transformations with natural base are applied.

Finding 4: Infrequent or large skip suggests higher perceived video difficulty

The number of forward seeks showed a negative linear effect ($\beta = -0.13$, 95% CI = [-0.19, -0.06], $p < .0005$) on perceived video difficulty. This is not surprising since it would be natural practice for the students to “jump” forward more often if they thought the videos were easy to comprehend. Frequently "jumping" forward in a video leads to skimming behaviors, which

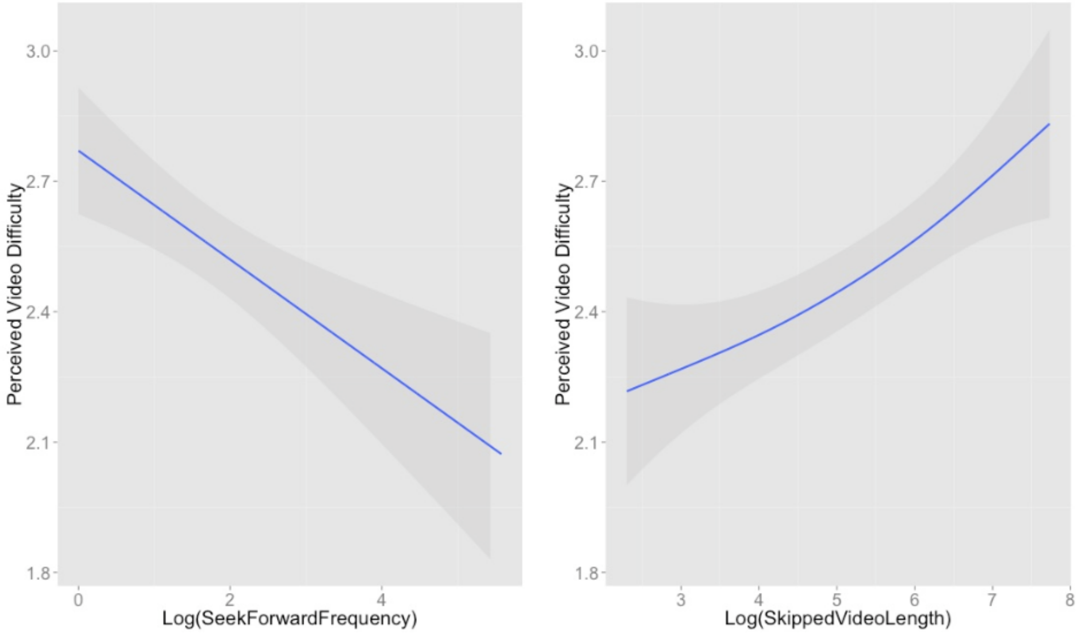


Figure 9.6: Model fit for skipping profile with confidence interval band

can be seen as an alternative way for speeding up the video. As presented in the *explicit-speeding* profile, more effective speed increases relate to lower video difficulty, which is in line with the result about the forward seeking frequency. Students who interacted in this way might have found skimming through the content sufficient for understanding the video.

On the other hand, when we hold the seeking frequency constant, we find the skipped video length exerted a positive non-linear effect (edf = 1.56, p < .0005). The estimated degree of freedom is quite close to 1, so the latter effect approximates a negative linear result, as depicted in Figure 9.6. This finding contradicts our expectation that more skipped content may indicate a video is boring and easy. In fact, this behavior may indicate higher video difficulty. Therefore, if frequently forward-seeking interactions can be understood as a way for quickly grasping the gist of the video, then large amount of skipped content perhaps implies "giving up" the video.

Replaying Profile

The *replaying* video sessions are analyzed in a similar way as we did for the skipping profile. The following two features are analyzed:

- **Number of backward seeks**. This is similar to the previously presented number of forward seeks, but in the opposite seeking direction.
- **Replayed video length**. The replayed video length refers to the video seconds that are re-watched by a student. The same parts of video can be watched several times. This measure accumulatively sums the total length of all replayed video seconds.

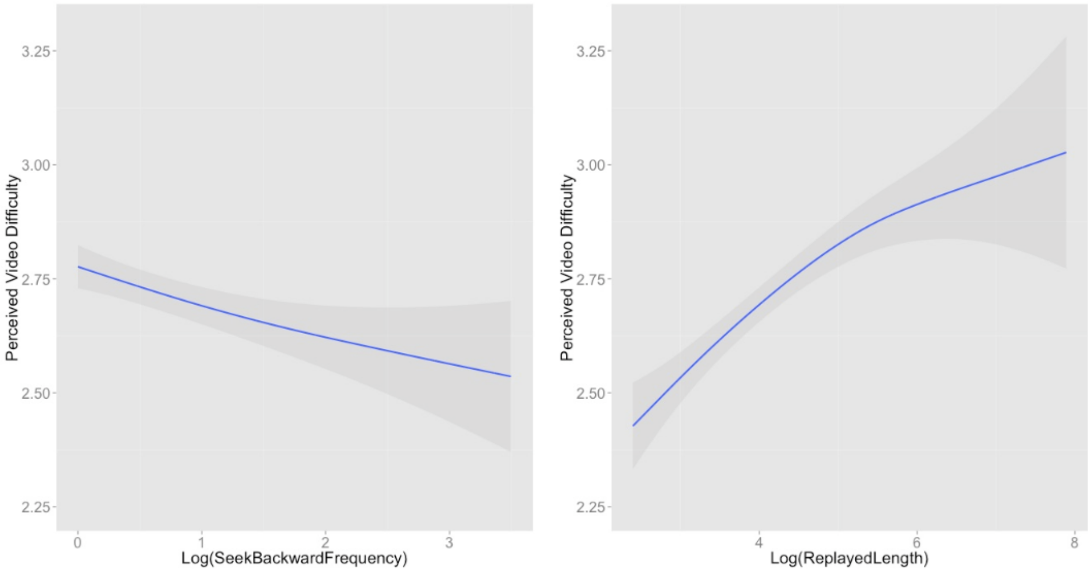


Figure 9.7: Model fit for replaying profile with confidence interval band

Similar to the analysis in the skipping profile, the above two features underwent logarithm transformation and were modeled as explanatory variables in a multiple regression GAMM for predicting perceived video difficulty.

Finding 5: Less frequent or large amount of re-watching indicates higher video difficulty

The replayed video length shows a positive effect on the perceived difficulty (edf = 2.20, $p < .0001$) as depicted in Figure 9.7 (Right). We can see that the curve has a monotonically sharp increasing trend until the value on the x-axis reaches around 6, which can be translated to 5-minute content being re-watched. After this point, the curve bends down a little bit. This finding indicates that the more a student replays the video, the more difficult they perceive the video. The effect is stronger if the replayed length is less than 5 minutes.

To our surprise, if the replayed video length is held constant, the number of backward seeks showed a significant effect on the perceived video difficulty (edf = 1.36, $p < .0005$). A similar finding was confirmed in the DSP dataset as well. The result suggests that on average higher replayed length per seek event is associated with higher video difficulty. In the video sessions containing high number of backward seeks, the events typically occurred within very short intervals, which may indicate that the students were deliberately looking for specific video frames. In this case, the frequent backward seeking behavior can be seen as more of "frame-seeking" rather than "re-watching".

Discussion

The previous analyses reveal the variation trends of the perceived video difficulty with respect to different types of video interactions. We extracted several features and find video interactions, including video speed decreasing, frequent or long pauses, infrequent or large-step skipping and re-watching are associated with higher video difficulty. These findings answered the first research question posed in Section 9.2.

Limitations

Although the results presented previously are statistically significant, the magnitudes of the effects are small, in terms of β value or variation slopes. In other words, we did not see the average perceived difficulty changes drastically within the variation range of any of the presented video features. Several reasons can possibly explain this phenomenon.

First, students study MOOCs with various motives, educational background, personal characteristics, habits and learning strategies. All these factors may also explain part of the variance in video difficulty. Second, MOOC students can externalize their perceived video difficulty in alternative ways. Instead of adapting video interaction accordingly, they may choose to tackle the problem in the forum or search in the Internet after watching the videos etc. Third, the analyses were conducted with a heterogeneous set videos from all weeks, the differences in video content were not considered. Our primary pursuit in the analysis was to generalize the effects of the *video interaction features* rather than *video content features.* The findings were actually similar in two different courses.

Impact of the Findings

As reviewed in Section 9.1.1, much existing MOOC research is devoted to predicting students' attrition or performance scores based on their video behaviors, and some (e.g. (Sinha et al., 2014a)) are promising. However, the relationships between video interactions and attrition or performance are definitely not causal. We believe other factors, such as learning motives and learning experiences, may be confounding. Since video lectures play a central role in MOOC learning, how students perceive the videos is a crucial measure of learning experiences. Despite the limitations presented before, the analyses presented before have identified a set of video features that are associated with students' perceived video difficulty. A natural next step is to combine these features to identify more general interaction patterns, which will be presented in the next section.

9.3.5 Inspecting Video Interaction Patterns

The video interaction profiles were strictly separated based on the type of video interactions. However, different types of interactions do not live in isolation. As Table 9.2 illustrates, *mixed-interacting* video sessions account for the largest proportion, but these sessions were not studied in the previous analyses. We argue a better segmentation of video interactions can be achieved by clustering the sessions with the video interaction features presented before. This section first explains how clustering methodologies are employed to identify video interaction patterns. This is followed by three follow-up analyses for investigating the relationships between the patterns and perceived video difficulty, video revisiting behaviors and performance, which provide deeper and more comprehensive insights for us to evaluate the zone of intervention for contextual information scent.

Video Interaction Clustering

The video features used for clustering are listed in Table 9.3. As discussed previously, these features characterize both the frequency and time dimensions for each interaction type. Most of the presented features that are significantly related to perceived video difficulty are included, except the number of speed-down events, because even in the *explicit-speeding* sessions, very few (no more than 5%) contain more than 3 speed-down events. Including this feature would not add much value to the clustering process.

Table 9.3: Video features used for clustering

1. number of pauses (NP)	5. number of backward seeks (NB)
2. median duration of pauses (MP)	6. replayed video length (RL)
3. number of forward seeks (NF)	7. average video speed (AS)
4. proportion of skipped video content (SR)	8. effective video speed change (SC)

The datasets contain a large number of video sessions with **in-video dropout**, i.e. the student left the video before they reached the end. Such behaviors are different from skipping with forward seeks, because in *in-video-dropout* situations the users never reached later part of the videos. When we compute the *proportion of skipped video content*, we only consider the proportion that is skipped by forward seeks, and unwatched content due to *in-video dropout* was not counted as skipped content.

When we cluster video interactions, we need to make sure the video sessions to be clustered are in the same vein. For example, a complete video session with one 1-minute pause should not be in the same cluster as an *in-video-dropout* session with the same interaction. Before proceeding to the clustering process, we separate complete video sessions from those containing *in-video-dropout*. While discarding the video sessions that do not reach the very end may be too strict, we group all the video sessions where the watchers did not reach the last 10% into the "in-video dropout" category, and our unsupervised clustering will be performed only on the remaining "complete" video sessions.

9.3.6 Clustering Pipeline

Our datasets contain a large number of video sessions without video interaction events (e.g. 17% for the RP). It makes little sense to include these data for clustering, because these video sessions form a natural cluster, which we call *Passive*. The *Passive* sessions are filtered out, so that clustering is performed on the remaining dataset of the two MOOCs independently with the 8 video features presented before. After preprocessing with PCA dimension reduction, we obtain 6 new uni-variance variables which account for 90% of the original variance.

For clustering, we use Neural Gas, a neural network-based convex clustering algorithm which is a robustly converging alternative to k-means. The goal of clustering is to obtain a minimal number of interpretable clusters explaining user behaviors. The Simple Structure Index (SSI) (Dolnicar et al., 1999) is used as a criterion for selecting the optimal number of clusters, since this index is known to multiplicatively combine several elements which influence the interpretability of a partition solution. For the RP course, we vary the number of clusters from 3 to 15, and find that 9 clusters maximize the SSI value (0.356 in [0,1] scale), as compared to the minimum value of 0.1 with 5 clusters. We then partition 9 video interaction clusters for the dataset. Similarly, 9 clusters are obtained for the DSP dataset as well.

9.3.7 Video Interaction Patterns

The centers of the 9 clusters for the RP dataset are shown in Table 9.4, and the results for the DSP course are analogous. The full names for the abbreviated feature names can be found in Table 9.3. We label each cluster with an intuitive name according to the corresponding dominating features (marked as bold) in the table. For example, the *LongPause(LP)* video sessions have an average median duration of pauses (MP) of 284.96 seconds. Note that the average number of pauses (NP) for this cluster is small (1.71). So it actually represents video sessions with infrequent long pauses.

Table 9.4: Cluster centers for the RP dataset

Pattern	Proportion	NP	MP	NF	NB	SR	RL	AS	SC
Replay (RP)	3%	4.73	62.58	5.86	12.84	0.05	**531.44**	1.10	-0.00
HighSpeed (HS)	10%	1.17	23.19	1.18	0.95	0.10	27.14	**1.66**	-0.01
SpeedUp (SU)	3%	1.38	27.16	1.66	1.04	0.09	25.13	1.53	**0.39**
SkimSkip (SS)	4%	1.00	30.73	**21.70**	4.94	**0.75**	17.46	1.14	0.00
Inactive (IA)	38%	1.93	39.05	0.71	1.28	0.03	36.65	1.05	-0.00
FrequentPause (FP)	4%	**13.39**	40.58	2.87	5.13	0.05	109.37	1.08	-0.00
JumpSkip (JS)	13%	0.45	11.62	**5.38**	1.10	**0.71**	9.40	1.06	0.00
LongPause (LP)	6$	1.71	**284.96**	1.34	1.26	0.08	44.62	1.07	0.00
SpeedDown (SD)	1%	2.13	42.93	1.61	1.73	0.08	44.42	1.24	**-0.58**

While Table 9.4 only presents the centroids of the clusters, the distributions of these features are illustrated as boxplot in Figure 9.8, in terms of standard scores (z-scores) of the variables. 50 percent of the data points are enclosed in the boxes. The upper and lower whiskers extend from the hinge to the highest or lowest value that is within 1.5 interquartile range of the hinge. Data beyond the end of the whiskers are considered as outliers and are not shown in the figure. The multidimensionality, continuity, skewed distribution, and inter-correlation natures of the features imply that clear separations are unlikely to be found based on the current feature sets, and this explains why the maximum SSI (0.356) of the partition solutions is relatively small. Nevertheless, the dominating features in each cluster are still prominent, as shown in Figure 9.8.

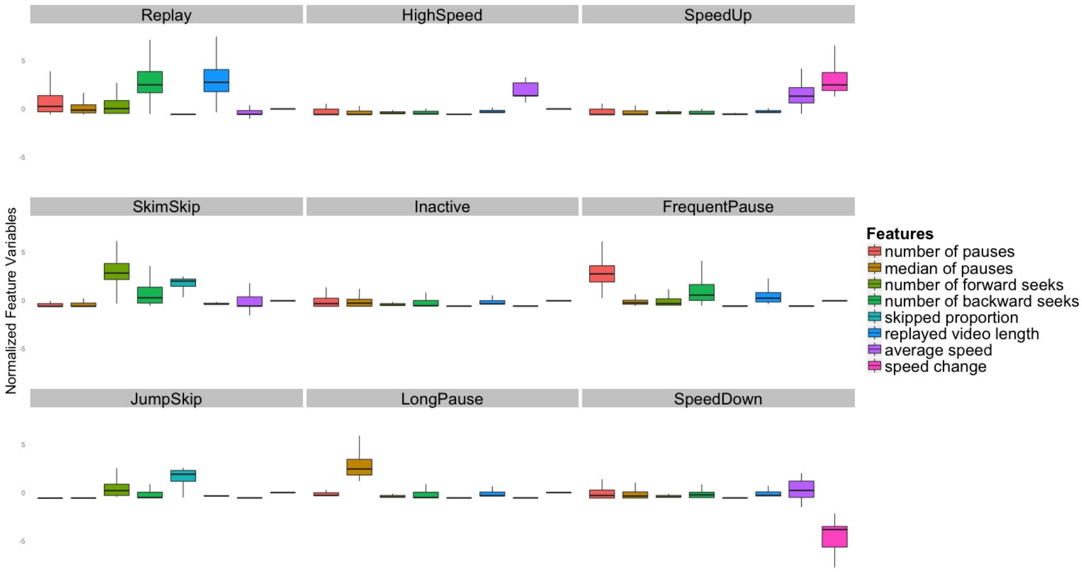


Figure 9.8: Cluster data distributions for the RP dataset

In addition to the presented 9 patterns, we have (17 %) *Passive (PS)* sessions. It should be noted that most video sessions contain few video events. The *PS*, *IA* and *HS* account for 65% of the dataset, indicating a small number of video interactions satisfy the students' need for most of the time. On the other hand it also implies the adoption of rarer patterns may reflect certain changes in the students' learning state. We will discuss them in the upcoming sections.

9.3.8 Perceived Video Difficulty

In Section 9.3.4, we studied the relationships between the video features of each interaction profile and the perceived video difficulty. Here we examine the video interaction patterns for the same question: i.e. ***How do the different video interaction patterns reflect different levels of perceived video difficulty?***. Since the interaction patterns are more naturally partitioned than the interaction profiles, the analysis would likely to give more valuable insights for recognizing the difficult situations for a student (zone of intervention).

Result

We built mixed-effect ANOVA models, where the students were modeled as random effects, to compare the perceived video difficulty among the video interaction patterns, which were found to be significantly different (RP: $F(9,124964) = 313$, $p<0.0001$; DSP: $F(9, 17505) = 24$, $p<0.0001$). We plot the Least-square mean difficulty with confidence interval in Figure 9.9. The colored labels underneath the name of each pattern along the x-axis depict the number of video sessions with difficulty ratings and the total number of video sessions belonging to the corresponding pattern. The two numbers are separated with a slash sign "/".

