Common Ground by Artefacts
– Everyday Collaborative Manipulation

Peter Ling
Common Ground through Artefacts – everyday collaborative manipulation

Submitted by Peter Ling to the University of Skövde as a final year project towards the degree of B.Sc. in the School of Humanities and Informatics. The project has been supervised by Tarja Susi.

6/6 2010

I hereby certify that all material in this final year project, which is not my own work has been identified and that no work is included, for which a degree has already been conferred on me.

Signed:
Preface

This thesis represents the culmination of a three-year Cognitive Sciences program, and encompasses 30 ECTS. The inspiration for this thesis was the shortcomings of classic classroom interactions and an intense wish that these shortcomings do not propagate into the next generation of computer supported teaching environments.

I would like to use this space to direct special thanks to my advisors Tarja Susi and Henrik Svensson, who both were a great help in refining the issue addressed in this thesis and whose encouragement has been invaluable throughout the process.

Also, thank you great master Halcolm.
Common ground through artefacts – everyday collaborative manipulation

Peter Ling

Abstract

This thesis explores how cognitive artefacts contribute to the process of reaching common ground within collaborative groups through a study of both Clark’s theory of Common Ground and of how artefacts are understood to be used, both by individuals and as seen in distributed cognition. This was accompanied by an ethnomethodologically inspired study in a natural setting of a kitchen to observe how artefacts are used when negotiating common ground. After the study, participants were interviewed in order to establish whether common ground was successfully established and to look for consistency between observations from the study and how the participants motivate their actions. The study which was video recorded was analysed in order to find patterns, of which four distinct kinds were identified; these categories were then related to facts established in the study of Common Ground and artefacts.

Keywords: Common Ground, Artefacts, Collaboration, Communication, Qualitative study
# Table of contents

1 Introduction........................................................................................................... 1

1.1 Overview................................................................................................................. 3

2 Common Ground........................................................................................................... 4

2.1 Origin of Common Ground...................................................................................... 4

2.1.1 Basic Tenets of Common Ground...................................................................... 4

2.1.2 Levels of Interaction......................................................................................... 7

2.2 Critique of Common Ground.................................................................................. 10

2.3 Scientific Applications............................................................................................ 12

2.4 Summary of Common Ground................................................................................ 12

3 Artefacts...................................................................................................................... 14

3.1 Cognitive Artefacts................................................................................................. 14

3.2 Social Usage of Artefacts....................................................................................... 17

3.3 Artefacts within Distributed Cognition................................................................... 21

3.4 Unsolved Issues with Artefacts.............................................................................. 23

3.5 Summary of Artefacts............................................................................................ 23

4 Issue to be Addressed.................................................................................................. 25

4.1 Areas of Interest..................................................................................................... 26

4.2 Scope of the Present Thesis.................................................................................... 26

5 Methodology, Methods and Procedure...................................................................... 27

5.1 The Scientific Method............................................................................................. 27

5.2 Setting and Social Context..................................................................................... 28

5.3 Structured Interviews............................................................................................. 29

5.4 Procedure and Overview....................................................................................... 29

6 Analysis and Results.................................................................................................... 32

6.1 Analytic Overview.................................................................................................. 32

6.2 Artefact Usage........................................................................................................ 32

6.2.1 Structures for Public use.................................................................................. 32

6.2.2 Structures for Private use................................................................................ 35

6.2.3 Indirect Usage................................................................................................... 37

6.2.4 Direct Usage...................................................................................................... 38

6.3 Summary of Results............................................................................................... 39
7 Discussion ................................................................. 41
  7.1 Trustworthiness ......................................................... 42
  7.2 Transferability .......................................................... 42
  7.3 Dependability ........................................................... 43
  7.4 Confirmability ........................................................... 43
  7.5 Future Considerations .................................................. 43

References

Appendix
Short Glossary for the Readers Convenience

Common Ground: Clark’s theory of language use, not to be confused with common ground, which is a generic term used to describe shared understanding.

Common ground#: notation in this thesis to highlight shared understanding understood through the mechanisms of Common Ground.

Icon: In Clarks “Language use” theory, an icon is a sign that demonstrates a thing. (p. 160, Clark, 1996)

Index: In Clarks “Language use” theory, index indicates a thing. (p. 160, Clark, 1996)

Joint project: an action inviting participants to partake in a project with a common goal.

Joint activity: a sequence of mutual joint actions between participants who are equals in order to reach a certain goal.

Joint action: any action, verbal or non-verbal, whose meaning is both co-constructed and shared between participants.

Signal: an action by a participant affirming the grounding, co-constructed if correctly identified and treated as such by other participants.

Symbol: In the context of Clark’s theory of “Language use” a symbol is a sign whose representative character is associated with an object by formal rules, such as other signs. Symbols describe things (p. 160, Clark, 1996). Examples range from the non-linguistic “toward stretched open hand”, which is an iconic symbol of friendship, to the linguistic “Et tu, Brute?” which has become a verbalised icon of betrayal even among people not speaking Latin. That being said, Clark (1996) is of the opinion that most symbols in joint actions are composite of both verbal and non-verbal signals, such as a shake of the head combined with the verbal “no”.