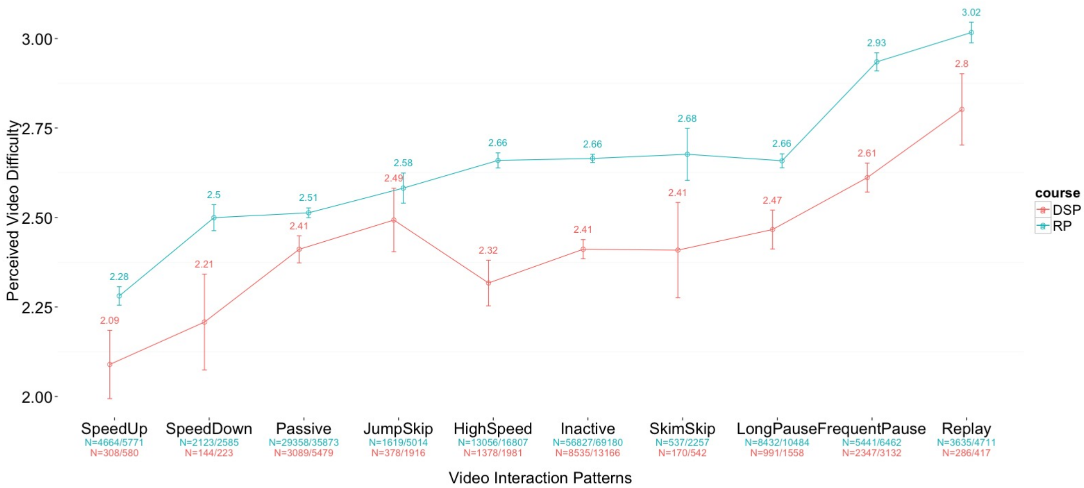


Figure 9.9: Video interaction patterns and perceived video difficulty

Figure 9.9 shows that relative differences in perceived video difficulty of the 10 video interaction patterns are more or less consistent across two courses (with a systematic difference attributed to the course intrinsics), though the clusters are generated independently. Therefore, we tend to believe empirical patterns emerged from another course are likely to follow a similar trend.

The *Replay(RP)* and *FrequentPause(FP)* patterns reflect significantly higher video difficulty than others. In other words, these two patterns were commonly employed as strategies to cope with difficult videos. Students may adopt the former pattern to clarify doubts within the videos by rehearing the explanation, whereas the latter may be used when the explanations cannot be found within the video (requiring external resources) or the verbal/visual explanations are too fast to be processed (requiring buffer time). The scenarios for pausing the videos are what we refer to as *internal interruption* in Section 2.1.3.

The *SpeedUp(SU)* pattern reflects significantly lower video difficulty compared to the other patterns, indicating explicitly increasing the speed during video playback are commonly used by MOOC students to watch to easy videos. This way, students can quickly grasp the gist of the video without skipping content.

As depicted in Figure 9.9, the average video difficulty of all the other patterns except the three discussed before fluctuates with small variations. The finding can be interpreted as follows: Most interaction patterns reflect a similar level of video difficulty, when students switch to either the *Replay(RP)* or the *FrequentPause(FP)* pattern, they may encounter problems. On the other hand, whey students switch to the *SpeedUp* pattern, they are likely to be viewing an easy video.

9.3.9 Video Revisiting Behaviors

Compared to traditional classroom lectures, MOOC videos are permanently preserved online, which makes revisiting certain videos a common practice. Students may revisit an older video for checking concepts while watching new videos or doing homework. They may also revisit a video if the first-time watch was not sufficient to comprehend the content. Kim et al. (2014) find out that first-watching sessions are more sequential while the revisiting sessions are more selective, i.e. the students selectively navigate the video into specific parts. All the analyses presented so far are based on first-time visiting video sessions. In Section 9.3.3, we presented that revisited videos were generally rated more difficult than first-time visited videos. Therefore, studying revisited video sessions may also shed lights on evaluating the zone of intervention. In this part of the analysis, we inspect video revisiting behaviors by asking: ***With which first-time video interaction patterns are the videos more likely to be revisited?***.

Table 9.5: Proportion of video revisiting for complete and in-video dropout sessions

	RP		**DSP**	
	Completed	**Dropped-out**	**Completed**	**Dropped-out**
Revisiting	20.6%	73.7%	23.5%	59.3%
No Revisiting	79.4%	26.3%	76.5%	40.7%
	χ^2(1,220875)=55805.1, p <.0001		χ^2(1,38825)=5114.1, p <.0001	

Before answering the above question, we first take an overview of video revisiting behaviors on all the video sessions, including those "in-video dropout" video sessions, which have been excluded from the previous analyses. We start by investigating how different *complete* and *in-video-dropout* video sessions are associated with revisiting sessions. As Table 9.5 illustrates, around one fifth of the completed videos (the "Completed" column) were revisited later. In comparison, videos that contained in-video dropout (the "Dropped-out" column) in the first-time watching sessions are significantly more likely to be revisited (73.7 % for the RP and 59.3% for the DSP), according to the reported Chi-squared statistics. Note that all the Chi-squared tests hereafter are actually conducted with frequency of occurrences, but are presented with percentage.

If we further focus on the video interaction patterns for complete video sessions only, then the results are as shown in Table 9.6. In each cell the percentage represents the observed probability of revisiting after the first view with the corresponding patterns. The expected probability for a video to be revisited for RP and DSP courses are 20.1% and 23.5% respectively, under the null hypothesis that video revisiting is independent of the interaction patterns. Chi-Squared tests show that the chances of revisiting significantly depend on the first-time video interaction patterns. Post-hoc residual analysis further reveals which cells contribute most to the Chi-Squared value. This is expressed by the adjusted standardized residuals, as shown below the percentage values in each cell. Significant positive residuals at $\alpha = 0.05$ (adjusted standardized residuals that are more than 2) are highlighted in bold. These highlighted cells indicate the frequency of occurrences for the corresponding patterns are significantly overly observed with respect to the expected frequency.

Table 9.6: Proportion of video revisiting for complete sessions

		RP	**HS**	**SU**	**SS**	**IA**	**FP**	**JS**	**LP**	**SD**	**PS**
RP	Revisiting	35.7%	17.2%	15.0%	25.6%	21.6%	26.1%	21.6%	21.1%	21.3%	20.6%
		25.9	-11.3	-10.5	**5.96**	**8.93**	**11.1**	1.86	1.31	0.99	-16.8
	No Revisiting	64.3%	82.8%	85.0%	74.4%	78.4%	73.9%	78.4%	78.9%	78.7%	79.4%
		-25.9	**11.3**	**10.5**	-5.96	-8.93	-11.1	-1.86	-1.31	-0.99	**16.8**
		$\chi^2(9,156517)=1293.7$, $p<.0001$									
DSP	Revisiting	41.5%	21.8%	16.5%	22.5%	22.1%	32.0%	23.9%	23.6%	26.6%	21.6%
		7.8	-1.6	-3.57	-0.47	-4.49	**10.7**	0.32	0.1	0.95	-3.24
	No Revisiting	58.5%	78.2%	83.5%	77.5%	77.9%	68.0%	76.1%	76.4%	73.4%	78.4
		-7.8	1.6	**3.57**	0.47	**4.49**	-10.7	-0.32	-0.1	-0.95	**3.24**
		$\chi^2(9,22717)=197.7$, $p<.0001$									

For both courses, the videos with *JumpSkip(JS)*, *LongPause(LP)* and *SpeedDown(SD)* do not show significance in revisiting behaviors. Interestingly, we find that the videos viewed with *Replay (RP)* and *FrequentPause (FP)* are significantly more likely to be revisited, while less revisiting probabilities are found with the *SpeedUp (SU)* and *Passive(PS)*. In section 5, *RP*, *FP* and *SU* are shown to reflect respectively the highest and lowest subjective difficulties. Therefore, we infer that the video difficulty may confound between the interaction patterns and the probability of video revisiting. However, as other patterns are weaker indicators of the perceived difficulty, the revisiting behaviors may in this case be confounded largely by other factors such as the course intrinsics. For example, the videos with *Inactive (IA)* pattern are significantly more likely to be revisited in the RP and less in the DSP. The potential reasons are hard to identify in this case. In this section we highlight the general finding that videos with *RP* and *FP* patterns are more likely to be revisited, and more follow-ups of this finding will be discussed later.

9.3.10 Student Performance

Students in MOOCs often have diverse background and learning abilities. Depending on their levels, MOOC students may watch video lectures in different ways. For example, we may hypothesize that strong students selectively watch MOOC videos whereas weak students spend more time with the learning materials. Our question is ***How do Strong and Weak students differ in lecture video viewing behaviors?*** The video interaction patterns provide us with a handy tool for diagnosing the students' video behaviors, so our analysis will be based on comparing the strategy of employing video interaction patterns.

Method

The foremost challenge for the analysis to pursue an answer to the posed question is to define *Strong* and *Weak* students. Considering MOOC is an open platform, students have different motives. A great proportion of the students drop out in the early or middle of the courses for various reasons. Even those who watch all the videos do not necessarily aim at obtaining a certificate or completing all the learning activities. This means the students who obtain 0 points in the final score are not necessarily weak in their learning abilities. As mentioned in Section 9.3.1, no exams were placed for the two courses in our datasets, and weekly quizzes are the only mean for student assessment. The quizzes for the RP course can be submitted unlimited times, and we do have seen that many students submitted more than 10 times for a quiz. The consequence is that 82% of the passed students got certificates of distinction, which is a quite inflated percentage. In order to compare the students who are strong and weak in learning abilities, we take a subset of the data which includes only the students who completed all the assignments. Thus we believe the remaining students have a similar learning goal, which is to complete the courses. As shown in Table 9.1, only 263 (less than 3%) of the total students passed the DSP course, and only 23 students obtained certificates with distinction, whilst the RP course achieved pretty high completion rate (23.15%). In the analysis hereafter, we only analyse the RP students who submitted all of the 6 assignments. To simplify the analysis, the students who obtained 80% of the total points in their *FIRST* quiz submissions are defined as *Strong*. Otherwise, they are labeled as *Weak*. The subset contains 4555 (86.3%) of the passed students, of which 35.3% are *Strong* students. Same as in many of the previous analyses, only the first-time watching patterns are concerned.

Result

In the targeted RP course, a video session has an expected probability of 37.6% to come from *Strong* students under the null hypothesis that the employment of video interaction patterns is independent of students' performance (Table 9.7). Chi-square test shows that the adoptions of video interaction patterns are significantly different between strong and weak students. Post-hoc residual analysis reveals that strong students tend to interact less with the videos, so the frequencies of *HighSpeed(HS)*, *SpeedUp(SU)*, *Passive(PS)* and *Inactive(IA)* sessions are significantly higher. On the other hand, weak students interact more with videos, they use significantly more *SkimSkip(SS)*, *JumpSkip (JS)*, *FrequentPause(FP)* and *LongPause(LP)*.

Table 9.7: Proportion of video interaction patterns based on students performance

	RP	HS	SU	SS	IA	FP	JS	LP	SD	PS
Strong	38.7%	46.1%	39.7%	31.6%	35.9%	32.4%	33.8%	35.5%	38.6%	39.0%
	-1.2	**17.3**	**2.1**	-4	-9	-6.6	-3.6	-3.2	0.7	**3.8**
Weak	61.3%	53.9%	60.3%	68.4%	64.1%	67.6%	66.2%	64.5%	61.4%	61.0%
	1.2	-17.3	-2.1	**4**	**9**	**6.6**	**3.6**	**3.2**	-0.7	-3.8
	χ^2(9,76094)=406.3, p <.0001									

Recall that both *FrequentPause(FP)* and *Replay(RP)* reflect the highest video difficulty. Interestingly, the usage of *Replay* pattern is not significantly different between the two student groups, indicating the frequency of replaying behaviors do not discriminate strong/weak students.

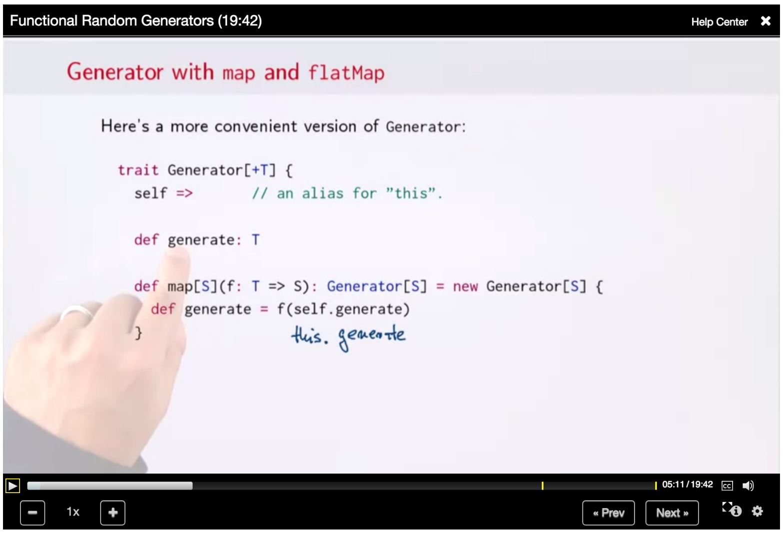


Figure 9.10: Example video frame with code snippet

We are especially interested in gaining deeper insights about the two pausing patterns, i.e. *LongPause* and *FrequentPause*. As discussed previously, pauses may occur when the presented information is overloaded so that the students require additional time or external material to comprehend the content. Weak students are found to use both *FrequentPause(FP)* and *LongPause(LP)* significantly more often than strong students. In order to understand how the weak students adopt the two pausing patterns, we randomly selected 50 video sessions with *FrequentyPause(FP)* pattern and another 50 with *LongPause(LP)* pattern from the weak students' interaction logs and manually examine the situations under which the associated 698 pauses happened. We categorize the pauses by the occasions when the professor was explaining *example codes* (46.7%), *programing grammar* (12.7%), *demos* (4.8%), *theories* (33.8%)

and *others* (2%)(e.g. in summary). Fisher's exact test shows no significant differences in the categories of pauses in sessions between *FrequentPause(FP)* and *LongPause(LP)* patterns (p=0.47). Nearly half of the pauses occur when example code snippets are shown in the video frame (e.g. in Figure 9.10), and more than half of the pauses are related to the presentation of code (code, grammar and demo). This result indicates that the weak students may have significant problems in understanding the code compared to the strong students.

9.3.11 Discussions

Our analyses show that MOOC students follow different video interaction patterns while watching lecture videos. The strategy of adopting the patterns may vary for different videos, depending on the students' perceived video difficulty, their capability and whether or not a video is watched for the first time. This section discusses how the results presented before help us evaluate the zone of intervention and provides design insights for contextual information scents.

Evaluating Zone of Intervention

For evaluating zone of intervention, we are actually evaluating when students are experiencing difficulty in videos, and how severe the problems are. In addition, we especially identified potential common problems for low-performing students. Proper interventions can then be introduced to help them. Through the analyses in this section, the following schemes can be employed to recognize the potential problematic situations:

1. Detect the change of video interaction patterns

As presented in Section 9.3.8, MOOC students seldom interact much with lecture videos, but once they do so, we have shown that different video interaction patterns reflect different levels of perceived video difficulty. The videos viewed with *SpeedUp* pattern are perceived to be easy, whereas videos viewed with *FrequentPause* or *Replay* patterns are perceived to be more difficult. Changing video interaction pattern to one of those that reflect higher video difficulty potentially creates a zone of intervention for contextual information scent.

2. Follow video re-watching patterns

Video sessions with in-video dropout have 60% -70% chances to lead to revisits. However, we are not sure whether or not the reasons behind such video revisits can attribute to video difficulty. Students may have left a video early simply due to time constraints. Nevertheless, considering the very high return rate, it is advisable to provide the students with short-cut access to the videos that are left earlier.

In addition, we also find out that complete video sessions with interaction patterns that reflect higher difficulty, such as the *Replay(RP)* and *FrequentPause(FP)* are more likely to be revisited.

This could be a signal of high difficulty, so that contextual help can be provided.

3. Make use of the pauses

Weak students tend to make more self-interruptions with *FrequentPause(FP)* or *LongPause(LP)* patterns than strong students. We have shown that in the programming course, these students especially paused more in video frames with code snippets, indicating they may have problems or simply need more time to understand the code. The paused periods in these cases, create zone of intervention for proper support.

Designing Intervention with Contextual Information Scents

Previously we discussed that the zone of intervention can be identified by means of an analysis of students' video interactions. A natural next step is to design the interventions. MOOC video players, in the current forms, play lecture videos in a traditional way. We argue contextual information scents can be designed to provide the students with timely support. In this case, the context comes from the video content being interacted. Since the videos were all pre-recorded, helping resources may be prepared in advance. Information scents can potentially be created from textbook pages, Web articles or discussion forums through either an instructor-supervised or a crowd-sourcing process.

Regarding where to display the information scents, the screen real-estate around a video player is a possible candidate. A second option is to design a video overlay with information, this especially has potential for "making use of pauses". An example for a programming MOOC as the RP could be displaying explanations of a piece of code as overlay when video is paused, so as to assist students' comprehension.

9.4 Conclusion

This chapter shed some light on the relationships between student video interaction and some important aspects in MOOC learning such as the perceived video difficulty, video revisiting behaviors and students' performance. We start by delivering an understanding of different types of video interactions, and then we extend the discussion to a more comprehensive analysis of video interaction patterns. The analyses aimed to help us evaluate zone of intervention, so that contextual information scents can be designed to help students solve problems.

However, this chapter provides only data inference rather than concrete implementations, as in previous chapters. Our pursuits are generalizable findings through statistical inferences. Therefore, the analysis did not consider the difference in video content. In addition, we only analyzed the first time video viewing sessions, which is another limitation of the analysis.

MOOC learning is a multi-faceted learning practice, watching lecture videos is central but it does not portray the complete picture. Activities in the forum, quiz, students' motivations all may also create zone of intervention. Future work may also incorporate these factors to gain a more comprehensive understanding about how students learn in MOOCs.

10 Reflection & Roundup

In this dissertation, we explored the design and implementation of contextual information scents, as well as the deployment and evaluation of them in learning activities. After a comprehensive literature review in theories and practices about human cognition and learning, information seeking and interaction, we conducted an initial survey study in order to understand information needs in learning and we further proposed a research framework consisting of design principles, design space and interaction phases for contextual information scents. This framework guided the follow-up explorations and analyses.

Throughout the thesis we presented four systems designed for contextual information scents: the RaindropSearch and TileSearch generate information scents based on conversation context; the MeetHub Search and the BOOC Player are based respectively on the context of groupware interactions and lecture video content being viewed.

Finally we extended our research scope to MOOC learning, and used data analytics to evaluate the zone of intervention for contextual information scents. In this concluding chapter, we reflect the lessons learned from previous studies, and point out our contributions, limitations and future work.

10.1 Reflection

We have discussed various learning activities so far. The initial survey study for understanding information needs was conducted in a seminar scenario; the prototypes for contextual information scents were designed for various collaborative learning activities; finally, data-driven approach for evaluating zone of intervention was applied in online learning. In this section, we look at a bigger picture, by reflecting the findings in the previous studies. The discussion will focus on three aspects: (1) information needs and zone of intervention, (2) adherence to design principles, and (3) the benefit of contextual information scents.

10.1.1 Information Needs and Zone of Intervention

The foremost assumption of our research is that people may have information needs while performing learning activities and such information needs potentially create *zone of intervention* for computer support.

Information Needs

The survey study presented in Section 5.1 offered preliminary insights about information needs in learning activities. Following the discussions of the survey study, we proposed to discriminate the elasticity of information needs. Inelastic information needs are well-articulated or must-satisfy needs. People often have strong will and clear goal to address the needs, sometimes with elaborated efforts. In contrast, elastic needs might not be fully aware of or not easy to articulate. In addition, it is sometimes improper to search for them due to social or environmental barriers during the performance of an activity, e.g in a seminar talk.

Information needs have different motives and modes (cf. Table 5.2). Those corresponding to lower layers of the Bloom's taxonomy (e.g. casual, lookup and learn) are likely to be elastic, whereas the modes close to the other end (e.g. investigate) are usually more associated with inelastic needs that may lead to a thorough and elaborated information seeking process. The elasticity of information needs is subjective and may be reflected by people's strategies in dealing with them. Inelastic information needs may be addressed by people by all means, elastic needs otherwise. The satisfaction of elastic informations may depend on the availability of convenient tools, and the contextual information scents are intended for this purpose.

Zone of Intervention

Information needs, regardless of elasticity, potentially create the so-called *zone of intervention*, which was originally defined as "area in which an information user can do with advice and assistance what he or she can hardly do alone or can do only with difficulty" (Kuhlthau, 2004). The term "intervention" refers to technological support to address the information needs. In principle, dedicated search tools can support both inelastic and elastic information needs. However, as discussed in Section 5.1, people who brought search devices to the seminars rarely use them for looking for information or failed to find the right information. Similarly, in the collaborative MOOC viewing study presented in Chapter 8, students seldom used the print textbook for help. However, with the introduction of the BOOC Player, significantly more use of the book were observed. This was an example that demonstrates the availability of proper interventions may influence peoples' behaviors with learning support materials, exhibiting great potentials for designing contextual information scents (e.g. the BOOC Player) as intervention.

Contextual information scents were intended for the zone of intervention created by elastic information needs. Inelastic information needs are expected to be supported by dedicated search tools, because the latter usually requires more elaborated efforts. However, this only speaks for the intended roles of the intervention technologies. In both RaindropSearch and MeetHub Search studies, participants are found to have employed the keyboard to search for unfamiliar terms and concepts. However, these needs could have been supported by the contextual information scents. Especially in the MeetHub Search study, some groups were obsessed with searching with keyboard. Indeed, this phenomenon may attribute to design flaws in the contextual information scents, but it also demonstrated the *principle of least effort* (cf. Section 3.1.2), which claims that information seekers prefer to acquire information with minimal efforts. In other words, searching information with keyboard was probably considered by the participants as the "minimal" solution, compared to interacting with the contextual information scents. In a similar vein, the MOOC study groups (cf. Chapter 8) resolved video difficulty mostly through group discussions, which was viewed by the groups as a more convenient means for tackling difficulty. In our experiment, the *zone of intervention* largely intersects with the *zone of proximal development*, but solving problems with a *more knowledgeable other* requires less effort.

In Chapter 9 we associated MOOC students' video interactions with perceived video difficulty and performance scores. This allowed us to evaluate the *zone of intervention* to support individual students by contextual information scents.

10.1.2 Adherence to Design Principles

In Section 5.2.1, we proposed a set of design principles to guide the exploration of contextual information scents in this thesis. This section summarizes how the design explorations adhere or fail to adhere to each of the principles.

Calm

According to the capacity theories of attention (cf. Section 2.1.2), if people focus their visual attentions on the primary task, then they would experience problems for simultaneously attending to peripheral objects with the same modality. That is to say, when a student is following a talk, or face-to-face discussing with others, or watching lecture videos, s/he cannot perform a secondary task such as searching information at the same time. However, the selective attention theories (cf. Section 2.1.1) posit that the student is still able to perceive and recognize peripheral objects visually, if these objects are relevant. The chances for successful recognitions are especially high if the peripheral objects are semantically related, which leads to priming effect (cf. Section 2.1.4). These theories lay a theoretical foundation for the display of contextual information scents in the learner's peripheral vision. A major concern is, however, to ensure minimal distractions and disturbances for the learning activity, which is described as the "calm" design principle.

Our very first prototype, the RaindropSearch achieved calmness by sacrificing "usability".

Noun words spoken in the conversation were enclosed in "rain drops" that fell down individually with a constant speed. A single "rain drop" definitely created "pop-up" effect (cf. Section 2.1.1), so that the moving object became easily noticeable. As group discussions became heated, the display was full of rain drops, thus eliminating the "pop-up" effect. In addition, searching with the rain drops required additional cognitive load for locating and eliciting words as query terms, resulting in distractions. In contrast, searching with keyboard required less effort, because the information seeking process has been habituated. As a follow-up design, the TileSearch displayed image or Wikipedia search results as contextual information scents, but animations were not employed. This prototype achieved calmness in a more successful way. In a similar vein, the Marquees in the MeetHub Search also employed Wikipedia and image search results as contextual information scents. As the relevance of results increased with the weighted-selection approach in week 3, the reported distractions were significantly reduced. This implies that animations may not be the determining factor for perceived distraction, whereas relevance is perhaps more important in this regard. The BOOC Player created highly relevant contextual information scents for MOOC video viewing with supervised book mappings. As a result, distractions were not an issue any more.

Context-awareness

As the name suggests, context-awareness is the key attribute of contextual information scents. In this dissertation we explored three types of context: group conversation, groupware interaction and video content. The goal was not to decide which context was best suited for generating information scents, but to approximate the use of context by gaining insights about how to design information scents out of different contexts.

Conversation, as explored in the RaindropSearch and TileSearch, has the advantage in quickly responding to situational collaborative learning context. The two prototypes both captured just-spoken words immediately and generate contextual information scents out of them. The disadvantage lies in the difficulty in the elicitation of keywords directly from conversation, because conversation usually contains noise (irrelevant or redundant information) and ambiguity. It is reasonable to believe that interacting with text in a groupware offers a less noisy and less ambiguous context, which gave birth to the MeetHub Search. The text created by the users in the MeetHub groupware were presented in a more streamlined way, but with the loss of timeliness. That is to say, group discussions usually precede groupware interactions, and the latter are usually the result of the discussions. Therefore, the information scents generated from the text were mostly not useful at the time when they were shown to the group. The BOOC Player displays the most relevant textbook pages at the right moment during video watching, but it lacked user-generated context, since the generation of information scents was independent of group discussions. Ideally different kinds of context can be combined to achieve higher context-awareness. For example, in Chapter 9 we evaluated the *zone of intervention* based on MOOC learners' video interactions, which identifies **when** students may need help. In addition, we can also identify **what** information can be potentially designed as information scents as intervention based on the content of the difficult video.

Redundant

Redundancy was proposed as a principle, because we view contextual information scents as an analogy to intermediaries in mediated search. According to the *principle of guaranteed result*, we believe contextual information scents cannot underline precision. Instead, they should present a max coverage of redundant but relevant information, from which users can select the most useful piece.

By following this principle, the RaindropSearch presented all the noun words captured from the conversation. The TileSearch and the Marquee UI in the MeetHub Search displayed multiple Wikipedia and image search results. For the latter system, we employed the *combinational approach* (i.e. the CA, cf. Section 7.1.2) to elicit different combinations of keywords for building search queries. However, it turned out that our group participants were overwhelmed by the "redundancy" created by the CA. In contrast, the weighted-selection approach (i.e. the WA) achieved more acceptable redundancy. The BOOC Player did not seem to contain redundancy at the first sight, since the book pages were seemingly precisely mapped. However, precise page mapping does not equal to the exposure of the exact piece of desired information. A textbook page may contain images, formulas, tables and text paragraphs, all of which constitute the "redundancy" of the contextual information scents in the BOOC Player, which cue the learners for potentially useful information.

Trigger-rich

Trigger-richness is one of the key considerations for inducing serendipity through facilitating the process of making connections (cf. Section 4.2.3). Considering the redundancy, we did not expect all of the displayed information scents to be helpful during the performance of learning activities, but serendipitous encountering was a desired attribute for the design. However, as the contextual information scents are not designed for focused attention as in other serendipity-inducing systems (e.g. the Bohemian Bookshelf (Thudt et al., 2012)), additional challenges were raised for the design to be peripherally trigger-rich.

Trigger-richness can be achieved visually and semantically. In the previous systems, trigger-richness was mainly associated with the capacity of the contextual information scents (cf. Section 5.2.2), which in turn relates to the type of serendipitous encounters. The RaindropSearch and the WordCloud in the MeetHub Search aimed to induce serendipity to trigger the recognition of query terms for searching information. The former attempted to achieve it by constantly visualizing a rich set of just-spoken words. However, the design rarely triggered recognitions, and participants were found to use the spoken words simply as an alternative input modality. The latter was designed for enhancing search experiences by presenting searchable keywords in a more advanced visualization, but the tool was still seldom used. Formulating automatic query terms and conducting searches are still at an early phase in an information seeking process, and people may be habituated to do it in a traditional way. Therefore, serendipity may be more effectively induced in later information seeking phases, such as during the elicitation of results, requiring the information scents to carry higher information capacity. The TileSearch and the Marquees designed higher capacity contextual information scents as images and Wikipedia snapshots, and the BOOC Player achieved it with even higher capacity, i.e. textbook pages.

When visual richness, in terms of information capacity was assured, semantic richness becomes the key. For example, the typing-triggered approach employed by the Marquee considered only instantly typed words, which lacked a global semantic context. More semantically rich information scents, such as the Maruqees generated with the weighted-selection approach, were more effective.

Multi-phase Interactive

Trigger-richness only facilitates making connections from the information scents, but means should also be provided for follow-up phases of serendipitous encountering, i.e. exploiting and reflecting the values of the connections. That is what we mean by Multi-phase Interactive.

All of the contextual information scents designed in this dissertation obeyed this principle, since they were all interactive, leading to the examinations of more detailed information. The information scents in the form of words in the RaindropSearch and WordCloud can be navigated to a list of Web search results. TileSearch, Marquee and the Querylist guided the users to the selected Web article. The BOOC Player displayed fine-grained information in a textbook as contextual information scents, but it also offered page navigation and interaction possibilities to further explore the information.

10.1.3 Added Value of Contextual Information Scents

The design of contextual information scents is motivated by the observation of situational information needs in learning activities. Some of the needs are elastic, and are usually not addressed timely. We argue contextual information scents would bring added value to the learning activities by addressing these elastic information needs. In Chapter 6 we recognized three potential roles of the contextual information scents in collaborative learning activities: (1) serendipity inducer (2) group facilitator (3) learning support. The RaindropSearch, Marquees generated with the typing-triggered and combinational approaches were almost not used. In this section we summarize how the other systems brought added values as each of the three roles.

Serendipity Inducer

At the lowest level, contextual information scents induce serendipity. In our studies, serendipity is measured objectively, by counting the frequency of interactions. We see it as the lowest level added value, because serendipitous encountering certain information and interacting with it may simply facilitate certain information needs occasionally, but does not necessarily lead to a systematic effect in the facilitation of group discussions or improved learning.

In fact, the TileSearch, the Marquees and Querylist generated with the weighted-selection approach in the MeetHub Search, as well as the BOOC Player successfully induced serendipitous encountering. Especially for the latter system, we have observed significant increase in the frequency of book interactions for the DD groups during video watching after the introduction of the BOOC Player. Other prototypes failed in this aspect. Whether or not a system induced serendipity can be seen as an assessment of how well the systems were designed according to the design principles. As discussed in Chapter 7, the RaindropSearch, the WordCloud and the Marquees generated with the typing-triggered approach did not achieve sufficient trigger-richness. The Marquees based on the combinational approach were overwhelmed with redundant information.

Group Facilitator

Serendipitous interactions with contextual information scents may exert systematic effects on group discussions in learning activities. In our studies, we measured such effects with survey questions that were concerned with participants' subjective ratings regarding various aspects of group discussions, such as the quality of discussions. The TileSearch system was found to reduce the imbalance of collaboration, which is a desired attribute of group work. However, it also reduced the effectiveness of communication, which we believe was caused by the low relevance of the information scents. We also found that the BOOC Player significantly increased the discussion quality for the MOOC study groups who shared the same display and video control.

Learning Support

By learning support, we mean interacting with the contextual information scents was helpful for the collaborative learning process, e.g. inspiring new ideas in brainstorming, or helping resolve difficulty in video watching etc. Our participants did not perceive the TileSearch to be helpful in the idea generation phase, perhaps because they were used to brainstorm alone or through discussions. In fact, the more they interacted with the system, the less ideas were generated. A first possible reason was more Web engaging time led to less time left for brainstorming. Secondly, the interactions were shown to improve imbalance by catching group's shared attention, but classical brainstorming research claimed that group participation might inhibit creative thinking. The Marquee and Querylist in the MeetHub Search did not show overall perceived usefulness, but the contextual information scents generated with the WA were significantly more useful than before, perhaps due to higher relevance of results produced by this query-building approach.

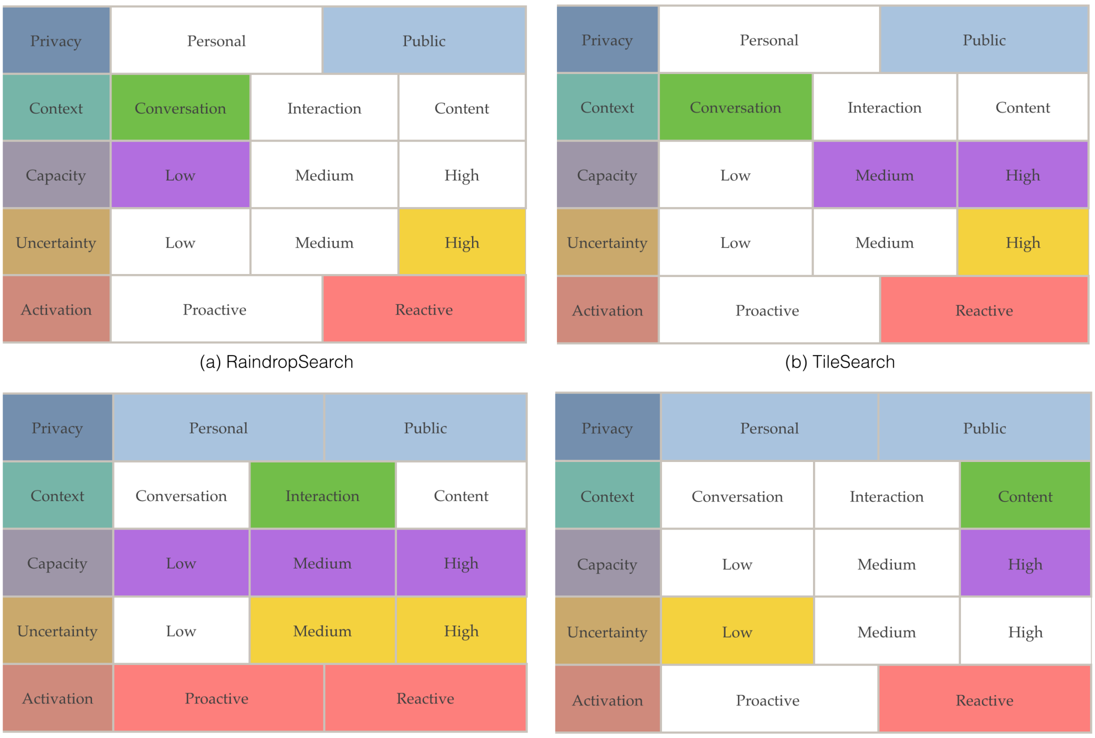
However, high relevance does not necessarily lead to increased learning support. In the collaborative MOOC viewing activity, the book pages in the BOOC Player were perfectly relevant to the video being played. For more difficult videos, the groups were found to have significantly more discussions rather than more interactions with the information scents. This finding actually reflects the group dynamics in collaborative learning. In Section 5.2.1 we modeled learning activities with the Activity Theory. Engeström's Activity System Triangle (cf. Figure 5.1) intuitively illustrates that students' completion of learning tasks can be mediated not only by tools, but also rules, community and division of labor. Clearly in the collaborative MOOC learning activity, the students are habituated to turn to the community, i.e. the group for help in case they encounter difficulties, and the contextual information scents were only found to have augmented group discussions. In other words, the *zone of intervention* largely intersects with the *zone of proximal development*, and students prefer to get help from *more knowledgeable others* rather than technological interventions. We further investigated how individual students learn MOOCs, and uncovered the *zone of intervention* for contextual information scents in Chapter 9.

10.2 Roundup

Finally, this section serves for rounding up the whole dissertation, by identifying its core contributions, limitations and future work.

10.2.1 Contributions

As described in the Introduction chapter, this thesis was set out to answer three research questions: (1) What types of situational information needs may arise during learning activities and what are the challenges, principles and potential design space for augmenting the activities with contextual information scents? (2) How can ambient technologies be designed as contextual information scents and what are the benefits and overall appeal of them? And (3) Can big educational data (MOOC) provide insights for designing contextual information scents? These questions were answered through the discussions from Chapter 5 to Chapter 9, and a reflection of them was presented in the preceding section. To sum up, the thesis overall delivers three contributions to the HCI and learning fields and we will discuss them in this section.



(a) RaindropSearch (b) TileSearch

(c) MeetHub Search (d) BOOC Player

Figure 10.1: Explorations of the design space

Design Framework

After identifying elastic information needs as the intended target and a few challenges to be addressed in this dissertation, we proposed a set of design principles, a design space and a specification of interaction phases (cf. Chapter 5), which altogether constitute the design framework for contextual information scents. The principles are derived from the identified challenges and are based on theories of information seeking, ambient information interaction and serendipitous encountering. The design space consists of 5 axes - privacy, context, capacity, uncertainty and activation. As context is the key concern of contextual information scents, we proposed three different contexts, i.e. conversation, interaction and content, and explored each of them by combining different characteristics in the other 4 axes. It is impractical to cover all possible combinations in this thesis, but we have developed several prototypes to exploit each type of context, as shown in Figure 10.1.

Designing contextual information scents to augment learning activities is a new research topic. Our explorations based on the framework offer a broad glance into the design space, with an in-depth view of different design choices and how they fit into learning activities. We believe the findings shed light on the practical design of contextual information scents. The design space may also guide practitioners and researchers to explore alternative solutions.

Design Implications

Throughout this thesis, no matter which dimension of the design space we explored, we always faced the challenges as presented in Section 5.2.1: distractions, relevance and timeliness. All the design efforts were devoted to minimize distractions, increase relevance and ensure timeliness. We summarize the design implications as follows:

- **Minimizing distractions**. Minimizing distractions were mainly achieved by varying information capacity and activation schemes. We find contextual information scents with proactive activation are prone to distractions, but the such distractions may fade away with increased relevance. Low capacity information scents usually led to the extreme case of calmness, i.e. ignorance. A more plausible explanation could be the information scents were only intended to facilitate early stages of the information seeking process, i.e. formulating queries. General implication to minimize distraction could be reactively displaying information of higher capacity, higher relevance to facilitate later stages of the information seeking process.