Symptom: lack of signals that would give evidence of grounding.
1 Introduction

Today’s interactions between humans are increasingly taking place between geographically separated individuals and more often than not, this interaction does not allow for the traditional face-to-face interaction that Clark asserts that we are designed for in his 1996 book “Using language”. To overcome the limitations that arise from less salient forms of communication we are dependent on artefacts, e.g., tools, in order to reach a shared understanding. While it is said (Ingold, 1993) that all tools are not artefacts and all artefacts are not tools, there will be no such functional distinction in this thesis. In this thesis the definition of artefact will be based on Norman’s information process centred definition (Norman, 1991), not because of any affection toward the computational model of human cognition but rather because of its ease of integration with a more holistic view of human activity as part of a distributed system. However, the intent is to narrow this rather broad definition down to physical instruments which wholly or partly represent the artefacts inherent information by the shape of the artefact. Incidentally, this does not call for any differentiation between tools and other artefacts as all cognitive tools will invariably affect how we manage information represented by cognitive artefacts.

Shared understanding is something which can be hard to achieve even under normal circumstances (Monk, 2003), and more so when it needs to take place outside of a commonly understood context (see Koschmann & LeBaron (2003) for a detailed example). Shared understanding is often understood as a shared state of beliefs reached through conversation, in this thesis this process will be referred to in terms of Herbert Clark’s theory of Common Ground (Clark, 1996 & Monk, 2003). Yet, conversation may not be the only means for the achievement of CG as artefacts have proven significant for shared cognitive processes (Hutchins, 2006). However, Clark’s framework does not make any mention of the usage of artefacts in establishing Common Ground (Monk, 2003) other than as means of communication, such as manipulation of an artefact in response to a joint action, in a conversation (Clark & Krych, 2003). According to Clark, artefact manipulation in itself does not differ from an exchange of words, only it is faster and more efficient (Clark & Krych, 2003), as there is less room for confusion when communication is embodied. Given that there were contemporary researchers such as Parker and Millbrath (Ingold, 1993) who established a connection between shared behaviour such as culture and language through artefacts (and that, language itself can be viewed as an artefact, but that’s outside the scope of this thesis), it is puzzling as to why Clark glosses over the role artefacts play when negotiating the common ground$^1$ during the grounding process.

A language community generates, transforms, selects, stores and transmits plans and procedures. In other words, a language community functions like a computer bulletin board interactively expanding and updating the available database. Language thereby removes planning from the pre-human domain of individual actions which are coordinated with or modelled after those of other animals through social facilitation and/or local

---

$^1$ For use of terminology in this thesis, see Section 1.1
and stimulus enhancement (Galef, 1988), and into the cultural domain (p.
322, Gibson & Ingold, 1993).

Although Ingold (1993) does not talk about Common Ground, or artefacts, the
given description of the cultural domain and its workings do have a striking
resemblance to Common Ground as “where a meeting of minds and
transferral of structures takes place”. It is also interesting to note the
similarities between the definition of a “language community” and Norman’s
definition of an artefact: “an artificial device designed to maintain, display, or
operate upon information in order to serve a representational function” (p. 17,
Norman, 1991). This seems to support the notion that language and other
cognitive artefacts are stored, transferred etc., in the cultural domain/Common
Ground. Additionally, researchers such as Ingold (1993) have suggested that
the purpose of tools, such as words (which in this thesis would be considered
cognitive artefacts), is to mediate between a task and the world. This is very
much in alignment with Norman’s (1991) view of artefacts. Furthermore,
Ingold (1993) advocates further study of the whole system of interaction
between users in their context in order to understand both artefact usage and
language (Gibson & Ingold, 1993) which is consistent with other cognitive
approaches like Distributed Cognition. As the purpose of Common Ground is
exactly to model and make formalised predictions in order to shape
communication that support humans in activity through the use of joint
actions. As it already describes most interactions, except those mediated
through artefacts, there is a need for a detailed inquiry of how artefacts fit into
the formalisation of a collaborative activity (Monk, 2003).

Despite increased reliance on technology supporting people working at
geographically distinct locations, the lion-share of work has focused on how
Common Ground is achieved in distributed domains such as; CMC
(Computer Mediated Communication), CSCL (Computer Supported
Collaborative Learning) and CSCW (Computer Supported Collaborative
Work). Those are areas where highly developed technologies mediate a
problem solving process, and less on how participants in everyday social
settings use artefacts in reaching Common Ground. Most research seems to
hurry past that particular point, satisfied with the conclusion that artefacts are
used. What is known is how we use artefacts in problem solving when we are
alone, e.g., Mordechai and Tzipppora (2006) or co-located, e.g., Arias et al
(2000), and how distance affects collaboration e.g., Bradner and Mark (2002)
and Fischer (2004). Other researchers, such as Hutchins (1999) and Perry
(1999) have, by using cognitive ethnography, shown that; yes, artefacts are a
large component in the interaction between humans, especially artefacts
distributing knowledge through time or locations, for example
language/culture. This allows for (among other things) coordination,
transferral and sharing of knowledge. Yes, the transformation of seemingly
complex cognitive tasks into simple information processing tasks – for
instance the act of arranging a physical puzzle instead of a mental image do
have significant advantages. However, the issue regarding the usage of
artefacts to reach a shared understanding in a collaborative task through a
Common Ground perspective seems to be largely unexplored.

This thesis will focus on how cognitive artefacts ease the process of achieving
Common Ground within small groups. Therefore, two areas of human
cognition are subject to study in this thesis. Those two are Herbert Clark’s
theory of Common Ground and human use of cognitive artefacts (and how we use those artefacts) to achieve a shared understanding. This is warranted as Common Ground is a relatively popular framework used to explain human interactions within areas such as CSCW as it provides a predictive model for how conversations are conducted, and what kind of information is sought and transmitted within them (Clark & Krych, 2003). The process of grounding and the usage of Common Ground have shown itself to have significant similarities to knowledge building within areas such as distributed cognition. Further studies of artefact usage could elaborate on this, which could add knowledge of how to determine what kind of information an artefact and how the artefact encodes and how it is used to transmit said knowledge (Hutchins, 2006) within the frame of a conversation - thereby facilitating the grounding process. Distributed cognition shown itself to be capable of describing relations between artefacts and humans, and it may be able to provide a structure to the fragmented usage of artefacts in conversations.