- **Increasing relevance**. In our explorations, we attempted to increase the relevance of contextual information scents by varying the context and information uncertainty. Conversations are sometimes noisy and out of context, it is difficult to extract keywords for searching information. Groupware interactions are more focused than conversations, but we find interactions create context that may lag behind the information needs. Video content is the most direct reflection of video viewing context, and supervised contextual information scents largely reduce the information uncertainty. As a result, the last prototype successfully played a role in group facilitation. The implication is that high relevance and low uncertainty are seemingly the desired attribute of contextual information scents. Note that relevance does not equal to precision. As one of the design principles suggests, relevant information can also be redundant. As technology advances, computer-driven information elicitation approaches may be able to mimic man's supervised effort. Then highly relevant contextual information scents can be achieved in all kinds of contexts.

- **Ensure timeliness**. Timeliness is solely based on the activation scheme. In our explorations, timeliness is ensured in the BOOC Player, because the intervention was pre-programmed in advance. This might not be the case for other learning activities. Our solution was to introduce redundancy by letting the system proactively deliver potentially useful contextual information scents to the users, e.g. in the MeetHub Search system. According to the *principle of guaranteed result*, this approach maximized the coverage. However, it also introduced distractions. A possible solution could be the employment of a weighting scheme as in the WA-generated Marquees, so that relevance is increased. Meanwhile the design may not be too disturbing.

An important concern apart from the perspective of the above three challenges is the influence of being in a group. This is especially noteworthy for the designs targeted for collaborative activities, since the *zone of proximal development* largely intersects with the *zone of intervention*, so that contextual information scents are not always essential. That explains our motivation to **augment** learning activities by supporting **elastic information needs**. For individual learning activities such as MOOC learning, we proposed a data inference approach to evaluate the *zone of intervention* for contextual information scents.

Evaluation Methodology

Contextual information scents are special. They inherit the core attributes from information seeking technologies, ambient information systems and serendipitous inducing systems. Ambient information systems usually attract occasional use within a short period of time. Serendipity inducing systems do not always successfully induce serendipity each time due to its opportunistic nature. We cannot evaluate contextual information scents with traditional quantitative evaluation methodologies, since we do not often have sufficient interaction data. The methodologies that are often applied to ambient information systems (cf. Section 4.1.3) are of qualitative focus. In this thesis, we adopted a different set of evaluation methodologies, which can be considered as part of our research contributions. These methodologies include:

- **Activity Theory**. The Activity Checklist (Kaptelinin et al., 1999) was applied as the major evaluation approach for the early design prototype, the RaindropSearch. The analysis with this tool allowed us to understand the role and context in a more structured manner. The intersection of *zone of proximal development* and *zone of intervention* in collaborative activities can also be identified by another Activity Theory tool, the Activity System Triangle (Engeström, 1987).
- **Longitudinal study**. The interactions with contextual information scents are not intense. So, in two of the presented studies we employed a longitudinal approach to study how contextual information scents were interacted with over weeks. Occasional interactions are coded by examining the video recordings, which allowed us to obtain both quantitative and qualitative findings.
- **Levels of benefits**. The studies of contextual information scents presented in this thesis were based on the analysis of their benefits in three levels. At the lowest level, the information scents attract serendipitous interactions. Next, these serendipitous interactions may have a systematic effect on facilitating group discussion, e.g. quality of discussion and balance of contributions. At the highest level, these interactions may support learning in certain aspects, e.g. help generate more ideas or help resolve difficulty. Categorization of the potential added value allows us to gain in-depth understanding of the benefits of the contextual information scents.
- **Data inference**. This approach is especially suitable for understanding large-scale learn-

ing activities such as MOOCs. We use data inference to evaluate the *zone of intervention* for contextual information scents.

Actually none of the above methodologies is new in the field of HCI, but we have demonstrated how these methods can be combined for analyzing ambient, serendipitous interactions anchored by the contextual information scents. This is the main methodological contribution made by this work.

10.2.2 Limitations and Future Work

With one survey study, two lab experiments, two longitudinal studies and one data analytics study, this dissertation has employed a rich set of HCI theories and methodologies for the exploration of contextual information scents. However, the research topic is new, so there is room for improvement. This section summarizes the limitations of our work and the prospects for future work.

Limitations

The limitations of the thesis can be summarized in two aspects:

- **Technical limitation**. When we explored the conversation context and interaction context, we used a simple approach to extract nouns as keywords for the prototypes (Chapter 6), and employed commercial keyword extraction for the prototypes (Chapter 7). The latter was a blackbox API and we could not control how it worked internally. In addition, Google and Bing search engines were used to return search results, but they are also blackboxes that increased uncertainty. All of these are technical limitations in the information elicitation process.
- **Methodological limitation**. There exist different kinds of learning activities. Our investigation began with seminar activities, but all the system designs were explored for collaborative learning activities. In addition, the learning tasks in different studies were not consistent, which hinders a more comprehensive comparison of the designs. Although we attempt to generalize our results, the some findings are specific to particular learning tasks and learning activities.
- **Lack of habituation**. The idea of augmenting learning activities with contextual information scent is new. Students in a study group usually discuss with other members first before turning to technologies or textbook for help. They also prefer to seek information on their own instead of relying on the suggested information. The design of information scent is an important factor that influences their appeal to the users. However, human behaviors in habituated tasks can not be easily changed. It is true that we presented two studies that lasted for weeks, but longer period of time may be required for users to habituate themselves to contextual information scents.

Future Work

Based on our research findings and limitations, future work can be directed to:

- **Combination of contexts**. In this dissertation we approximated the use of contextual information scents generated by different contexts separately. Clearly, there is room left for combing different contexts, given the pros and cons of each context uncovered by the studies. For example, groupware interactions are more focused, but our study showed that they might lag behind the occurrences of information needs. Probably interaction contexts can be combined with conversational context to balance between timeliness and relevance. In a similar vein, the content context as studied in the collaborative MOOC viewing study was solely based on video content, without considering group discussions and video interactions. A combination of context of different kinds may increase the chance for identifying the complement between the *zone of intervention* and *zone of proximal development*, so that contextual information scents may also be effective in helping the group resolve difficulty.
- **Focused learning activity**. As studies in this work are concerned with different learning activities, future work may focus on one type of learning activity, e.g. MOOC learning. Specific design principles and evaluation methodologies can then be developed. We have provided some design insights in Chapter 9, but no concrete designs were implemented and experimented. This can be a next step.
- **Long-term habituation**. Future work may include long-term studies to examine how users habituated themselves to contextual information scents. Probably MOOC platforms are potential testbeds for conducting such long-term experiments.

Closing Remarks

My research presented in this dissertation offers a first glance into contextual information scent. Context-awareness is a hot research topic in the field of HCI, and contextual information scents are designed to combine context-awareness with ambient information and serendipity inducing systems. We believe that in learning activities the learners from time to time have elastic information needs, which can potentially be supported by contextual information scents with minimal efforts. In the course of this thesis we presented several design prototypes to explore how contextual information scents can be designed, deployed and evaluated, and revealed their benefits and overall appeal of use. In fact, true context-awareness is difficult to be achieved technically. However, we argue it can be approximated by considering certain contextual factors. This dissertation covers conversation, interaction and content contexts, but other types of contexts can also be explored in the future. As technologies (e.g. data science or machine learning) advance, human behaviors can be better predicted. Relevant and timely information scents can then reliably augment learning activities of many kinds.

A Information Need Survey

The following survey was distributed to the participants of the 6 seminar talks presented in Section 5.1.

A.1 Information Need Survey in Seminars

1. I know very much about the topic before the presentation.
   Strongly Disagree 1 2 3 4 5 Strongly Agree

2. I understand very well of the topic after the presentation.
   Strongly Disagree 1 2 3 4 5 Strongly Agree

3. I am interested in this topic.
   Strongly Disagree 1 2 3 4 5 Strongly Agree

4. This topic is related to my work / research.
   Strongly Disagree 1 2 3 4 5 Strongly Agree

5. Please tick (✔) as appropriate (you may have multiple choices)
   I have a ***(1) smart phone (2) tablet (3) laptop (4) paper and pen (5) nothing*** *with me during the presentation*

6. Have you had any information needs (you feel the need to search something for whatever reasons, no matter if you actually made the search or not) during the presentation? Please note them down.
   (1) ______________________
   (2) ______________________
   (3) ______________________
   (4) ______________________
   (5) ______________________

The following section is repeated for each information need specified above

For information need (1), please answer the following questions.

Where is the need from, the slides or the oral presentation?

____

Why did you have this information need?

____

Did you search for it during the presentation on your laptop / tablet / phone? Specify the tool you used.

____

If no, then why not? Please tick (✔) as appropriate (you may have multiple choices)
(1) I have no tools. Yes No
(2) I have search tools, but I didn't use them for searching, because:
* I felt impolite to search during the presentation. ***Strongly Disagree 1 2 3 4 5 Strongly Agree***
* I didn't want to show my ignorance in front of others. ***Strongly Disagree 1 2 3 4 5 Strongly Agree***
* I was using the tool for note taking in the meanwhile. ***Strongly Disagree 1 2 3 4 5 Strongly Agree***
* I was too lazy to make the search. ***Strongly Disagree 1 2 3 4 5 Strongly Agree***
* I didn't want to miss part of the talk, searching needs time. ***Strongly Disagree 1 2 3 4 5 Strongly Agree***
* For Other reasons, please specify ______________________________________

If you did make searches, answer question (1), otherwise answer question (2)
(1) I searched for it during the presentation, which improved my understanding of the talk.
Strongly Disagree 1 2 3 4 5 Strongly Agree

(2) I searched for it during the presentation, which improved my understanding of the talk.
Strongly Disagree 1 2 3 4 5 Strongly Agree

Does the need still exist after the presentation? If no, why?

____

If yes, then are you going to search for it after the seminar or ask the presenter?

____

B RaindropSearch Study: Questionnaire

The following questionnaire was distributed to the participants after they completed the group task with the RaindropSearch system.

B.1 RaindropSearch Questionnaire

Section I: About yourself

1. What age group are you in?

☐ 19 and under
☐ 20 – 29
☐ 30 – 39
☐ 40 – 49
☐ 50 – 59

2. What age group are you in?

☐ Male
☐ Female

3. What age group are you in?

☐ Natural Sciences and Mathematics
☐ Engineering / Technology and Computer Science
☐ Life and Agriculture Science
☐ Clinical Medicine and Pharmacy
☐ Social Sciences
☐ Art
Other, please specify:

4. Rank your knowledge about power plants (1 – 5 from least to most):

(1) Nuclear plant	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
(2) Wind farm	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
(3) Solar plant	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
(4) Fossil fuel plant	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
(5) Hydroelectric plant	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
(6) Tidal plant	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
(7) Sea wave power plant	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

Section II: Feedback about the System

1. Do you think the task is easy? What do you find most difficult about the tasks?

2. Did the system provide good key words for search from your conversation? Give examples. Were there any words that you searched because you saw them on the table and which you would not have searched otherwise?

3. What do you think of interacting with the infrared pen? Do you have any difficulties?

4. How do you feel the paper browser interaction?

5. Leave some feedback about the system. Are there any general interaction problems?

C TileSearch Study: Instruction and Questionnaire

The following appendices include instructions of the experiment and pre-test questionnaire that were distributed to the participants before the start of the experiment, as well as the post-questionnaires that were asked to filled out after the first and second task.

C.1 Experiment Instructions

1. Introduction

In this experiment you will perform two brainstorming tasks in groups of three. Each brainstorming task should be done in two phases.

- *Divergent thinking phase*. The goal of the divergent thinking phase is to produce a large quantity of ideas, with no judgment of ideas in terms of good or bad.
- *Convergent thinking phase*. In the convergent thinking phase, you will need group, further develop, classify and prioritize suggestions obtained in divergent thinking.

2. Time

This experiment will take approximately 40 minutes.

- You will have 2 mins to fill in a Pre-test questionnaire
- You will have 7 mins to do the divergent thinking phase for the first task.
- You will have 7 mins to do the convergent thinking phase for the first task.
- You will have 5 mins to fill in a Post-test questionnaire for the first task.
- You will have 7 mins to do the divergent thinking phase for the second task.
- You will have 7 mins to do the convergent thinking phase for the second task.
- You will have 5 mins to fill in a Post-test questionnaire for the second task.

3. Instructions

We introduce you to the experimental task through the following welcoming procedure:

1. *During the divergent phase, each of you should write down the idea on a Post-it note and quickly paste it close to you on the table*

2. *During the convergent phase, you have to collect all the ideas your group had, judge them and present the best set of ideas in the end.*

4. Tasks

(1) Future Car Task

Imagine cars in the future may have lots of novel features that are not available nowadays. Your task is to brainstorm in your group about what innovative features a car would have in the future. Please brainstorm as many ideas as possible in the divergent thinking phase and judge which of the ideas are possible to be realized in 20 years during the convergent thinking phase.

(2) Future Home Task

Homes will probably become smart in the future, with the advancement of technologies. Your task is to brainstorm in your group about the innovative features, which you can imagine for a smart home. Please brainstorm as many ideas as possible in the divergent thinking phase and judge which of the ideas are possible to be realized in 20 years during the convergent thinking phase.

C.2 Pre-test Questionnaire

1. Which group are you in? Please indicate your group name.

☐ Apple
☐ Apricot
☐ Avocado
☐ Cherry

2. Please indicate your sitting position with respect to facing the lamp.

☐ Left (user 1)
☐ Middle (user 2)
☐ Right (user 3)

3. Your age is ______________

4. Gender

☐ Male
☐ Female

5. Background

☐ Natural science and Mathematics
☐ Engineering / technology and computer science
☐ Life and agriculture science
☐ Clinical medicine and pharmacy
☐ Social science
☐ Art
☐ Others

6. I prefer working :

☐ in groups
☐ individually

7. I see myself an _______ person, when it comes to group work.

Extrovert 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Introvert

8. Rank your familiarity with user 1 (left user)

Not known 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Very familiar

9. Rank your familiarity with user 2 (middle user)

Not known 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Very familiar

10. Rank your familiarity with user 3 (right user)

Not known 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Very familiar

11. How often do you attend meetings (e.g. project meetings, brainstorming etc.)

☐ Natural science and Mathematics

☐ Engineering / technology and computer science

☐ Life and agriculture science

12. What is the typical duration of a meeting that you usually attend?

☐ Less than 30 mins

☐ 30 mins – 2 hours

☐ More than 2 hours

13. It is best when all members of a group participate equally in the meeting task.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

14. How often do you use your mobile phone / tablet to search information during a meeting?

Very often 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Never

15. How often do you use your laptop to search information during a meeting?

Very often 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Never

16. If you remember you searched information on the Web during a meeting. Please (1) name the type of the meeting; (2) describe what you searched and why.

____

____

____

17. Have you ever used an interactive tabletop ?

Very often 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Never

C.3 Post-test Questionnaire 1 (After the first task)

1. Which group are you in? Please indicate your group name.

☐ Apple
☐ Apricot
☐ Avocado
☐ Cherry

2. Please indicate your sitting position with respect to facing the lamp.

☐ Left (user 1)
☐ Middle (user 2)
☐ Right (user 3)

3. I actively participated in the task.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

4. How many ideas have you contributed in the first divergent thinking phase? ______________

5. How many ideas have you finalized in the convergent thinking phase? ______________

6. Ideas that I contributed have been included in the final list

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

7. How many ideas do you think user 1 contributed? ______________

8. How many ideas do you think user 2 contributed? ______________

9. How many ideas do you think user 3 contributed? ______________

10. Someone dominated the brainstorming session for a certain period.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

11. I work closely with the other members of my group to accomplish this task.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

12. The members of the group communicated with each other effectively.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

13. The group worked effectively as a team.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

14. The tabletop system was a source of inspiration of new ideas for me as an individual.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

15. The tabletop system was a source of inspiration of new ideas for the group as a whole.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

16. The system was useful in the DIVERGENT thinking phase for this task.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

17. The system was useful in the CONVERGENT thinking phase for this task.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

18. Touch interaction with the search view was easy and intuitive.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

19. Touch interaction with the Web view was easy and intuitive.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

20. Sometimes I wanted to make a search with my own keywords.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

C.4 Post-test Questionnaire 2 (After the second task)

1. Which group are you in? Please indicate your group name.

☐ Apple
☐ Apricot
☐ Avocado
☐ Cherry

2. Please indicate your sitting position with respect to facing the lamp.

☐ Left (user 1)
☐ Middle (user 2)
☐ Right (user 3)

3. I actively participated in the task.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

4. How many ideas have you contributed in the first divergent thinking phase? ______________

5. How many ideas have you finalized in the convergent thinking phase? ______________

6. Ideas that I contributed have been included in the final list

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

7. How many ideas do you think user 1 contributed? ______________

8. How many ideas do you think user 2 contributed? ______________

9. How many ideas do you think user 3 contributed? ______________

10. Someone dominated the brainstorming session for a certain period.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

11. I work closely with the other members of my group to accomplish this task.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

12. The members of the group communicated with each other effectively.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

13. The group worked effectively as a team.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

14. The tabletop system was a source of inspiration of new ideas for me as an individual.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

15. The tabletop system was a source of inspiration of new ideas for the group as a whole.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

16. The system was useful in the DIVERGENT thinking phase for this task.
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

17. The system was useful in the CONVERGENT thinking phase for this task.
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

18. Touch interaction with the search view was easy and intuitive.
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

19. Touch interaction with the Web view was easy and intuitive.
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

20. Sometimes I wanted to make a search with my own keywords.
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

21. Compared to the previous condition, image stimuli is more useful than Wikipedia stimuli in the DIVERGENT thinking phase.
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

22. Compared to the previous condition, image stimuli is more useful than Wikipedia stimuli in the CONVERGENT thinking phase.
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

21. What is your overall impression about the system? Would you use such a system in meetings? If yes, in what type of meetings?

____

____

____

22. Do you need multiple browsers? Do you need multi-touch features, e.g. scale or rotate the browser?

____

____

____

23. Do you like the interaction with a single display horizontal display? Or you prefer a vertical screen? Or you like to have both kind in the system?

____

____

____

D MeetHub Search Study: Tasks and Questionnaires

The following appendices include descriptions of the three experiment tasks as well as the post-experiment questionnaire.

D.1 Task 1: The Thumb Problem

We don't think this is very likely to happen, but imagine for a moment what would happen if everyone born after 2013 had an extra thumb on each hand. This extra thumb will be built just as the present one is, but located on the other side of the hand. It faces inward, so that it can press against the fingers, just as the regular thumb does now. Here is a picture to help you see how it will be.

Now the question is:

What **practical benefits** and **difficulties** will arise when people start having this extra thumb? Please brainstorm in your group. You will have **TWO** phases for this task. In the first 20 minutes, you should generate as many ideas as possible **WITHOUT** judgment, criticism or evaluation, and in the last 10 minutes you will have to **JUSTIFY** your idea list and come out with a final list of the most practical ones. Free free to adjust the time allocated for each phase during your discussion with the time management tool on your iPad.