The aim of this thesis is to bring together the understandings of cognitive artefact usage, to ease the process of reaching common ground, by applying Common Ground theory to ethno-methodologically collected material in the form of video analysis. The intention is to take a closer look at two separate but possibly complementary theories, and to identify in what ways they complement each other and further our understanding of the process of achieving common ground.

1.1 Overview

Chapter two describes Common Ground in Clark’s theory of language use, how it is achieved and its relevance to the next chapter. The third chapter describes both artefacts and cognitive artefacts. Then it goes on to explore how people use cognitive artefacts in activities and why artefacts in activity are relevant when negotiating Common Ground in collaborative processes. Chapter four is a detailed description and motivation of the inquiry “How do cognitive artefacts contribute to the process of reaching common ground within collaborative groups?” which is addressed by this thesis. Chapter five details the methodology, the methods and the procedure of the ethno-methodologically inspired study used in this thesis. Chapter six elaborates on the on the analysis of the study, the structured interviews and their results. The seventh, and final, chapter contains the discussion and suggestions for future research, which marks the end of this thesis.
2 Common Ground

Common ground offers a solid and well-elaborated framework for how co-situated persons create a shared understanding, and to some extent how peripheral participants affect the process (Monk, 2003) and this chapter will describe its origin and content.

2.1 Origin of Common Ground

Herbert Clark’s first cohesive presentation of his theory of Common Ground was in his 1996 book “Using language”, and it has been improved ever since. The basic idea is that humans cooperatively reach a common understanding through conversation, where the added social model is supposed to add the “common understanding” earlier cognitive models of communication had failed to account for. Clark assumes (Monk, 2003) conversations to have their roots in face-to-face interaction between two individuals. Conversation is, by Clark, assumed to be an implicit iterative negotiation, which determines the understanding of the communicated representations exchanged by the participants, which results in an integration of the representation as understood by the recipient. This is something Hutchins mirrors in the following quote:

…it is apparent that verbal behavior is richly multimodal in the sense that it integrates many layers of structure, each consisting of a distinct set of patterns, and all coordinated with one another by a very complex set of relations’ (p. 1, Hutchins, 2006).

Monk (2003) describes grounding as a combination of previous common ground and the assumptions of knowledge the conversationalists have made about each other; the grounding process is where participants, as previously mentioned, exchange and integrate representations, which results in Clarks (1996) Common Ground. In grounding every utterance is both an attempt to assure the listener that indeed both explicit and implicit implications of utterances are understood, and an invitation to engage in further joint actions, now grounded in a shared understanding also known as Common Ground (Monk, 2003). As both “shared understanding” and “common ground” in the literature commonly describe any situation in which an idea has propagated past an individual, common ground# is used throughout this thesis in order to clarify that the subject at hand is specifically a shared understanding, as understood through Clark’s theory of Common Ground. An example of a joint action initiated through this invitation would be a correction prompted by a miss-match between the intended message from original speaker, and the understanding the original listener signalled in his or her response through either signals or symptoms.

2.1.1 Basic Tenets of Common Ground

It is common to conceptualise language in communication as utterances produced by a sender and decoded by a receiver, Clark, however, is of the opinion that something more is required. That “something” is how language works as a social structure when participants use language in concert. One major reason social structure is needed is that communication is often fragmented and imprecise, which a traditional cognitivist information-
transfer model is not only ill equipped to describe, it also lacks an ability of transfer of e.g., intent or history (Monk, 2003). Clark is of the opinion that rules governing that kind of knowledge has its basis in experiences, accordingly spoken language is efficient because, not despite, it communicates only what is relevant to the participants in their current activity. In “Using Language” Clark (1996) states that his focus in language is on the use of signals, and identify three basic methods of signal use: “describing-as”, “indicating”, and “demonstrating”.

However, he starts with the assertion that language use and joint actions are inseparable and that we cannot understand one of them without the other. So, what are joint actions? Joint actions are, according to Clark (1996), essentially activities, ongoing processes, which can be either simultaneous or intermittent. As such they need something to provide coordination; one example (Monk, 2003) is conventions which are part of the common ground shared between the participants. Or, as phrased by Clark (Monk, 2003), “The key characteristics of a joint action are that both people involved intend to do their part and they believe that the joint action includes their part and the other’s” (p. 275, Monk, 2003). Both Clark and his predecessors (Clark 1996) admit that activities can appear arbitrary, as they consist of several varying dimensions; however, this is supposedly not true. They are supposed to follow “a few basic principles, in particular a rational organization around a dominant goal” (p. 31, Clark, 1996) in order to create a structure that can identify these principles. In order to facilitate this process, Clark argues (Monk, 2003) that we (as participants), within moments, are able to make assumptions about the extent of the common ground among the participants. While this accommodates efficiency in face-to-face interactions, it does assume that a speaker talks with the intention of facilitating understanding, it is here ambiguities arise when the Common Ground process fails due to signal failure (Monk, 2003). Clark (1996) also reasons that to understand common ground, there is a need to identify all participants and what roles they fill. Additionally common ground, in relation to an activity, expands through the cumulative nature of joint activities (Clark, 1996).