D.2 Task 2: The Energy Crisis Problem

Much of central China is enduring its worst energy crisis, with factories and residents facing power cuts as supply runs short of demand. This fast-growing country has long experienced periodic power shortages, especially in winter and summer when weather extremes boost demand for heating and cooling. Thermal power station is still the major type of power stations, and the coal fuels contribute to about three-quarters of the country's electricity generation.

China Electricity Council has published statistics about power shortfall in the past five years as illustrated in the following table:

Year	Power Shortage (beyond current installed capacity)
2007	14 million Kilowatts
2008	19 million Kilowatts
2009	22 million Kilowatts
2010	26 million Kilowatts
2011	30 million Kilowatts

Considering China's economic and population growth, it is estimated that the power shortfall will be steadily increasing for the next five years and then reach its saturation point in the year 2016. Suppose you are a group of consultants hired by China Energy Council and your task is to analyze the given statistics, estimate the power needs (only the power shortage) and design an energy plan to solve the energy crisis by the end of 2021. Please decide on the types and the numbers of power plants to be built and an estimated cost. You have to present good enough reasons both for the government and power companies. That being said, although you are not limited to a certain amount of money or a certain area in China to build power plants, you must consider environmental factors and the cost/benefit ratio.

	Installed Capacity	Construction cost	Operation cost	Selling Price
Nuclear	1-3 million Kilowatts	13000 ￥/Kilowatt	0.5 ￥/Kwh	1.3 ￥/Kwh
Wind farm	0.1-0.3 million Kilowatts	6500 ￥/Kilowatt	0.4 ￥/Kwh	1.2 ￥/Kwh
Solar Energy	0.01-0.15 million Kilowatts	9000 ￥/Kilowatt	0.6 ￥/Kwh	1.4 ￥/Kwh
Fossil fuel (coal)	0.6-4 million Kilowatts	4800 ￥/Kilowatt	0.8 ￥/Kwh	1.2 ￥/Kwh
Fossil fuel (gas)	0.5-2.3 million Kilowatts	3400 ￥/Kilowatt	0.9 ￥/Kwh	1.4 ￥/Kwh
Hydroelectric	0.5-3 million Kilowatts	3000 ￥/Kilowatt	0.7 ￥/Kwh	1.3 ￥/Kwh
Tidal	0.0005-0.004 million Kilowatts	8000 ￥/Kilowatt	0.8 ￥/Kwh	1.3 ￥/Kwh

D.3 Task 3: The Neuroscience Task

This time you will be given 15 minutes to read a neuroscience text before starting the experiment.

Every neuron has an electrical voltage on both sides of the membrane that is called the "membrane potential". The neuron at rest (which does not transmit nerve impulses) generally has a membrane potential of about -65 mV. The membrane potential of a non-stimulated neuron is called the "resting potential". This negative potential can be explained by the fact that the interior of the neuron is negatively charged while its exterior is positively charged. Thus it is said that the neuron is polarized.

The resting potential exists only across the membrane; in other words, the liquids that are inside and outside the neuron are electrically neutral. The resting potential is generated by differences in the ionic composition of interior and exterior environments. Thus, the inside of the neuron contains a smaller concentration of sodium (Na +) and a higher concentration of potassium (K +) than the outside. In the extracellular fluid, the positive charges of sodium ions are generally balanced by chloride ions (Cl-). In the intracellular fluid, negatively charged proteins (A-) facilitate the equalization/balancing of the positive charges of potassium ions (K +).

The ionic differences arise on the one hand from the difference in ionic permeability of the membrane, and on the other hand from the operation of the sodium-potassium pump. In the resting state, the membrane is about 75 times more permeable to K + than to Na +. This resting permeability is bound to the properties of passive ion channels in the membrane.

The concentration gradients of K + and Na + ions explain their diffusion from the medium where they are most concentrated to the medium where they are least concentrated, that means towards the exterior of the neuron for the K + ions and towards the interior for the Na + ions. Furthermore, K + ions diffuse more rapidly than sodium ions. From this follows that the positive ions that diffuse outward are a little more numerous than those which diffuse inward, leaving a small surplus of negative charges inside the neuron; this phenomenon leads to an imbalance of electric charges (electrical gradient) which causes the resting potential.

As there is always a certain quantity of K + leaving the cell and a certain amount of Na + that enters it, one might think that the concentration of Na + and K + ions on both sides of the membrane will equalize, which would lead to the disappearance of their respective concentration gradients. However, this is not the case because the sodium-potassium pump exchanges (the) Na + ions from the interior with the K+ ions from the exterior of the neuron. In other words, the K + ions are pumped into the neuron at the same time as the Na + ions are released to the outside.

You have the following three tasks to complete:

1. Compare the roles of Na+ and K+ in neuro-transmission.
2. Draw a schematic neuron which illustrates the generation of resting potential.
3. Assume that you are a group of TAs for a neuron science course. Please design an assignment to check whether or not your students understand the concept illustrated in this article.

D.4 Post-test Questionnaire

Section I : About content sharing

The following questions are about content sharing in your meeting.

- Which tool did you use more for writing.
  ☐ Keyboard ☐ iPad
  ☐ Both iPad and Keyboard ☐ Neither of them (Didn't Write)

- Which tool did you use more for creating new objects, moving and deleting.
  ☐ Mouse ☐ iPad
  ☐ Both iPad and Mouse ☐ Neither of them

- Which input tool would you prefer in meetings.
  ☐ iPad ☐ Mouse and Keyboard
  ☐ Both ☐ No Preference

- The usage of Pen/Stylus was intuitive with the iPad.
  Strongly Disagree ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 Strongly Agree

- The group reached a consensus at the end of the meeting.
  Strongly Disagree ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 Strongly Agree

- I feel that my contributions were taken into account by the group.
  Strongly Disagree ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 Strongly Agree

- At which display did you look more during the experiment.
  ☐ iPad ☐ Whiteboard (Public Display)

- During the discussions, in which direction most of your gestures were made to.
  ☐ iPad ☐ Whiteboard (Public Display)
  ☐ Other participants

- I think that the meeting environment (Table, Public Display, iPads and Stylus, Mouse & Keyboards) facilitated group coordination effectively (Disagree … Agree).
  Strongly Disagree ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 Strongly Agree

Section 2 : About time management

The following questions are about awareness of time management in your meeting.

- In how many parts did you split your meeting ?
  ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

- Did you observe the blinking time management notification on the public display ?
  ☐ Yes ☐ No

- What did you mostly do when time was up for one of the phases ?
  ☐ We continued to discuss regardless of bypassing the allocated time.
  ☐ We extended the discussion a little, but kept in mind that we needed to move on.
  ☐ We quickly discussed/noted down some conclusions in order to move on fast.
  ☐ We moved on immediately because there was already a consensus.
  ☐ We moved on immediately even though information or consensus was still lacking.

- For how long did you typically extend your discussion over the allocated time ?
  ☐ 0 minutes (not at all)
  ☐ At most 5 minutes
  ☐ 5 to 15 minutes
  ☐ We were not aware that we bypassed the allocated time.

- At the end of the meeting, did you feel that you managed your time proficiently?
  ☐ Yes, the pre-suggested time allocation actually corresponded to our management.
  ☐ We tended to over-discuss items but the awareness helped us keep on time track.
  ☐ We tended to finish our collaboration faster than the allocated times.
  ☐ We did not care about the time management.

- Please rank the overall utility/efficiency/ appropriateness of having a time management awareness element in your meeting.
  ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

- Please let us know if you have any comments regarding how you addressed the time management in your meeting.
  ____
  ____
  ____

Section III : About search activities

The following questions are about the searches in your meetings. The term "search suggestion" refers to the "moving rectangle" on the whiteboard, which contains two image blocks and one Wikipedia block.

- The search suggestions are NOT disturbing or intrusive to the discussion (Disagree ... Agree)
  Strongly Disagree ☐ ☐ ☐ ☐ ☐ Strongly Agree
  1 2 3 4 5

- The search suggestions helped me with accomplishing the task (Disagree ... Agree)
  Strongly Disagree ☐ ☐ ☐ ☐ ☐ Strongly Agree
  1 2 3 4 5

- I am satisfied with the number of suggestions for each search (Disagree ... Agree)
  Strongly Disagree ☐ ☐ ☐ ☐ ☐ Strongly Agree
  1 2 3 4 5

- I think Wikipedia suggestions were more helpful than images (Disagree ... Agree)
  Strongly Disagree ☐ ☐ ☐ ☐ ☐ Strongly Agree
  1 2 3 4 5

- I think the keyword suggestions extracted for websites (on the browser window) were useful (Disagree ... Agree)
  Strongly Disagree ☐ ☐ ☐ ☐ ☐ Strongly Agree
  1 2 3 4 5

- From the search suggestions, list the image/Wikipedia link that you thought were useful during the discussion (if there were any), and explain why it was useful. Was the suggestion (image/wikipedia) itself useful or the web page containing the suggestion is useful?
  ____
  ____
  ____
  ____

- List keywords that you thought might be useful for making searches during the discussion (if there were any)
  ____
  ____
  ____
  ____

E BOOC Player Study: Questionnaires

The following appendices include the pre-test questionnaire that was filled out before the study session in the first week, as well as a repeating post-test questionnaire that were answered by the subjects each week after the study session.

E.1 Pre-test Questionnaire

Section A : Personal Information

The following questions are concerned about your personal details.

Full Name: ______________

Gender:

☐ Male

☐ Female

Age: ________

Study Major

☐ Mechanical Engineering

☐ Communication Systems

☐ Computer Science

☐ Physics

☐ Management of Technology

☐ Other: __________

Your current semester: _________

Section B : Personality Questions

The following questions assess your personality. For each statement, please place your opinion on the scale, ranging from "Strongly disagree" to "Strongly agree"

I consider myself as extraverted, enthusiastic.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I consider myself as critical, quarrelsome.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I consider myself as dependable, self-disciplined.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I consider myself as anxious, easily upset.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I consider myself as open to new experiences, complex.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I consider myself as reserved, quiet.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I consider myself as sympathetic, warm.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I consider myself as disorganized, careless.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I consider myself as calm, emotionally stable.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I consider myself as conventional, uncreative.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

Section C : Familiarity with other group members

How well do you know each person in this study group? For each person, please rate the degree to which you know him/her. For your own name, please ignore the question.

Please indicate your current seat label: __________ (Labels from A-E are pasted on the table)

I know person A very well.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I know person B very well.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I know person C very well.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I know person D very well.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

I know person E very well.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

Section D : Studying in groups

This section is concerned with your group studying experiences.

During the semester, how often do you study as part of a group?

☐ Never (I always study alone)

☐ Less often than once a month

☐ Once a month or more

☐ Once a week or more

☐ Once a day or more

If you have been involved in any study groups before, please describe your typical group practices. E.g. how often you meet, for which course, with how many people, where you meet and what you usually do.

____

____

Section E : MOOC use and perceptions

This section is concerned with MOOC learning experiences.

Before the current MOOC, how many MOOC have you attempted or completed?

☐ None, this is my first MOOC

☐ One

☐ Two

☐ Three

☐ Four or more

If this is NOT your first MOOC, please tell us about the previous MOOCs you took. E.g. which courses you registered, and whether or not you completed them, etc.

____

____

I regard MOOC-based learning as being more effective and efficient than attending real classes.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ Strongly agree

Section F : Use of iPad and other tablet devices

This section is concerned about experiences with tablet devices.

Access to iPad or other tablet device (multiple selections possible)

☐ I never used to own or have access to an iPad or other tablet device

☐ I currently have an iPad for personal use

☐ I currently have another kind of table device (e.g. Android) for personal use

☐ I currently have access to a shred iPad or other kinds of tablet devices

☐ I used to have access to an iPad or other tablet device but I am not currently using one

How long have you used a tablet device (iPad or other)?

☐ I have never used a tablet device

☐ I have tested a tablet device a few times

☐ I have used a tablet device for less than a half year

☐ I have used a table device for half a year to 1 year

☐ I have used a tablet device for 1 to 2 years

☐ I have used a table device for 2 to 3 years

☐ I have used a tablet device for more than 3 years

E.2 Post-test Questionnaire

Section A : Learning and participation of the study session

This week's MOOC lecture content were:

Very easy 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Very difficult

This week's quizzes / homework were:

Very easy 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Very difficult

I am happy with what I learned during the study session:

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Strongly agree

The quality of discussions in the group were :

Very bad 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Very good

Everyone contributed to the discussion :

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Strongly agree

Section B : Emotional State Assessment

Please indicate your OVERALL emotional feelings (with respect to MOOC content and your team work) to the study session with the following emotional scale

Suspicious	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	Accepting
Disinterested	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	Interested
Dissatisfied	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	Satisfied
Irritated	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	Cheerful
Guided	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	In control
Dominant	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	Submissive
Apathetic	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	Excited
Sleepy	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	Wide-awake
Despared	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐	6 ☐	7 ☐	Hopeful

Section C1 : Use of the textbook (Before the BOOC Player, i.e. Week 1-3)

Since last week's study group session, did you use the book at home?

☐ Yes

☐ No

Did you use the book during today's study session?

☐ Yes

☐ No

If yes, did the book help you understand the videos?

Not at all 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Very much

If yes, please note down the pages of the book that you remember reading in the study session.

____

If you have never used the book during the study session, can you tell us why?

____

Section C2 : Usage of the textbook (After the BOOC Player, i.e. Week 4-5)

Since last week's study group session, did you use the book at home?

☐ Yes

☐ No

Please indicate the level of use of the new video-book mapping feature

☐ I did not pay attention to the feature

☐ I glanced at the page content a few times

☐ I spent time reading the page content

☐ Other: _______________

I find the video-book mapping feature very useful.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Strongly agree

Please give us your opinions about the video-book mapping feature.

____

Did you use the PRINT book during today's study session?

☐ Yes

☐ No

If yes, I was influenced by the recommendation of the video-book mapping feature.

Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Strongly agree

If yes, did the textbook help you understand the videos?
Note at all 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Very much

If yes, please note down the pages of the book that you remember reading in the study session.

If you have never used the book during the study session, can you tell us why?

Section D : Watching MOOCs

During today's study session, when a video was difficult to understand, what did you do?

I replayed the segment which I did not understand
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Strongly agree

I paused the video and though by myself
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Strongly agree

I paused the video and discussed with others
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Strongly agree

I consulted the book (week 1-3)
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Strongly agree

I consulted the print book because the PDF display guided me to specific pages (week 4-5)
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Strongly agree

I consulted the print book without being affected by the PDF display on the iPad (week 4-5)
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Strongly agree

Please explain your answers to the questions above.

So far, I am highly satisfied with learning through MOOCs in general
Strongly disagree 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ Strongly agree

The number of quiz problems I have solved during the study session was: ______
(please leave a number here (In one "quiz" set, there might be more than one problems))

The number of courser homework problems I have solved during the study session was : ____
(for Analyse Numerique students only)

Bibliography

Adamczyk, P. D. and Bailey, B. P. (2004). If not now, when?: the effects of interruption at different moments within task execution. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 271–278. ACM.

Alavi, H. S. and Dillenbourg, P. (2012). An ambient awareness tool for supporting supervised collaborative problem solving. *Learning Technologies, IEEE Transactions on*, 5(3):264–274.

Aleven, V., Mclaren, B., Roll, I., and Koedinger, K. (2006). Toward meta-cognitive tutoring: A model of help seeking with a cognitive tutor. *International Journal of Artificial Intelligence in Education*, 16(2):101–128.

Aleven, V., Stahl, E., Schworm, S., Fischer, F., and Wallace, R. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research*, 73(3):277–320.

Allport, D. A., Antonis, B., and Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *The Quarterly journal of experimental psychology*, 24(2):225–235.

Alt, F., Schneegaß, S., Schmidt, A., Müller, J., and Memarovic, N. (2012). How to evaluate public displays. In *Proceedings of the 2012 International Symposium on Pervasive Displays*, page 17. ACM.

Altmann, E. M. and Trafton, J. G. (2002). Memory for goals: An activation-based model. *Cognitive science*, 26(1):39–83.

Altmann, E. M. and Trafton, J. G. (2004). Task interruption: Resumption lag and the role of cues. Technical report, DTIC Document.

Ames, R. and Lau, S. (1982). An attributional analysis of student help-seeking in academic settings. *Journal of Educational Psychology*, 74(3):414.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., and Qin, Y. (2004). An integrated theory of the mind. *Psychological review*, 111(4):1036.

Anderson, J. R. and Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, 96(4):703.

Anderson, L. W., Krathwohl, D. R., and Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* Allyn & Bacon.

André, P., Teevan, J., Dumais, S. T., et al. (2009). Discovery is never by chance: designing for (un) serendipity. In *Proceedings of the seventh ACM conference on Creativity and cognition*, pages 305–314. ACM.

Attfield, S., Blandford, A., Dowell, J., and Cairns, P. (2008). Uncertainty-tolerant design: Evaluating task performance and drag-and-link information gathering for a news-writing task. *International journal of human-computer studies*, 66(6):410–424.

Bachour, K., Kaplan, F., and Dillenbourg, P. (2008). Reflect: An interactive table for regulating face-to-face collaborative learning. In *Times of Convergence. Technologies Across Learning Contexts*, pages 39–48. Springer.

Bailey, B. P. and Konstan, J. A. (2006). On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Computers in Human Behavior*, 22(4):685–708.

Bailin, S., Case, R., Coombs, J. R., and Daniels, L. B. (1999). Conceptualizing critical thinking. *Journal of Curriculum Studies*, 31(3):285–302.

Bakker, S., Van Den Hoven, E., and Eggen, B. (2013). Fireflies: physical peripheral interaction design for the everyday routine of primary school teachers. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, pages 57–64. ACM.

Bakker, S., Van Den Hoven, E., Eggen, B., and Overbeeke, K. (2012). Exploring peripheral interaction design for primary school teachers. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, pages 245–252. ACM.

Balaam, M., Fitzpatrick, G., Good, J., and Luckin, R. (2010). Exploring affective technologies for the classroom with the subtle stone. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1623–1632, New York, NY, USA. ACM.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online review*, 13(5):407–424.

Bates, M. J. (2002). Toward an integrated model of information seeking and searching. *The New Review of Information Behaviour Research*, 3:1–15.

Bdker, S. (1991). *Through the Interface: A Human Activity Approach to User Interface Design.* L. Erlbaum Associates Inc., Hillsdale, NJ, USA.

Belanger, Y., Thornton, J., and Barr, R. C. (2013). Bioelectricity: A quantitative approach–duke university's first mooc. *EducationXPress*, 2013(2):1–1.

Benjamin, W., Chandrasegaran, S., Ramanujan, D., Elmqvist, N., Vishwanathan, S., and Ramani, K. (2014). Juxtapoze: supporting serendipity and creative expression in clipart compositions. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 341–350. ACM.

Benway, J. P. and Lane, D. M. (1998). Banner blindness: Web searchers often miss “obvious” links. *Itg Newsletter*, 1(3).

Bergstrom, T., Harris, A., and Karahalios, K. (2011). Encouraging initiative in the classroom with anonymous feedback. In *Proceedings of the 13th IFIP TC 13 International Conference on Human-computer Interaction - Volume Part I*, INTERACT’11, pages 627–642, Berlin, Heidelberg. Springer-Verlag.

Bergstrom, T. and Karahalios, K. (2009a). Conversation clusters: grouping conversation topics through human-computer dialog. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2349–2352. ACM.

Bergstrom, T. and Karahalios, K. (2009b). Social mirrors as social signals: transforming audio into graphics. *Computer Graphics and Applications, IEEE*, 29(5):22–32.

Blandford, A. and Attfield, S. (2010). Interacting with information. *Synthesis Lectures on Human-Centered Informatics*, 3(1):1–99.

Bloom, B. S. et al. (1956). Taxonomy of educational objectives. 1. cognitive domain. *New York: David McKay*.

Bogunovich, P. and Salvucci, D. (2011). The effects of time constraints on user behavior for deferrable interruptions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3123–3126. ACM.

Bonnard, Q., Lemaignan, S., Zufferey, G., Mazzei, A., Cuendet, S., Li, N., and Dillenbourg, P. (2013). Chilitags 2: Robust fiducial markers for augmented reality and robotics.

Boswell, J. (1873). *The Life of Samuel Johnson*. William P. Nimmo.

Branch, P., Egan, G., and Tonkin, B. (1999). Modeling interactive behaviour of a video based multimedia system. In *Communications, 1999. ICC’99. 1999 IEEE International Conference on*, volume 2, pages 978–982. IEEE.

Brehm, S. S. and Brehm, J. W. (2013). *Psychological reactance: A theory of freedom and control*. Academic Press.

Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., and Seaton, D. T. (2013). Studying learning in the worldwide classroom: Research into edx’s first mooc. *Research & Practice in Assessment*, 8(1):13–25.

Brignull, H. and Rogers, Y. (2003). Enticing people to interact with large public displays in public spaces. In *Proceedings of INTERACT*, volume 3, pages 17–24.

Brinton, C. G., Chiang, M., Jain, S., Lam, H., Liu, Z., and Wong, F. M. F. (2014). Learning about social learning in moocs: From statistical analysis to generative model. *Learning Technologies, IEEE Transactions on*, 7(4):346–359.

Brixey, J. J., Robinson, D. J., Johnson, C. W., Johnson, T. R., Turley, J. P., and Zhang, J. (2007). A concept analysis of the phenomenon interruption. *Advances in Nursing Science*, 30(1):E26–E42.

Broadbent, D. E. (1958). *Perception and communication.* Pergamon Press, New York, NY.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Univ of California Press.

Buzikashvili, N. (2005). Information searching behavior: Between two principles. In *Context: Nature, Impact, and Role*, pages 79–95. Springer.

Byron, M. (2004). *Satisficing and maximizing: Moral theorists on practical reason.* Cambridge University Press.

Cadiz, J., Venolia, G. D., Jancke, G., and Gupta, A. (2001). Sideshow: Providing peripheral awareness of important information. *Microsoft Research, Collaboration, and Multimedia Group, Tech. Rep.*

Cadiz, J. J., Balachandran, A., Sanocki, E., Gupta, A., Grudin, J., and Jancke, G. (2000). Distance learning through distributed collaborative video viewing. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, pages 135–144. ACM.

Campos, J. and Figueiredo, A. D. d. (2001). Searching the unsearchable: Inducing serendipitous insights. In *Proceedings of the Workshop Program at the Fourth International Conference on Case-Based Reasoning, ICCBR.*

Capdeferro, N. and Romero, M. (2012). Are online learners frustrated with collaborative learning experiences? *The International Review of Research in Open and Distributed Learning*, 13(2):26–44.

Capobianco, A. and Carbonell, N. (2001). Contextual online help: elicitation of human experts' strategies. In *Proceedings of HCI*, volume 1, pages 824–828. Citeseer.

Card, S. K., Mackinlay, J. D., and Robertson, G. G. (1990). The design space of input devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 117–124. ACM.

Carenini, G. and Moore, J. D. (1993). Generating explanations in context. In *Proceedings of the 1st international conference on Intelligent user interfaces*, pages 175–182. ACM.

Carroll, J. M. (1991). *Designing interaction: Psychology at the human-computer interface*, volume 4. CUP Archive.

Carroll, J. M., Neale, D. C., Isenhour, P. L., Rosson, M. B., and McCrickard, D. S. (2003). Notification and awareness: synchronizing task-oriented collaborative activity. *International Journal of Human-Computer Studies*, 58(5):605–632.

Carton, A. M. and Aiello, J. R. (2009). Control and anticipation of social interruptions: Reduced stress and improved task performance1. *Journal of Applied Social Psychology*, 39(1):169–185.

Case, D. O. (2012). *Looking for information: A survey of research on information seeking, needs and behavior*. Emerald Group Publishing.

Case, Donald, O. (2002). Looking for information.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical population biology*, 9(2):129–136.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, 25(5):975–979.

Collins, A. M. and Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review*, 82(6):407.

Connaway, L. S., Dickey, T. J., and Radford, M. L. (2011). "if it is too inconvenient i'm not going after it:" convenience as a critical factor in information-seeking behaviors. *Library & Information Science Research*, 33(3):179–190.

Cool, C. and Xie, H. I. (2004). How can ir help mechanisms be more helpful to users? *Proceedings of the American Society for Information Science and Technology*, 41(1):249–255.

Cunha, M. P. E. (2005). Serendipity: Why some organizations are luckier than others.

Curtis, D. D. and Lawson, M. J. (2001). Exploring collaborative online learning. *Journal of Asynchronous learning networks*, 5(1):21–34.

Cutrell, E., Czerwinski, M., and Horvitz, E. (2001). Notification, disruption, and memory: Effects of messaging interruptions on memory and performance. In *Human-Computer Interaction–Interact '01*, pages 263–269. IOS Press.

Czerwinski, M., Horvitz, E., and Wilhite, S. (2004). A diary study of task switching and interruptions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 175–182. ACM.

De Bruijn, O. and Spence, R. (2001). Serendipity within a ubiquitous computing environment: A case for opportunistic browsing. In *Ubicomp 2001: Ubiquitous Computing*, pages 362–369. Springer.

De Bruijn, O. and Spence, R. (2008). A new framework for theory-based interaction design applied to serendipitous information retrieval. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 15(1):5.

Derr, R. L. (1983). A conceptual analysis of information need. *Information Processing & Management*, 19(5):273–278.

Dervin, B. (1983). *An Overview of Sense-making Research: Concepts, Methods, and Results to Date.* Sense-making packet. The Author.

Dervin, B. (1997). Given a context by any other name: Methodological tools for taming the unruly beast. *Information seeking in context*, 13:38.

Deutsch, J. A. and Deutsch, D. (1963). Attention: some theoretical considerations. *Psychological review*, 70(1):80.

Dewey, J. (1931). Context and thought.

Dey, A. K., Kortuem, G., Morse, D. R., and Schmidt, A. (2001). Situated interaction and context-aware computing. *Personal and Ubiquitous Computing*, 5(1):1–3.

Dey-Sircar, J. K., Salehi, J. D., Kurose, J. F., and Towsley, D. (1994). Providing vcr capabilities in large-scale video servers. In *Proceedings of the second ACM international conference on Multimedia*, pages 25–32. ACM.

Dillenbourg, P. (1999). What do you mean by collaborative learning? *Collaborative-learning: Cognitive and Computational Approaches.*, pages 1–19.

Dillenbourg, P., Baker, M. J., Blaye, A., and O'Malley, C. (1995). The evolution of research on collaborative learning. *Learning in Humans and Machine: Towards an interdisciplinary learning science.*, pages 189–211.

DiMicco, J. M. and Bender, W. (2007). Group reactions to visual feedback tools. In *Persuasive Technology*, pages 132–143. Springer.

DiMicco, J. M., Pandolfo, A., and Bender, W. (2004). Influencing group participation with a shared display. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 614–623. ACM.

Dolnicar, S., Grabler, K., Mazanec, J. A., Woodside, A., Crouch, G., Oppermann, M., Sakai, M., et al. (1999). *A tale of three cities: perceptual charting for analysing destination images.* CABI Publishing.

Dworman, G. and Rosenbaum, S. (2004). Helping users to use help: improving interaction with help systems. In *CHI'04 extended abstracts on Human factors in computing systems*, pages 1717–1718. ACM.

Eisenberg, D., Downs, M. F., Golberstein, E., and Zivin, K. (2009). Stigma and help seeking for mental health among college students. *Medical Care Research and Review*, 66(5):522–541.

Eisenberg, D., Golberstein, E., and Gollust, S. E. (2007). Help-seeking and access to mental health care in a university student population. *Medical care*, 45(7):594–601.

Elliot, A. J. and Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of personality and social psychology*, 72(1):218.

Ellis, D. (1989). A behavioural approach to information retrieval system design. *Journal of documentation*, 45(3):171–212.

Elsweiler, D., Mandl, S., and Kirkegaard Lunn, B. (2010). Understanding casual-leisure information needs: a diary study in the context of television viewing. In *Proceedings of the third symposium on Information interaction in context*, pages 25–34. ACM.

Engeström, Y. (1987). Learning by expanding.

Erdelez, S. (1997). Information encountering: a conceptual framework for accidental information discovery. In *Proceedings of an international conference on Information seeking in context*, pages 412–421. Taylor Graham Publishing.

Erdelez, S. (1999). Information encountering: It's more than just bumping into information. *Bulletin of the American Society for Information Science and Technology*, 25(3):26–29.

Erdelez, S. (2004). Investigation of information encountering in the controlled research environment. *Information Processing & Management*, 40(6):1013–1025.

Eyrolle, H. and Cellier, J.-M. (2000). The effects of interruptions in work activity: Field and laboratory results. *Applied ergonomics*, 31(5):537–543.

Ferla, J., Valcke, M., and Schuyten, G. (2010). Judgments of self-perceived academic competence and their differential impact on students' achievement motivation, learning approach, and academic performance. *European Journal of Psychology of Education*, 25(4):519–536.

Fischler, I. (1977). Semantic facilitation without association in a lexical decision task. *Memory & Cognition*, 5(3):335–339.

Fisher, J. (1999). Technical communicators and online help: The developers' and users' perspectives. *Technical Communication: Journal of the Society for Technical Communication*, 46(3):360–65.

Fisher, J. D., Nadler, A., and Whitcher-Alagna, S. (1983). Four conceptualizations of reactions to aid. *New directions in helping*, 1:51–84.

Fjeld, M., Morf, M., and Krueger, H. (2004). Activity theory and the practice of design: evaluation of a collaborative tangible user interface. *International Journal of Human Resources Development and Management*, 4(1):94–116.

Flores, F. and Winograd, T. (1986). Understanding computers and cognition. *Norwood, NJ, Ablex, USA.*

Foster, A. and Ford, N. (2003). Serendipity and information seeking: an empirical study. *Journal of Documentation*, 59(3):321–340.

Fourney, A., Lafreniere, B., Chilana, P., and Terry, M. (2014). Intertwine: creating interapplication information scent to support coordinated use of software. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, pages 429–438. ACM.

Frese, M. and Zapf, D. (1994). Action as the core of work psychology: A german approach. *Handbook of industrial and organizational psychology*, 4:271–340.

Freund, L. and Toms, E. G. (2013). Contextual search: from information behaviour to information retrieval. In *Proceedings of the Annual Conference of CAIS/Actes du congrès annuel de l'ACSI*.

Freund, L. S. (1990). Maternal regulation of children's problem-solving behavior and its impact on children's performance. *Child development*, 61(1):113–126.

Gagné, R. M. (1985). *The conditions of learning and theory of instruction*. Holt, Rinehart and Winston New York.

Gibbons, J. F., Kincheloe, W. R., and Down, K. S. (1977). Tutored videotape instruction: A new use of electronics media in education. *Science*, 195(4283):1139–1146.

González, V. M. and Mark, G. (2004). Constant, constant, multi-tasking craziness: managing multiple working spheres. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 113–120. ACM.

Gray, J. A. and Wedderburn, A. A. I. (1960). Shorter articles and notes grouping strategies with simultaneous stimuli. *Quarterly Journal of Experimental Psychology*, 12(3):180–184.

Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., and Wang, M. (2011). Proxemic interactions: The new ubicomp? *interactions*, 18(1):42–50.

Groome, D. (2013). *An introduction to cognitive psychology: Processes and disorders*. Psychology Press.

Grossman, T. and Fitzmaurice, G. (2010). Toolclips: an investigation of contextual video assistance for functionality understanding. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1515–1524. ACM.

Gueddana, S. and Roussel, N. (2009). Effect of peripheral communication pace on attention allocation in a dual-task situation. In *Human-Computer Interaction–INTERACT 2009*, pages 111–124. Springer.

Guo, P. J. and Reinecke, K. (2014). Demographic differences in how students navigate through moocs. In *Proceedings of the first ACM conference on Learning@ scale conference*, pages 21–30. ACM.

Hacker, W. (2003). Action regulation theory: A practical tool for the design of modern work processes? *European Journal of work and organizational psychology*, 12(2):105–130.

Halawa, S., Greene, D., and Mitchell, J. (2014). Dropout prediction in moocs using learner activity features. *Experiences and best practices in and around MOOCs*, page 7.

Hall, E. T. (1969). *The hidden dimension*, volume 1990. Anchor Books New York.

Hangal, S., Nagpal, A., and Lam, M. (2012). Effective browsing and serendipitous discovery with an experience-infused browser. In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*, pages 149–158. ACM.

Hausen, D. and Butz, A. (2011). Extending interaction to the periphery. *Embodied Interaction: Theory and Practice in HCI, in conjunction with CHI*.

Heer, R. (2015). Blooms revised taxonomy (text-only), center for excellence in learning and teaching.

Heift, T. (2006). Context-sensitive help in call. *Computer Assisted Language Learning*, 19(2-3):243–259.

Hennessey, M. G. (1999). Probing the dimensions of metacognition: Implications for conceptual change teaching-learning.

Heyman, G. D. (2008). Children's critical thinking when learning from others. *Current Directions in Psychological Science*, 17(5):344–347.

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American psychologist*, 44(3):513.

Hollan, J., Hutchins, E., and Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(2):174–196.

Huang, J., Dasgupta, A., Ghosh, A., Manning, J., and Sanders, M. (2014). Superposter behavior in mooc forums. In *Proceedings of the first ACM conference on Learning@ scale conference*, pages 117–126. ACM.

Hunter, S. and Maes, P. (2008). Wordplay: A table-top interface for collaborative brainstorming and decision making. *Proceedings of IEEE Tabletops and Interactive Surfaces*, pages 2–5.

Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive science*, 19(3):265–288.

Inkpen, K., Whittaker, S., Czerwinski, M., Fernandez, R., and Wallace, J. (2009). Groupbanter: Supporting serendipitous group conversations with im. In *Collaborative Computing: Networking, Applications and Worksharing*, pages 485–498. Springer.

Iqbal, S. T., Grudin, J., and Horvitz, E. (2011). Peripheral computing during presentations: perspectives on costs and preferences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 891–894. ACM.

Ishii, H., Wisneski, C., Brave, S., Dahley, A., Gorbet, M., Ullmer, B., and Yarin, P. (1998). ambientroom: integrating ambient media with architectural space. In *CHI 98 Cconference Summary on Human Factors in Computing Systems*, pages 173–174. ACM.

James, W. (1890). 1950. the principles of psychology.

Jean-Louis, L., Zouaq, A., Gagnon, M., and Ensan, F. (2014). An assessment of online semantic annotators for the keyword extraction task. In *PRICAI 2014: Trends in Artificial Intelligence*, pages 548–560. Springer.

Jett, Q. R. and George, J. M. (2003). Work interrupted: A closer look at the role of interruptions in organizational life. *Academy of Management Review*, 28(3):494–507.

Jiang, S., Williams, A., Schenke, K., Warschauer, M., and O'dowd, D. (2014). Predicting mooc performance with week 1 behavior. In *Educational Data Mining 2014*.

Jin, J. and Dabbish, L. A. (2009). Self-interruption on the computer: a typology of discretionary task interleaving. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1799–1808. ACM.

Johnson, J. D. (2003). On contexts of information seeking. *Information Processing & Management*, 39(5):735–760.

Jonassen, D. H. and Rohrer-Murphy, L. (1999). Activity theory as a framework for designing constructivist learning environments. *Educational Technology Research and Development*, 47(1):61–79.

Kahneman, D. (1973). *Attention and effort*. Citeseer.

Kaptelinin, V. (2006). *Acting with technology: Activity theory and interaction design*. Mit Press.

Kaptelinin, V., Nardi, B., Bodker, S., Carroll, J., Hollan, J., Hutchins, E., and Winograd, T. (2003). Post-cognitivist hci: second-wave theories. In *CHI'03 extended abstracts on Human factors in computing systems*, pages 692–693. ACM.

Kaptelinin, V., Nardi, B. A., and Macaulay, C. (1999). Methods & tools: The activity checklist: a tool for representing the "space" of context. *interactions*, 6(4):27–39.

Karabenick, S. A. (1998). *Strategic help seeking: Implications for learning and teaching*. Routledge.

Kelley, H. H. et al. (1972). *Causal schemata and the attribution process*. General Learning Press Morristown, NJ.

Kelly, G. A. (1963). *A theory of personality: The psychology of personal constructs*. W.W., New York.

Khan, S. A. and Shafique, F. (2011). Information needs and information-seeking behavior: A survey of college faculty at bahawalpur.

Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., and Miller, R. C. (2014). Understanding in-video dropouts and interaction peaks inonline lecture videos. In *Proceedings of the first ACM conference on Learning@ scale conference*, pages 31–40. ACM.

Kim, T., Chang, A., Holland, L., and Pentland, A. S. (2008). Meeting mediator: enhancing group collaborationusing sociometric feedback. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pages 457–466. ACM.

Kizilcec, R. F., Piech, C., and Schneider, E. (2013). Deconstructing disengagement: analyzing learner subpopulations in massive open online courses. In *Proceedings of the third international conference on learning analytics and knowledge*, pages 170–179. ACM.

Knight, S. and Spink, A. (2008). *Toward a Web search information behavior model.* Springer.

Konig, C. J., Buhner, M., and Murling, G. (2005). Working memory, fluid intelligence, and attention are predictors of multitasking performance, but polychronicity and extraversion are not. *Human Performance*, 18(3):243–266.

Kop, R., Fournier, H., and Mak, J. S. F. (2011). A pedagogy of abundance or a pedagogy to support human beings? participant support on massive open online courses. *The International Review Of Research In Open And Distributed Learning*, 12(7):74–93.

Korpela, M. (1997). Activity analysis and development in a nutshell. *Retrieved July*, 24:2006.

Korpela, M., Soriyan, H. A., and Olufokunbi, K. C. (2001). Activity analysis as a method for information systems development. *Scandinavian Journal of Information Systems*, 12(1-2):191–210.