According to Clark (1996), at the initiation of the joint action, people ratify nearby as participants, although persons can, with some negotiation, leave and join during the activity. The leaving and joining of activities relates to boundaries, more about that later. Joint activities cause participants to perform an activity role, which shape both what the individual participant do and the understanding of what they are supposed to do, which affects the interpretations of their actions. Participants in joint activities are perform actions with a purpose, as either procedural goals, which are goals to accomplish in order to reach the overarching goal, or domain goals, which are what they want to achieve. Aside from procedural and domain goals, participants may also have other kind of goals; interpersonal goals, which relates to our self-image and private agendas, which are related to why we strive to achieve the domain goal. Domain goals contrast with public goals, which are openly established and are recognisable by all participants in the joint activity (Clark, 1996). This means that in joint activities participants establish activity goals, grounded in the assigned activity roles, in order to determine the division of labour; however, how well this works in practice is subject to how well the public goal meshes with the participant’s domain
goals. This means that when, e.g., deceiving others, not only are the public goals in the joint activity incrementally increased by every joint action, some of the joint actions are also/or meant to advance private goals. Joint activities themselves are realised through sequences of actions, some of which are joint actions, creating joint activity through a hierarchy of joint actions, some joint actions are what Clark call joint projects. Joint projects are suggestions of concerted action, which participants signify by questions and other signals seeking a response (Clark, 1996). The construction of this hierarchy is in many aspects similar to grounding, the most important similarity probably being that all participants need to believe that they are participating in the same activity as the other person. To tell joint activities apart they all have boundaries (Clark, 1996), which identify a jointly constructed entry and exit for any successful joint action. With boundaries established, a joint action may expand, contract according to the number of participants and their relative participation, or it can divide into two or more joint actions, which Clark name dynamics (1996).

Since Clark’s (1996) publication of “Using Language”, Clark have elaborated the theory and made further additions in various publications, either to clarify issues that have garnered his framework criticisms or natural development as he has done further research, notably further elaboration of the concepts of levels. One way to describe Joint actions is by using primarily four levels of interaction. To describe a joint action, which is one step in the joint activity of a conversation, there is a need for an initiator and one or more attendees. This text describes initiator and attendee as such in to signify:

1) Joint actions are co-constructed signals, of which some can be linguistic (words, sentences, etc.) or non-linguistic (gestures, body language, etc.), some of which are symbols, and thus do not necessarily need to be vocalised and

2) A better differentiation between the roles individuals play and their actions performed in the capacity of their role. Joint actions completes through a joint project suggested by the initiator, which the attendee has to decide if it wants to engage in.

An example of a joint project used to invite potential participants to a joint activity is a simple greeting such as “Hello”. Potential participants who do not wish to participate simply ignore the inherent suggestion of a joint project in the “Hello” symbol, even if, technically, it is more polite to accept the joint project and engage in another joint action replying “Sorry, not now”, thus providing closure to the joint activity. By providing evidence that the content of the proposed project has been acknowledged, considered and replied to, in accordance with the (primarily) four levels of joint actions the closure is accomplished. However, do not interpret that as a closure in itself, as being a definitive end to interactions. Joint closures are an integral part of conversations where participants through joint actions establish earlier propositions as facts before moving on in the conversation (Clark, 1996), an example of joint closure would be “A: Bring me the book, please’ B: ‘You mean the blue book?’ A: ‘Yes”, in which A and B through a simple exchange both clarifies something which had been uncertain and provided evidence that B had understood A. That means that any further joint activities involving A and B can rely on that knowledge as a part of their shared ground. During a nested joint activity, of which a conversation is an example of, participants
will take turns initiating joint actions, as the subject at hand requires (Norman, 1991).

When breaking down joint actions to levels there are a few things to keep in mind: Symptoms are expressions of failure on the part of the attendee, at any level of the conversation, to act in response to the initiator. A suitable response at any level is to send signals. The purpose of the signals is to let the initiator know that common ground has been lost or that common ground this maintains the conversation and thus keeps it aligned. Signals in their turn are much like artefacts, only useful as long as the underlying intention is shared, hence all signals are co-constructed (e.g., a signal sent by the attendee not caught by, or shared by, the initiator would effectively be a symptom). When the signals are not only furthering participants understanding, but also understanding in the current context, anchoring is taking place, and signals turn into symbols (Clark 1996). This is why we monitor five areas, both those of ourselves and others, for signals and composite symbols: Voice, workspace, faces, body language, and shared environments (Clark & Krych, 2004). All participants monitor those five areas for signals and symbols at any level, described in the next section, while the participants engage in a joint action. As Clark (1996) asserts that most signals are composite signals (consisting of both linguistic and non-linguistic signals), any co-constructed signals are acceptable forms of answers.

2.1.2 Levels of Interaction

In an interaction, joint actions need to be successful, and for measuring their success, Clark (1999) has devised a system of levels that details what an action need to succeed with in order to become a joint action. This chapter will present four of these levels, detailing why and their use in the grounding process and some examples of how participants repair a conversation if there is a failure at any of the presented levels.

First level: The initiator needs the attendee to attend to him or her in order to establish common ground#. Here, a signal to establish the joint action would be any action or symbol of the attendee indicating that they did not pay attention, causing the initiator to attempt a restart. An oft-cited example is the greeting of a customer towards a cashier; the customer uttering the greeting acts as the initiator, the uttered greeting is an action that becomes subject to co-construction as the cashier responds with a signal. If the cashier at this point does not greet the costumer, the costumer will treat it as a symptom of failure. A symptom at this level would be no reaction at all from the attendee, causing the initiator once again to attempt to gain the attention of the intended attendee through a restart. A signal from the cashier can be a verbal or non-verbal action which indicates that the cashier, in his role as attendee, has acknowledged the initiators proposed joint action (transferral of goods in exchange for monetary compensation) (Clark & Krych, 2004).