Kuhlthau, C. (2005). Information search process.

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *JASIS*, 42(5):361–371.

Kuhlthau, C. C. (1993). A principle of uncertainty for information seeking. *Journal of documentation*, 49(4):339–355.

Kuhlthau, C. C. (2004). *Seeking meaning. A process approach to library and information services.* Libraries Unlimited, Westport, Conn.[u.a.], 2. edition.

Kuhlthau, C. C., Maniotes, L. K., and Caspari, A. K. (2007). *Guided inquiry: Learning in the 21st century.* Greenwood Publishing Group.

Kunifuji, S., Kato, N., and Wierzbicki, A. P. (2007). Creativity support in brainstorming. In *Creative Environments*, pages 93–126. Springer.

Kuutti, K. (1991). Activity theory and its applications to information systems research and development. *Information systems research: Contemporary approaches and emergent traditions*, pages 529–549.

Kuutti, K. (1996). Activity theory as a potential framework for human-computer interaction research. *Context and consciousness: Activity theory and human-computer interaction*, pages 17–44.

Lafreniere, B., Grossman, T., and Fitzmaurice, G. (2013). Community enhanced tutorials: improving tutorials with multiple demonstrations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1779–1788. ACM.

Lamberty, K. K., Adams, S., Biatek, J., Froiland, K., and Lapham, J. (2011). Using a large display in the periphery to support children learning through design. In *Proceedings of the 10th International Conference on Interaction Design and Children*, IDC '11, pages 62–71, New York, NY, USA. ACM.

Lee, C. J. (2007). Academic help seeking: theory and strategies for nursing faculty. *The Journal of nursing education*, 46(10):468–475.

Leibowitz, H. W. and Post, R. B. (1982). The two modes of processing concept and some implications. *Organization and representation in perception*, pages 343–363.

Leont'ev, A. (1978). N.(1978). activity, consciousness and personality.

Leshed, G., Perez, D., Hancock, J. T., Cosley, D., Birnholtz, J., Lee, S., McLeod, P. L., and Gay, G. (2009). Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 537–546. ACM.

Li, V. O., Liao, W., Qiu, X., and Wong, E. W. (1996). Performance model of interactive video-on-demand systems. *Selected Areas in Communications, IEEE Journal on*, 14(6):1099–1109.

Liestman, D. (1992). Chance in the midst of design: approaches to library research serendipity. *RQ*, pages 524–532.

Liu, Z. and Yang, Z. Y. L. (2004). Factors influencing distance-education graduate students' use of information sources: A user study. *The Journal of Academic Librarianship*, 30(1):24–35.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological review*, 95(4):492.

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review*, 7(4):618–630.

Luckin, R., Du Boulay, B., et al. (1999). Ecolab: The development and evaluation of a vygotskian design framework. *International Journal of Artificial Intelligence in Education*, 10(2):198–220.

Lugmayr, A., Risse, T., Stockleben, B., Laurila, K., and Kaario, J. (2009). Semantic ambient media—an introduction. *Multimedia Tools and Applications*, 44(3):337–359.

MacArthur, R. H. and Pianka, E. R. (1966). On optimal use of a patchy environment. *American Naturalist*, pages 603–609.

Macdonald, J. S. and Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics*, 73(6):1780–1789.

MacIntyre, B., Mynatt, E. D., Voida, S., Hansen, K. M., Tullio, J., and Corso, G. M. (2001). Support for multitasking and background awareness using interactive peripheral displays. In *Proceedings of the 14th annual ACM symposium on User interface software and technology*, pages 41–50. ACM.

Mack, A. and Rock, I. (1998). Inattentional blindness: Perception without attention. *Visual attention*, 8:55–76.

Makri, S. and Blandford, A. (2012). Coming across information serendipitously-part 1: A process model. *Journal of Documentation*, 68(5):684–705.

Makri, S., Blandford, A., Woods, M., Sharples, S., and Maxwell, D. (2014). "making my own luck": Serendipity strategies and how to support them in digital information environments. *Journal of the Association for Information Science and Technology*, 65(11):2179–2194.

Mamykina, L., Mynatt, E., and Terry, M. A. (2001). Time aura: Interfaces for pacing. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 144–151. ACM.

Mankoff, J., Dey, A. K., Hsieh, G., Kientz, J., Lederer, S., and Ames, M. (2003). Heuristic evaluation of ambient displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 169–176. ACM.

Manning, C. D. and Schütze, H. (1999). *Foundations of statistical natural language processing*. MIT press.

Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, 49(4):41–46.

Marchionini, G. and Komlodi, A. (1998). Design of interfaces for information seeking. *Annual review of information science and technology*, 33:89–130.

Marchionini, G. and White, R. (2007). Find what you need, understand what you find. *International Journal of Human [# x02013] Computer Interaction*, 23(3):205–237.

Mark, G., Gonzalez, V. M., and Harris, J. (2005). No task left behind?: examining the nature of fragmented work. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 321–330. ACM.

Mark, G., Gudith, D., and Klocke, U. (2008). The cost of interrupted work: more speed and stress. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 107–110. ACM.

Marshall, A. (1890). *Principles of Economics*. Number v. 1 in Principles of Economics. Macmillan and Company.

Martin, F. G. (2012). Will massive open online courses change how we teach? *Communications of the ACM*, 55(8):26–28.

Martins, L. E. G. and Daltrini, B. M. (1999). An approach to software requirements elicitation using the precepts from activity theory. In *2011 26th IEEE/ACM International Conference on Automated Software Engineering (ASE 2011)*, pages 15–15. IEEE Computer Society.

Matthews, T., Rattenbury, T., and Carter, S. (2007). Defining, designing, and evaluating peripheral displays: An analysis using activity theory. *Human–Computer Interaction*, 22(1-2):221–261.

Matthews, T., Rattenbury, T., Carter, S., Dey, A., and Mankoff, J. (2003). A peripheral display toolkit.

Matthews, T. L. (2007). Designing and evaluating glanceable peripheral displays.

Maxwell, D., Woods, M., Makri, S., Bental, D., Kefalidou, G., and Sharples, S. (2012). Designing a semantic sketchbook to create opportunities for serendipity. In *Proceedings of the 26th Annual BCS Interaction Specialist Group Conference on People and Computers*, pages 357–362. British Computer Society.

McCay-Peet, L. and Toms, E. G. (2010). The process of serendipity in knowledge work. In *Proceedings of the third symposium on Information interaction in context*, pages 377–382. ACM.

McCay-Peet, L. and Toms, E. G. (2015). Investigating serendipity: How it unfolds and what may influence it. *Journal of the Association for Information Science and Technology*.

McCay-Peet, L., Toms, E. G., and Kelloway, E. K. (2015). Examination of relationships among serendipity, the environment, and individual differences. *Information Processing & Management*, 51(4):391–412.

McCrickard, D. S., Chewar, C. M., Somervell, J. P., and Ndiwalana, A. (2003). A model for notification systems evaluation—assessing user goals for multitasking activity. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 10(4):312–338.

Mcleod, P. (1977). A dual task response modality effect: Support for multiprocessor models of attention. *The Quarterly Journal of Experimental Psychology*, 29(4):651–667.

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. Psychology Press.

Meho, L. I. and Tibbo, H. R. (2003). Modeling the information-seeking behavior of social scientists: Ellis's study revisited. *Journal of the American society for Information Science and Technology*, 54(6):570–587.

Mellon, J. R. A. R. K. et al. (2007). *How can the human mind occur in the physical universe?* Oxford University Press.

Melucci, M. (2012). Contextual search: A computational framework. *Foundations and Trends in Information Retrieval*, 6(4-5):257–405.

Merrill, M. D. (1983). Component display theory. *Instructional-design theories and models: An overview of their current status*, 1:282–333.

Meyer, D. E. and Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations. *Journal of experimental psychology*, 90(2):227.

Miyata, Y. and Norman, D. A. (1986). Psychological issues in support of multiple activities. *User centered system design: New perspectives on human-computer interaction*, pages 265–284.

Mizzaro, S. (1997). Relevance: The whole history. *JASIS*, 48(9):810–832.

Mongy, S., Djeraba, C., and Simovici, D. A. (2007). On clustering users' behaviors in video sessions. In *DMIN*, pages 99–103. Citeseer.

Monsell, S. (2003). Task switching. *Trends in cognitive sciences*, 7(3):134–140.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly journal of experimental psychology*, 11(1):56–60.

Morris, M. R. (2013). Collaborative search revisited. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1181–1192. ACM.

Morris, M. R., Lombardo, J., and Wigdor, D. (2010). Wesearch: supporting collaborative search and sensemaking on a tabletop display. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, pages 401–410. ACM.

Morville, P. and Callender, J. (2010). *Search patterns: design for discovery.* " O'Reilly Media, Inc.".

Mwanza, D. (2002). Conceptualising work activity for cal systems design. *Journal of Computer Assisted Learning*, 18(1):84–92.

Myrhaug, H. I. and Göker, A. (2003). Ambiesense–interactive information channels in the surroundings of the mobile user. In *Univeral Access in HCI, 10th International Conference on Human-Computer Interaction*, volume 4, pages 1158–1162.

Nadler, A. and Fisher, J. D. (1986). The role of threat to self-esteem and perceived control in recipient reaction to help: Theory development and empirical validation. *Advances in experimental social psychology*, 19:81–122.

Nardi, B. A. (1996). Activity theory and human-computer interaction. *Context and consciousness: Activity theory and human-computer interaction*, pages 7–16.

Navon, D. and Gopher, D. (1979). On the economy of the human-processing system. *Psychological review*, 86(3):214.

Neisser, U. (1963). Decision-time without reaction-time: Experiments in visual scanning. *The American Journal of Psychology*, pages 376–385.

Neisser, U. (1964). Visual search. *Scientific American*.

Neisser, U. (1967). *Cognitive psychology*. Century psychology series. Appleton-Century-Crofts.

Neisser, U. (1979). The control of information pickup in selective looking. *Perception and its development: A tribute to Eleanor J. Gibson*, pages 201–219.

Neisser, U. and Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive psychology*, 7(4):480–494.

Neisser, U., Novick, R., and Lazar, R. (1963). Searching for ten targets simultaneously. *Perceptual and motor skills*, 17(3):955–961.

Nelson-Le Gall, S. (1981). Help-seeking: An understudied problem-solving skill in children. *Developmental Review*, 1(3):224–246.

Nelson-Le Gall, S. (1985). Help-seeking behavior in learning. *Review of research in education*, pages 55–90.

Newman, R. S. (1994). Adaptive help seeking: A strategy of self-regulated learning.

Newman, R. S. (2002). How self-regulated learners cope with academic difficulty: The role of adaptive help seeking. *Theory into Practice*, 41(2):132–138.

Newman, R. S. (2008). Adaptive and nonadaptive help seeking with peer harassment: An integrative perspective of coping and self-regulation. *Educational Psychologist*, 43(1):1–15.

Nielsen, J. (2005). Ten usability heuristics.

Nielsen, J. and Molich, R. (1990). Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 249–256. ACM.

Nordlie, R. (1996). Unmediated and mediated information searching in the public library. In *Proceedings of the ASIS Annual Meeting*, volume 33, pages 41–46. ERIC.

Norman, D. A. (1980). Twelve issues for cognitive science. *Cognitive science*, 4(1):1–32.

Norman, D. A. (1988). *The psychology of everyday things*. Basic books.

Norman, D. A. and Shallice, T. (1986). *Attention to action*. Springer.

Occhialini, V., van Essen, H., and Eggen, B. (2011). Design and evaluation of an ambient display to support time management during meetings. In *Human-Computer Interaction–INTERACT 2011*, pages 263–280. Springer.

O'Day, V. L. and Jeffries, R. (1993). Orienteering in an information landscape: How information seekers get from here to there. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, CHI '93, pages 438–445, New York, NY, USA. ACM.

Ogan, A., Walker, E., Baker, R., Rodrigo, M. M. T., Soriano, J. C., and Castro, M. J. (2014). Towards understanding how to assess help-seeking behavior across cultures. *International Journal of Artificial Intelligence in Education*, pages 1–20.

Piaget, J. (1970). Science of education and the psychology of the child. trans. d. coltman.

Pirolli, P. (1997). Computational models of information scent-following in a very large browsable text collection. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, pages 3–10. ACM.

Pirolli, P. (2006). The use of proximal information scent to forage for distal content on the world wide web. *Adaptive Perspectives on Human-Technology Interaction: Methods and Models for Cognitive Engineering and Human-Computer Interaction*, pages 247–266.

Pirolli, P. and Card, S. (1999). Information foraging. *Psychological review*, 106(4):643.

Polson, M. C. and Richardson, J. J. (2013). *Foundations of intelligent tutoring systems*. Psychology Press.

Pongnumkul, S., Dontcheva, M., Li, W., Wang, J., Bourdev, L., Avidan, S., and Cohen, M. F. (2011). Pause-and-play: automatically linking screencast video tutorials with applications. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 135–144. ACM.

Pousman, Z. and Stasko, J. (2006). A taxonomy of ambient information systems: four patterns of design. In *Proceedings of the working conference on Advanced visual interfaces*, pages 67–74. ACM.

Prante, T., Röcker, C., Streitz, N., Stenzel, R., Magerkurth, C., Van Alphen, D., and Plewe, D. (2003). Hello. wall–beyond ambient displays. In *Adjunct Proceedings of Ubicomp*, pages 277–278. Citeseer.

Previc, F. H. (1998). The neuropsychology of 3-d space. *Psychological bulletin*, 124(2):123.

Purchase, H. C. and Worrill, J. (2002). An empirical study of on-line help design: features and principles. *International Journal of Human-Computer Studies*, 56(5):539–566.

Pyke, G. H., Pulliam, H. R., and Charnov, E. L. (1977). Optimal foraging: a selective review of theory and tests. *Quarterly Review of Biology*, pages 137–154.

Quek, A. and Shah, H. (2004). A comparative survey of activity-based methods for information systems development. In *ICEIS (5)*, pages 221–232.

Redström, J., Skog, T., and Hallnäs, L. (2000). Informative art: using amplified artworks as information displays. In *Proceedings of DARE 2000 on Designing augmented reality environments*, pages 103–114. ACM.

Reigeluth, C. M. and Moore, J. (1999). Cognitive education and the cognitive domain. *Instructional-design theories and models: A new paradigm of instructional theory*, 2:51–68.

Remer, T. G. (1965). *Serendipity and the three princes: From the Peregrinaggio of 1557.* Norman, U. Oklahoma P.

Robertson, S. E. (1977). Theories and models in information retrieval. *Journal of Documentation*, 33(2):126–148.

Rocchi, C., Tomasini, D., Stock, O., and Zancanaro, M. (2008). Fostering conversation after the museum visit: a woz study for a shared interface. In *Proceedings of the working conference on Advanced visual interfaces*, pages 335–338. ACM.

Röcker, C. (2009). Toward smart office environments-benefits and drawbacks of using ambient intelligence technologies in knowledge-based enterprises. In *Proc. of the Intern. Conference on Economics, Business, Management and Marketing (EBMM 2009)*, pages 17–21.

Roman, F. (2013). *Interaction Mirrors for Face-to-Face Meetings.* PhD thesis, IC, Lausanne.

Rubin, V. L., Burkell, J., and Quan-Haase, A. (2011). Facets of serendipity in everyday chance encounters: a grounded theory approach to blog analysis. *Information Research*, 16(3).

Russell-Rose, T., Lamantia, J., and Burrell, M. (2011). A taxonomy of enterprise search and discovery. *Proceedings of HCIR.*

Ryan, A. M., Gheen, M. H., and Midgley, C. (1998). Why do some students avoid asking for help? an examination of the interplay among students' academic efficacy, teachers' social–emotional role, and the classroom goal structure. *Journal of educational psychology*, 90(3):528.

Ryan, A. M., Pintrich, P. R., and Midgley, C. (2001). Avoiding seeking help in the classroom: Who and why? *Educational Psychology Review*, 13(2):93–114.

Salvucci, D. D., Monk, C. A., and Trafton, J. G. (2009a). A process-model account of task interruption and resumption: When does encoding of the problem state occur? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 53, pages 799–803. SAGE Publications.

Salvucci, D. D. and Taatgen, N. A. (2008). Threaded cognition: an integrated theory of concurrent multitasking. *Psychological review*, 115(1):101.

Salvucci, D. D., Taatgen, N. A., and Borst, J. P. (2009b). Toward a unified theory of the multitasking continuum: From concurrent performance to task switching, interruption, and resumption. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1819–1828. ACM.

Sangin, M., Molinari, G., Nüssli, M.-A., and Dillenbourg, P. (2011). Facilitating peer knowledge modeling: Effects of a knowledge awareness tool on collaborative learning outcomes and processes. *Computers in Human Behavior*, 27(3):1059–1067.

Saracevic, T. (2007). Relevance: A review of the literature and a framework for thinking on the notion in information science. part ii: nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology*, 58(13):1915–1933.

Sasse, A., Johnson, C., et al. (1999). Coordinating the interruption of people in human-computer interaction. In *Human-computer interaction, INTERACT*, volume 99, page 295.

Satyanarayanan, M. (2001). Pervasive computing: vision and challenges. *Personal Communications, IEEE*, 8(4):10–17.

Schilit, B., Adams, N., and Want, R. (1994). Context-aware computing applications. In *Mobile Computing Systems and Applications, 1994. WMCSA 1994. First Workshop on*, pages 85–90. IEEE.

Schraw, G. and Moshman, D. (1995). Metacognitive theories. *Educational psychology review*, 7(4):351–371.

Schunk, D. H., Pintrich, P. R., and Meece, J. L. (2008). Motivation in education: Theory, research, and applications.

Schwartz, S. H. and Tessler, R. C. (1972). A test of a model for reducing measured attitude-behavior discrepancies. *Journal of Personality and Social Psychology*, 24(2):225.

Seaton, D. T., Bergner, Y., Chuang, I., Mitros, P., and Pritchard, D. E. (2014). Who does what in a massive open online course? *Communications of the ACM*, 57(4):58–65.

Shaffer, L. (1975). Multiple attention in continuous verbal tasks.

Shah, P. K., Perez-Iratxeta, C., Bork, P., and Andrade, M. A. (2003). Information extraction from full text scientific articles: Where are the keywords? *BMC bioinformatics*, 4(1):20.

Shenoy, P. J. and Vin, H. M. (1995). Efficient support for scan operations in video servers. In *Proceedings of the third ACM international conference on Multimedia*, pages 131–140. ACM.

Shiffrin, R. M. and Schneider, W. (1977). Controlled and automatic human information processing: Ii. perceptual learning, automatic attending and a general theory. *Psychological review*, 84(2):127.

Shneiderman, B., Byrd, D., and Croft, W. B. (1997). Clarifying search: a user-interface framework for text searches. *D-lib magazine*, 3(1):18–20.

Siemens, G. (2006). *Knowing knowledge*. Lulu.com.