---

2 Actually, there are further levels than the four described as Clark (1996) admits that extended joint projects may require further more to be accurately described. However, Clark has not described those further, and as they fall outside of the scope of this thesis there will be no further inquiry into what they may or may not describe.
Second level: The *initiator* must get the *attendee* to identify the presented words, phrases, and sentences. If grounding fails at this level, *initiator and attendee* will try to re-establish common ground through the *initiator* by attending to any signs or symptoms of the *attendee*. An example of a signal at this level would be the *attendee* asking the *initiator* for additional more familiar context in which they can relate any unfamiliar expressions, or vice versa (Clark & Krych, 2004). An example of a symptom at this level would be if the *attendee* attended to the wrong subject in the *initiators* narrative. To return to the greeting as an example, as the *initiator* says “Hello, are you Arthur?”, the *attendee* notices that something has been said but not what. By either signalling through an explicit action, such as “Sorry mate, didn’t quite catch that” or displaying a symptom such as a quizzical look, or if missing the latter part, assuming there was only a greeting, replying to the greeting alone. A failure in the grounding process at this level might lead the initiator to, erroneously, believe, e.g., that the *attendee* is Arthur.

Third level: The *initiator* must have the *attendee* to understand the intended meaning of the used words. At this level, the *attendee* would attempt further grounding through signals through the act of asking for further details and the *initiator* would add more details until they reach a shared understanding (Clark & Krych, 2004). The third level mostly addresses ambiguities; an example would be a priest referring to divine signs simply as “signs”, whereas the handyman deals strictly in mundane signs. Hypothetically, the priest might say, “I am in need of a sign”. At this point the handyman can signal by asking what kind of sign the priest is in need of, or the priest might take the awkward silence (the handyman awaiting to be told what kind of sign the priest need him to craft) as a symptom and elaborate with further details in order to establish common ground.

Fourth level: The *initiator* must get the *attendee* to consider answering, which the speaker usually achieves by finishing their sentence with something prompting a response (Clark & Krych, 2004). A common strategy is for the *initiator* to ask the *attendee* for his or her opinion, regardless if they are actually interested in it or not, to make sure the *attendee* is up to speed, which in itself is a form of alignment. Alignment is the process of comparing the participants common ground, when a discrepancy in the common ground between participants becomes public, it brings attention to the fact that there is a need for joint actions called repairs to facilitate grounding closing the gap (Clark, 1996). Using the four levels of interaction it is possible to analyse a failed conversation by using downward evidence and find at what point in the conversation grounding failed (Clark & Krych, 2004). An example of using downward evidence would be to look at a joint action, and look at it with the highest identifiable level as a starting point. Did the *attendee* answer the *initiator*? Ok, did the *attendee* understand the meaning behind the words used? Did the *attendee* identify the words used? “Oh they said, cat – not hat, no small wonder it seemed strange!” This identifies that the grounding process failed at level two, and armed with this new knowledge, the content of the common ground is open to new interpretations that may allow for easier grounding in further joint actions, or might give evidence that there is a need for further grounding.

While engaged in any level of the conversation, the participants of the dialogue are engaged in two kinds of monitoring: self-monitoring and
monitoring of others. The purpose of self-monitoring is to identify and correct our own mistakes while the monitoring of others is to allow us to either send signals or cause symptoms of lacking grounding (Clark & Krych, 2004). Signalling is, according to Clark (1996), although necessary for language to proceed, not linguistic; as he explains:

> The traditional assumption is that signals are “linguistic” objects – utterances of speech sounds, words, sentences – that work via their conventional meanings. \(\ldots\) …signals are either linguistic or nonlinguistic. Saying “yes” is linguistic, and nodding yes is non-linguistic. This wouldn’t be right either. It isn’t signals that are linguistic or nonlinguistic, but methods of signalling. Most signals are composite signals, the artful fusion of two or more methods of signalling (p. 155, Clark, 1996).

Furthermore, Clark (1996) explains that these signals are symbols, anchored in the situation by their indexical elements. However, signals are only meaningful if participants use them to co-construct joint-actions (Clark 1996).

Layers, which mean domains that build upon other domains, see use in expressing irony, jokes, etc. (Clark, 1996). For an example of a layered in-joke, “An orch walks into a bar with a parrot on his shoulder. The Bartender asks, ‘Where did you get that?’ And the parrot answers ‘Doruthar, they got them all over the place.’” The first layer is between the author of this text and its readers, on the subject of communication, enabling the pretense of the actors; everyone knows it is fictional. The second layer is the imaginary domain created when a person imagines an orch with a parrot walking into a bar, with the third layer being the interaction between the bartender and as it turns out, the parrot. To this, it is possible to add a fourth layer. This is common to so-called “in-jokes”. They require some knowledge about the domains involved that is not described in the public domain (in this case it is not just the fact that it is the parrot talking, and it is not only that it is a parody of another well known joke, but it depends on the relations between parrots and orchs in its original context). Without domain knowledge in-jokes fail, which means that it is also true that whenever participants in a conversation are unable to separate layers or layers depend on common ground, which is not there, conversation fails. Either because they plain “don’t get it” or because they think they “get it” but has only partially penetrated the layers (see “third level”, page 8).

Aside from that, layers are also an important factor for closures and repairs. Layers are important for closures as people are opportunistic about closures in their conversations. As it would be burdensome to explicitly declare “I’m done” people tend to opportunistically treat signals as evidence of closure, which means that layers are important in order to align the participants, which allows for joint-closures. However, according to Clark (1996) joint closures are impossible to reach without repairs. When and how to repair is decided, as mentioned earlier, by participants through monitoring of themselves and others, and it it is for this part where layers are important. By monitoring the layers as they speak, the participants can repair their speech opportunistically by initiating new joint actions as needed (Clark, 1996). Monitoring and alignment can also cause a restart; i.e., when someone starts saying something, changing their mind mid-sentence, stops, and starts over (Monk, 2003). An example would be “Earth is spher- actually it’s slightly oval due to gravity.” That would indicate a speaker who at first intends to utter the
simplified view that the planet Earth is spherical, but while monitoring other participants, the speaker realises that a higher degree of precision is called for.

As if that is not complicated enough, Clark also introduces the concept of tracks. There may be several tracks in a conversation, although Clark asserts that tracks beyond four are rare, and Clark himself only describes the first two (Clark, 1996). The first track is the best understood one, as it is describing the joint activity (being the conversation proper). The second, and subsequent tracks, is subject to a lesser degree of understanding. Clark defines the second track as one that describes the meta-actions of the conversation, such as um, ah, utterances that usually are categorised as noise or performance errors. Clark’s view, however, is that they are systematic and essential for language use.

2.2 Critique of Common Ground

In “Reconsidering Common Ground: Examining Clark's Contribution Theory in the OR”, Koschmann and LeBaron (2003) attributed a number of problems to Common Ground. Among those were: ambiguous responses that made it difficult to tell signals and symptoms apart, silences are not well explained (silences can be constructed as either acceptance or the withholding of acceptance), and a supposed lack of grounding of joint actions in spatial deixis. They also find Clarks terms of acceptance and presentation less than clear-cut when the environment at hand is changing rather than static, as assumed by Clark, causing some confusion among the two. According to Koschmann and LeBaron (2003) it is also a deficit of the theory that when a signal of acceptance is subject to rejection, it causes further explanation instead of allowing the conversation to move on. In light of the difficulties of differentiating rejections of acceptance and the withholding of acceptance, they ask for specified methods of withholding acceptance (Koschmann & LeBaron, 2003). They also question whether participants achieve Common Ground, as there is no way to access the mental states of the participants.

Clark (1996), in his more recent writings on common ground discusses it as a distributed form of mental representation. It is the superset of all of the sets of inferred understandings of each participant in a conversation independently maintained with respect to every other participant to the conversation. The superset is an abstraction, however, only observable by a god-like, omniscient outsider with privileged access to the participants’ representational inventories. Serious problems arise when one begins to treat common ground as if it were a singularity, a possession of the participants, a place, an arrived at state, in short, as a noun instead of as a verb (p. 13, Koschmann & LeBaron, 2003).

The 2004 article by Clark and Krych may well try to address the points raised by Koschmann and LeBaron (2003). Starting with silences, participants in a conversation probably ought to treat them as symptoms as Clark’s (Clark & Krych, 2004) definition of levels includes a requirement of downward evidence. This means that being silent is an option for the listener, but unless there is some exchange of joint-actions between the listener and the speaker, the responsibility of a loss in common ground at this point is entirely on the speaker. Using Clark’s “levels of interaction”, it is clear that the speaker need to gain attention and be acknowledged by the listener already at the first level (Clark & Krych, 2004).
This brings us to the “signals are ambiguous” part; if signals are ambiguous, it is possible to argue that the cause is the opportunistic nature of conversations, where participants, Clark predicts (Clark, 1996), will make assumptions about the extent of each other’s knowledge (Carroll, 2003). In addition, one would do well to keep in mind that Clark has defined Common Ground as “a proposition \( p \) is only common ground if: all the people conversing know \( p \); and they all know that they all know \( p \)” (p. 270, Carroll, 2003), which quite clearly gives that perfect common ground is seldom achieved. Occasional failures in a Common Ground process should not be surprising; rather they ought to be expected.

When failure in grounding happens through ambiguity, it is not a failing of the framework, but a failure of communication. That might be a bold statement to make, but considering the definitions of levels Clark gives from 2003 and onward, it should be safe to assume that at the third level the listener is obliged to ask for clarifications and the speaker is obliged to keep explaining until he or she is satisfied with the listener’s expressed understanding. In the same vein, a failure to heed a signal of acceptance does not have to be a failure of the framework as both level two and three require the listener to not only identify and understand (Clark & Krych, 2003), but also require him or her to signal the understanding in a form accepted by the speaker. Otherwise, joint actions would not consist of co-constructed signals. While speaking of joint actions, the notion of withholding acceptance ought to be subject to the interpretation of a failed joint action to achieve grounding, not as something that is open to a later revisit. The reason for this is that any grounding is subject to anchoring in the moment, and thus once lost a moment cannot be re-created, requiring anyone intending to send a signal acceptance to initiate a joint action, explicitly accomplishing this. To support this, Clark (1996) makes the assertion that it is possible to embed joint actions in other joint actions. To do anything otherwise would be, as Koschmann and LeBaron (2003) note, to invite the speaker to attribute the acceptance to the latest utterance. Additionally, people only participate as much as they need in order to satisfy their own goals (Carroll, 2003); if the listener’s lack of common ground does not have any negative impact on the speaker then the speaker has no incentive to initiate any repair behaviour.