Siemens, G. (2012). Moocs are really a platform. *ELearnspace,[online] Available at: http://www.elearnspace. org/blog/2012/07/25/moocs-are-really-a-platform.*

Simons, D. J. and Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception-London*, 28(9):1059–1074.

Simons, D. J. and Jensen, M. S. (2009). The effects of individual differences and task difficulty on inattentional blindness. *Psychonomic Bulletin & Review*, 16(2):398–403.

Sinha, T., Jermann, P., Li, N., and Dillenbourg, P. (2014a). Your click decides your fate: Inferring information processing and attrition behavior from mooc video clickstream interactions. In *2014 Empirical Methods in Natural Language Processing Workshop on Modeling Large Scale Social Interaction in Massively Open Online Courses*, number EPFL-TALK-202095.

Sinha, T., Li, N., Jermann, P., and Dillenbourg, P. (2014b). Capturing" attrition intensifying" structural traits from didactic interaction sequences of mooc learners. *arXiv preprint arXiv:1409.5887*.

Sipusic, M. J., Pannoni, R. L., Smith, R. B., Dutra, J., Gibbons, J. F., and Sutherland, W. R. (1999). Virtual collaborative learning: a comparison between face-to-face tutored video instruction (tvi) and distributed tutored video instruction (dtvi).

Skaalvik, E. M. (1997). Self-enhancing and self-defeating ego orientation: Relations with task and avoidance orientation, achievement, self-perceptions, and anxiety. *Journal of Educational Psychology*, 89(1):71.

Smith, G. G., Sorensen, C., Gump, A., Heindel, A. J., Caris, M., and Martinez, C. D. (2011). Overcoming student resistance to group work: Online versus face-to-face. *The Internet and Higher Education*, 14(2):121–128.

Smith, R. B., Sipusic, M. J., and Pannoni, R. L. (1999). Experiments comparing face-to-face with virtual collaborative learning. In *Proceedings of the 1999 conference on Computer support for collaborative learning*, page 68. International Society of the Learning Sciences.

Sohn, T., Li, K. A., Griswold, W. G., and Hollan, J. D. (2008). A diary study of mobile information needs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 433–442, New York, NY, USA. ACM.

Sonnentag, S. and Zijlstra, F. R. (2006). Job characteristics and off-job activities as predictors of need for recovery, well-being, and fatigue. *Journal of Applied Psychology*, 91(2):330.

Speier, C., Vessey, I., and Valacich, J. S. (2003). The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. *Decision Sciences*, 34(4):771–797.

Spelke, E., Hirst, W., and Neisser, U. (1976). Skills of divided attention. *Cognition*, 4(3):215–230.

Spindler, M. and Dachselt, R. (2009). Paperlens: advanced magic lens interaction above the tabletop. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, page 7. ACM.

Stasko, J., McColgin, D., Miller, T., Plaue, C., and Pousman, Z. (2005). Evaluating the infocanvas peripheral awareness system: A longitudinal. *Situ Study*.

Stone, H. R. (1990). Economic development and technology transfer: Implications for video-based distance education. *Contemporary issues in American distance education*, pages 231–242.

Streitz, N. A., Röcker, C., Prante, T., Stenzel, R., and van Alphen, D. (2003). Situated interaction with ambient information: Facilitating awareness and communication in ubiquitous work environments. In *Tenth International Conference on Human-Computer Interaction (HCI International 2003)*. Citeseer.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18(6):643.

Sturm, J. and Terken, J. (2009). Relational cockpit. In *Computers in the Human Interaction Loop*, pages 257–270. Springer.

Suchman, L. and Reconfigurations, H.-M. (1986). Plans and situated actions. *New York, Cambridge University*.

Sullivan, L. (1976). Selective attention and secondary message analysis: A reconsideration of broadbent's filter model of selective attention. *The Quarterly Journal of Experimental Psychology*, 28(2):167–178.

Sun, X., Sharples, S., and Makri, S. (2011). A user-centred mobile diary study approach to understanding serendipity in information research. *Information Research-An International Electronic Journal*, 16(3).

Sutcliffe, A. and Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with computers*, 10(3):321–351.

Syeda-Mahmood, T. and Ponceleon, D. (2001). Learning video browsing behavior and its application in the generation of video previews. In *Proceedings of the ninth ACM international conference on Multimedia*, pages 119–128. ACM.

Tang, K. (1993). Spontaneous collaborative learning: A new dimension in student learning experience? *Higher Education Research and Development*, 12(2):115–130.

Tate, T. and Russell-Rose, T. (2012). The information needs of mobile searchers: A framework. In *Proceeding of Searching4Fun Workshop*.

Taylor, D. W., Berry, P. C., and Block, C. H. (1958). Does group participation when using brainstorming facilitate or inhibit creative thinking? *Administrative Science Quarterly*, pages 23–47.

Thayer-Bacon, B. J. (2000). *Transforming critical thinking: Thinking constructively*. Teachers College Press.

Thudt, A., Hinrichs, U., and Carpendale, S. (2012). The bohemian bookshelf: supporting serendipitous book discoveries through information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1461–1470. ACM.

Tipper, S. P. and Cranston, M. (1985). Selective attention and priming: Inhibitory and facilitatory effects of ignored primes. *The Quarterly Journal of Experimental Psychology*, 37(4):591–611.

Toms, E. G. (2000). Serendipitous information retrieval. In *DELOS Workshop: Information Seeking, Searching and Querying in Digital Libraries*. Zurich.

Trafton, J. G., Altmann, E. M., Brock, D. P., and Mintz, F. E. (2003). Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human-Computer Studies*, 58(5):583–603.

Treisman, A. (1988). Features and objects: The fourteenth bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology*, 40(2):201–237.

Treisman, A. and Davies, A. (1973). Divided attention between eye and ear. *Attention and Perjormance IV, S. KORNBLUM (Editor). Academic Press, New York.*

Treisman, A. M. (1964). Selective attention in man. *British medical bulletin*.

Treisman, A. M. and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, 12(1):97–136.

Tucker, B. (2012). The flipped classroom: Online instruction at home frees class time for learning. *Education Next*, 12(1):82.

Uden, L. and Willis, N. (2001). Designing user interfaces using activity theory. In *System Sciences, 2001. Proceedings of the 34th Annual Hawaii International Conference on*, pages 11–pp. IEEE.

Underwood, G. (1993). *The psychology of attention*, volume 2. NYU Press.

Van Andel, P. (1994). Anatomy of the unsought finding. serendipity: Orgin, history, domains, traditions, appearances, patterns and programmability. *The British Journal for the Philosophy of Science*, 45(2):631–648.

Van den Bussche, E., Van den Noortgate, W., and Reynvoet, B. (2009). Mechanisms of masked priming: a meta-analysis. *Psychological bulletin*, 135(3):452.

Verma, H. (2015). *Latent Social Information in Group's Interactions with a Shared Workspace*. PhD thesis, IC, Lausanne.

Visser, T., Vastenburg, M., and Keyson, D. (2010). Snowglobe: the development of a prototype awareness system for longitudinal field studies. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, pages 426–429. ACM.

Vogel, D. and Balakrishnan, R. (2004). Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 137–146. ACM.

Vygotsky, L. S. (1978). Mind and society: The development of higher mental processes.

Wacker, R. R. and Roberto, K. A. (2008). *Patterns of Service Use and Theories of Help-Seeking Behavior*, pages 35–51. SAGE Publications, Inc., 0 edition.

Walster, E., Berscheid, E., and Walster, G. W. (1973). New directions in equity research. *Journal of personality and social psychology*, 25(2):151.

Wang, H.-C., Cosley, D., and Fussell, S. R. (2010). Idea expander: Supporting group brainstorming with conversationally triggered visual thinking stimuli. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, pages 103–106. ACM.

Weiser, M. (1991). The computer for the 21st century. *Scientific american*, 265(3):94–104.

Weiser, M. and Brown, J. S. (1996). Designing calm technology. *PowerGrid Journal*, 1(1):75–85.

Weiser, M. and Brown, J. S. (1997). The coming age of calm technology. In *Beyond calculation*, pages 75–85. Springer.

Weisz, J. D., Kiesler, S., Zhang, H., Ren, Y., Kraut, R. E., and Konstan, J. A. (2007). Watching together: integrating text chat with video. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 877–886. ACM.

Welford, A. T. (1952). The ‘psychological refractory period’and the timing of high-speed performance—a review and a theory. *British Journal of Psychology. General Section*, 43(1):2–19.

Wickens, C. (1984). Processing resources in attention. i parasuraman, r. & davies, dr (eds): Varieties of attention.

Wickens, C. D. (1980). The structure of attentional resources. *Attention and performance VIII*, 8.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3):449–455.

Wickens, C. D. and McCarley, J. S. (2007). *Applied attention theory*. CRC press.

Wild, P. J., McMahon, C., Darlington, M. J., Liu, S., and Culley, S. (2010). A diary study of information needs and document usage in the engineering domain. *Design Studies*, 31(1):46–73.

Wilson, M. L., Alhodaithi, B., and Hurst, M. (2012). A diary study of information needs produced in casual-leisure reading situations.

Wilson, T. D. (1981). On user studies and information needs. *Journal of documentation*, 37(1):3–15.

Wilson, T. D. (1997). Information behaviour: an interdisciplinary perspective. *Information processing & management*, 33(4):551–572.

Wilson, T. D. (1999). Models in information behaviour research. *Journal of documentation*, 55(3):249–270.

Wilson, T. D. (2000). Human information behavior. *Informing science*, 3(2):49–56.

Wilson, T. D. (2005). A re-examination of information seeking behaviour in the context of activity theory. *Information Research*, 11(4):1.

Yuan, L. and Powell, S. (2013). Moocs and open education: Implications for higher education. *JISC cetis*, pages 1–21.

Zelkha, E., Epstein, B., Birrell, S., and Dodsworth, C. (1998). From devices to ambient intelligence. In *Digital living room conference*, volume 6.

Zijlstra, F. R., Roe, R. A., Leonora, A. B., and Krediet, I. (1999). Temporal factors in mental work: Effects of interrupted activities. *Journal of Occupational and Organizational Psychology*, 72(2):163–185.

Zipf, G. K. (1949). Human behavior and the principle of least effort.

Zufferey, G., Jermann, P., and Dillenbourg, P. (2008). A tabletop learning environment for logistics assistants: activating teachers. In *Proceedings of the Third IASTED International Conference on Human Computer Interaction*, pages 37–42. ACTA Press.

Zufferey, G., Jermann, P., Lucchi, A., and Dillenbourg, P. (2009). Tinkersheets: using paper forms to control and visualize tangible simulations. In *Proceedings of the 3rd international Conference on Tangible and Embedded interaction*, pages 377–384. ACM.

NAN LI

HCI RESEARCHER

Date of Birth : 21.03.1984

Nationality: Chinese

Email : nan.li@epfl.ch

Phone: (+41) 789613371

Address: Ch. de la Chiésaz 1, Ecublens, Switzerland

PROFILE

I am a motivated and enthusiastic person with interdisciplinary skills, including interaction design, computer science, statistics and psychology. I've got experiences working in the field of HCI for several years and I am familiar with various qualitative and quantitative methods to conduct user research. My objectives include but are not limited to a career in user experience, technology enhanced learning or analytics.

EDUCATION

PhD in Computer Science (*majored in HCI*)

06.2010 - 11.2015

Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
Advised by Prof. Pierre Dillenbourg

MSc in Computer Science (IDEA League Research Program)

06.2009 - 12.2009

Eidgenössische Technische Hochschule Zürich (ETHZ), Switzerland
Thesis Co-Advisor Prof. Moira Norrie

MSc in Media Informatics (*majored in HCI*)

10.2006 - 12.2009

Rheinisch-Westfälische Technische Hochschule Aachen (RWTH), Germany
Thesis Advisor Prof. Wolfgang Prinz

BSc in Information & Computing Science

09.2002 - 07.2006

Beijing University of Posts and Telecommunications (BUPT), China
Thesis Advisor Prof. Yixian Yang

SKILLS AND EXPERTISE

User Research

- *Qualitative: questionnaire, interview, user observation and coding*
- *Quantitative: observational and experimental studies, data-driven investigation*

Data Analysis

- *Statistical methods for data exploration, correlation, regression, classification and clustering; Experiences with large data set*

Interaction Design

- *Storyboarding, paper prototyping, wire-framing, design patterns, usability testing*

Programming and Coding

- *Java, Python, R (Proficient); C#, C++ (Knowledgable); SQL, Objective-C, javascript (Capable)*

Language

- *Chinese (Native); English (Full professional proficiency); German (Limited working proficiency); French (Basic)*

EXPERIENCES

Research Assistant — 06.2010 - 11.2015

Computer-Human-Interaction in Learning and Instruction Lab, EPFL, Switzerland

- ***Modelled how MOOC students interact with lecture videos***
  - *Python, R and SQL*
  - *statistical methods especially additive models, PCA, high dimensional clustering etc.*
- ***Conducted observational study on how study group students watch and use MOOC videos***
  - *analysis with recorded speech coding, video interaction log*
- ***Explored the design space for contextual information scents***
  - *C++, C#, Java, Python, computer vision, Kinect, speech recognition, Objective-C and R*
  - *interview, questionnaire, interaction log analysis, video coding, experiment design*

Visiting Researcher — 06.2009 - 12.2009

Information Systems Lab, ETH Zurich, Switzerland

- ***Designed and implemented tools that bridge interaction between paper and interactive tabletop***
  - *Java, Flash and Anoto technologies*

Student Assistant (Hilfswissenschaftler) — 12.2008 - 02.2009

Mobile and Multimedia Processing Group, RWTH Aachen, Germany

- ***Developed computer vision app on the first generation of the iPhone to track moving objects***
  - *C and Objective-C, computer vision*

Student Assistant (Hilfswissenschaftler) — 04.2008 - 09.2008

Virtual Environment Lab, Fraunhofer IAIS, Sankt Augustin, Germany

- ***Built the user interface for tracking indoor robots for the EU project INT-MANUS***
  - *C++, Qt and Ubisense indoor tracking technologies*

Intern (Praktikum) — 11.2007 - 02.2008

Information in Context Lab, Fraunhofer FIT, Sankt Augustin, Germany

- ***Built an innovative remote media control application on the mobile phone***
  - *Java and Ubisense indoor tracking technologies*

Intern (Praktikum) — 04.2007 - 07.2007

User Interface Engineering Lab, Fraunhofer IAIS, Sankt Augustin, Germany

- ***Built an innovative location-based context-aware application on a mobile platform***
  - *Java and GPS tracking*

ONLINE EDUCATION CERTIFICATION (MOOCs)

- Human Computer Interaction: 100% achieved — Stanford University
- Data Analysis & Statistical Inference : 91.7% achieved (Distinction) — Duke University
- Getting and Cleaning Data : 100% achieved (Distinction) — Johns Hopkins University
- The Data Scientists' Toolbox : 100% achieved (Distinction) — Johns Hopkins University
- R Programming : 100% achieved (Distinction) — Johns Hopkins University

AWARDS & GRANTS

- IDEA league research grant. 2009
- Shared task competition winner at EMNLP MOOC workshop. 2014
- Best paper award at EC-TEL. 2015

PUBLICLICATIONS

Journal articles:

[1] ***Li, Nan***, Himanshu Verma, Afroditi Skevi, Guillaume Zufferey, Jan Blom, and Pierre Dillenbourg. "Watching MOOCs together: investigating co-located MOOC study groups." Distance Education 35, no. 2 (2014): 217-233.

[2] Blom, Jan, Himanshu Verma, ***Nan Li***, Afroditi Skevi, and Pierre Dillenbourg. "MOOCs are more social than you believe." eLearning Papers (2013).

Conference papers:

[3] ***Li, Nan,*** Łukasz Kidziński, Patrick Jermann, and Pierre Dillenbourg. "MOOC Video Interaction Patterns: What Do They Tell Us?." In Design for Teaching and Learning in a Networked World, pp. 197-210. ECTEL 2015.(**best paper award**)

[4] ***Li, Nan***, Lukasz Kidzinski, and Pierre Dillenbourg. "Augmenting Collaborative MOOC Video Viewing with Synchronized Textbook." In Proceedings of INTERACT 2015

[5] ***Li, Nan***, Lukasz Kidzinski, Patrick Jermann, and Pierre Dillenbourg. "How Do In-video Interactions Reflect Perceived Video Difficulty?." In Proceedings of the European MOOCs Stakeholder Summit 2015, pp. 112-121. PAU Education, 2015.

[6] ***Li, Nan,*** Himanshu Verma, Afroditi Skevi, Guillaume Zufferey, and Pierre Dillenbourg. "MOOC Learning in Spontaneous Study Groups: Does Synchronously Watching Videos Make a Difference?." In Proceedings of the European MOOC Stakeholder Summit 2014, pp. 88-94. PAU Education, 2014.

[7] Roman, Flaviu, Himanshu Verma, ***Nan Li***, Patrick Jermann, and Pierre Dillenbourg. "Mobile technology for awareness of time progression and its impact on meetings." In Collaboration Technologies and Systems (CTS), 2013 International Conference on, pp. 260-267. IEEE, 2013.

[8] ***Li, Nan***, and Pierre Dillenbourg. "Designing conversation-context recommendation display to support opportunistic search in meetings." In Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia, p. 12. ACM, 2012.

[9] Ispas, Adriana, ***Nan Li***, Moira C. Norrie, and Beat Signer. "Paper-digital meeting support and review." In Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 2010 6th International Conference on, pp. 1-10. IEEE, 2010.

Workshop papers:

[10] Sinha, Tanmay, Patrick Jermann, ***Nan Li***, and Pierre Dillenbourg. "Your click decides your fate: Inferring information processing and attrition behavior from mooc video clickstream interactions." In 2014 Empirical Methods in Natural Language Processing Workshop on Modeling Large Scale Social Interaction in Massively Open Online Courses, 2014.

[11] Sinha, Tanmay, ***Nan Li***, Patrick Jermann, and Pierre Dillenbourg. "Capturing "attrition intensifying" structural traits from didactic interaction sequences of MOOC learners." EMNLP 2014 (2014): 42. (**shared task competition winner**)

Posters:

[12] ***Li, Nan***, Frédéric Kaplan, Omar Mubin, and Pierre Dillenbourg. "Supporting opportunistic search in meetings with tangible tabletop." In CHI'12 Extended Abstracts on Human Factors in Computing Systems, pp. 2567-2572. ACM, 2012.

[13] ***Li, Nan*** and Pierre Dillenbourg. "Ambient Intelligence in Meetings:Can Conversation-sensing Ambient Display Support Opportunistic Search? ". International Symposium on Pervasive Displays, 2012. (non-archived)

Demos:

[14] ***Li, Nan*** and Pierre Dillenbourg. "BOOC: A MOOC Video Player with Synchronised Textbook." ACM International Conference on Interactive Experiences For Television and Online Video. 2012 . (non-archived)