The remaining problems identified are the multiple ways in which the grounding process in the operation room is hampered and the fact that they were able to determine that the least knowledgeable participant did not achieve Common Ground during the procedure. The participant did not achieve common ground despite the participant making all the sounds of someone understanding (Koschmann & LeBaron, 2003). Clark does seemingly not address these problems, along with the issue of separation of acceptance and presentation, even in later writings. What needs to be remembered in this case is that it was applied to human interaction mediated by an artefact, rather than the face to face interaction Common Ground was based upon, and that the theory is not equipped for deception (intentional or otherwise). In this scenario, it would probably have been more productive to try to extend Common Ground to include artefacts rather than condemning Common Ground to the annals of history.
2.3 Scientific Applications

Clark & Brennan, (1991) has established a use for applied Common Ground when they used it in a study to determine the cost of grounding through different mediums, and determine which constraints mediating technology needs to fulfill (Monk, 2003); due to the nature of their studies their findings are applicable to any scenario involving some form of Computer mediated Communication (CMC). Watts and Monk (1998) applied it to a CSCW scenario where they included peripheral participants, earlier interaction had been limited to doctor, patient and specialist.

Other concrete examples are implementations of collaborative augmented reality by Nilsson, Johansson and Jönsson (2009), groupware for students by Singley, Singh, Fairweather, Farrell and Swerling (2000), emergency management by Convertino, Mentis, Rosson, Carroll, Slavkovic and Ganoe (2008), naturalistic human-computer interaction such as studied by DeVault, Sagee and Traum (2009), human-robotic interaction as done by Stubbs, Hinds and Wettergreen (2006) and many other applications. This is only a selection of papers and presentations where the researches have worked with tangible objects and realistic environments, but they seem to be a fair representation of areas in which Clark’s theory of Common Ground is applicable.

2.4 Summary of Common Ground

To summarise this chapter; the basic tenets are that interactions are co-constructed joint-actions, most of which most of us never even consider, to construct meaningful interactions. This enables knowledge building through exchange and interaction of representations, which is the grounding process. This creates a shared understanding, also known as the Common Ground, which never is entirely available to any participant— as they individually each got their own interpretation which is only as good as it needs to be to accomplish the task. The relevance to the thesis lies in the understanding of what conversations entail, where artefacts would have to fit in and provide a context in which artefact usage can be related. Common Ground is an attempt by Clark to bridge cognitivist information processing theory and theories of social interaction, which he accomplishes by creating structure and formal rules for face-to-face communication.

Common Ground also formalises collaboration as joint actions (Monk, 2003) and has shown to be, not without problems, but as a tested framework, usable for predicting both needed exchanges in conversations and identifying where humans look for non-verbal communication to fill knowledge gaps. Which are something formal structures where meaningful interactions are deemed to be co-constructive accomplishes, in these actions have meaning because all participants ascribe roughly the same meaning to the action. Joint actions in turn, are what joint activities consists of, often nested, allowing Clark to use the same mechanics to explain the overarching interaction with the same mechanics as the singular joint action. The actions people take, jointly constructed or not, have a purpose, either a public or a private one and the commitment to the goal, plus the effectiveness of the grounding, are what decide the outcome of the activity. When it comes to grounding, the most important factor seems to be how well the participants follow the conversation on the different levels of interaction. During the grounding people are
assumed to show breakdowns in the interaction through signals (active responses) or symptoms (lack of responses), which is supposed to cause the speaker to, using downward evidence, go back to an earlier level until the problem has been identified and corrected.

The next chapter introduces artefacts, cognitive artefacts in particular, and their use in social settings. As a part of the social aspect, distributed cognition is briefly explored in order to illuminate how individual and social cognition exploits artefacts. The chapter rounds off with some identified problems with artefacts and a discussion detailing why a combination of distributed cognition and Common Ground could yield valuable results.
3 Artefacts

The earliest recorded artefacts are clay tokens and tables from Sumer, dated to 3300-3100 B.C. At first representations were quite direct, one to one mappings engravings on clay tablets, but as their usage increased, the representations grew increasingly abstract, allowing representations to map one token to larger numbers, such as ten or a hundred. These representations foreshadowed the writings later produced by scholars such as mathematicians and philosophers, whose writings are considered to have transformed the way people thought (Schmandt-Besserat, 1999). According to Norman (1991), it has been an area of interest for psychologists and social scientists since the early 20th century, with the majority of the research being conducted in Scandinavia, Germany and the former USSR. This research is the foundation for research areas such as situated action theory and distributed cognition, which both seek to explain social interactions between individuals and the world (artefacts) (Hutchins, 1999).

3.1 Cognitive Artefacts

Most of the explorations of cognitive artefacts have been from the perspective of individuals, and how they use them in social settings, as the following chapter will show. In order to understand the usage of cognitive artefacts in collaborative social settings, we must first understand how the individual interacts with them. Cognitive artefacts can be either building blocks of thought, ready to use in a mental structure at a moment’s notice, or they can be tools performing tasks such as computations, saving us the cognitive overhead while we perform other tasks (Norman, 1991). The only limits to the utility of cognitive artefacts are other processes in the cultural system that organises them and the exploitability of the results of their usage (Hutchins, 1999). Cognitive artefacts are commonly understood as objects that offer cognitive offloading through abstraction, which means that they allow for more complex cognitive efforts to be completed without increasing the complexity of the task. Norman defines a cognitive artefact as “an artificial device designed to maintain, display, or operate upon information in order to serve a representational function” (p. 17, Norman, 1991), in short: artefacts facilitate information processes. In this thesis cognitive artefacts will be viewed as physical instruments which carry both formal and informal knowledge, which are wholly or partly represented by the physical attributes of the artefact. The motivation for this is that artefacts tend to set the agenda for a conversation, and they are often at the centre of the discussion, even if they are only indirectly referenced (Streeck, 1996).

Common cognitive artefacts can be lists of different kinds, mnemonics, maps, or spreadsheets. However, it is worth noting that these items are not only artefacts by themselves, the co-joined arrangement of two or more items can also be a cognitive artefact (Hutchins, 1999). It is noted that basic value placed arithmetic operations display this; a structure is created by operations, the results are examined and the resulting structure is re-used by applying further operations to it (Hutchins, 1999). Kirsh is the author of the article “Intelligent use of Space”, where he argues that structures becomes cognitive artefacts, which refers to our ability to create structures through spatial placements of objects in our environment (Hutchins, 1999). To elaborate on
the concept of intelligent use of space; it is grounded in the assumption that as humans get bodies, they are occupying space which means they always face a direction which influences what they see and what they can reach – in short: what they may interact with. Therefore, Kirsh (1995) is of the opinion that spatial arrangement of items is not an afterthought or coincidence; it is something humans do intentionally and is an integral part of thinking, planning and behaviour. Kirsh (1995) defines three categories of spatial arrangements: those that simplify choice, those that simplify perception and those that simplify internal computation.

When we use space well we can often bring the time and memory demands of our tasks down to workable levels. We can increase the reliability of execution, and the number of jobs we can handle at once. The techniques we use are not always obvious, or universal. Some were taught to us, some naturally evolved as we improved our performance through practice, some are inevitable consequences of having the type of bodies, manipulators and sensors we have (p. 32, Kirsh, 1995).

The explanation to why this works so well for humans, according to Kirsh (1995), is in human psychophysics, through mechanisms that make it easier for humans to, e.g., estimate length rather than area or volume. This supports execution of tasks through the easy to track cues (visual or otherwise) which people add to the environment through various techniques. When discussing behaviour and techniques Kirsh uses three categories grounded in temporal distance. However, for the purpose of this thesis only short and medium term techniques are considered. The reason for that decision is that long term planning is at the level where an individual has complete control over the environment, which is seldom the case – many have to work in areas build and provided by others in parts of the world they come to live in without making an active choice to end up there, etc. On the other hand, medium term planning involves behaviours such as adapting the work place, bringing forth appropriate tools, etc., and short term planning involves the arrangement of objects, which are to be used in the workplace.

When observing, as Kirsh (1995) puts it, agents, in the short term, Kirsh noted that agents constantly re-arrange items to make it easy to: track the state of the task or realising the properties signalling what to do next, and to predict the outcome of the performed actions. All these are things that experts can leverage, as cues likely trigger both skill and rule based behaviours. However, it is important to note that Kirsh (1995) emphasises that most tasks consists of everyday actions that implies that over time anyone may achieve expertise. An example of behaviour, which Kirsh defines as explicit for experts, is the ability to prepare their work area. These preparations serve the above-mentioned perception of all things in the working area, constraints as to what do with items in it, and as reminders of what to do next. The result is that people make changes to the world in order to save themselves costly and potentially error-prone mental computations. That is not all, when talking about organisation and categorisation as a way of adding cues, Kirsh (1995) adds that these cues are not only available and usable to the person who creates them, but those cues are also available to others who might encounter them.

This means that cognitive artefacts do not have to be physical, as we are also able to use both mental structures and pattern-like behaviours by groups in
social settings as an object symbol (see: end of section 3.1). As the last example showcases, usage of artefacts does not need to be something premeditated, as it also includes opportunistic usage of natural structures (Hutchins, 1999). What information on a piece of paper does is to allow the usage of a different functional skill to complete the task involving the information (Hutchins, 1999). Norman actually explains it in somewhat plainer terms: "...artefacts do not actually change an individual’s capabilities; rather they change the nature of the task performed by the person" (p. 19, Norman, 1991). As previously mentioned, Norman (1991) defines artefacts as information processes; he also views human activity as information processes, which is, in a sense, rather handy when you want to explain human-artefact interaction as a system. Norman (1991) views the cognitive capabilities of a system consisting of: the human mind, a task and any artefacts, as one unit. This causes something that many mathematics teachers complain about; students come to rely on calculators which makes them skilled at using calculators but they become less skilled at performing the actual arithmetic operations. This is because the calculator, as a tool, displaces the functional skills required to form the arithmetic operations (Hutchins, 1999), causing the underutilised skills to deteriorate.

Through pre-existing skills and knowledge supporting a conceptual model, a user of artefacts reach an understanding of the artefacts uses by contextualising the conceptual model in an activity (Mordechai & Tzippora, 2006). Norman (1991) discusses three aspects of cognitive artefacts. The first aspect is two differing views of artefacts: the system view and the personal view. The second aspect is levels of directness and engagement. The third aspect is their representational properties, this determines the relationship between the system’s state and the system’s representation thereof, and e.g., if the system has a way of producing sounds, it may produce audible error codes for a human operator to interpret. Additionally Norman (1991) writes that the power of cognitive artefacts lies in their role as representational devices and he divides the representations into two different kinds: surface representations and internal representations.

Surface representations are artefacts that are only capable of the displaying and maintaining of symbols, such as notepads and displays (Norman, 1991). Artefacts with internal representations, such as calculators, maintain symbols within the device. These artefacts are therefore in need of an interface that is able to transform the internal representations into surface representations (Norman, 1991). This brings us to a special kind of artefact: the object symbol. An object symbol is an artefact that, though a surface representation represents the relationship between control operation and system state. This means that the surface representation indicates both that the artefact is operable and its current state, when a change in the artefact’s state occur, the surface representation of the artefact change in reflection of this, levers being a classic example (Norman, 1991). As representations inherent in artefacts support both the execution and evaluation phase of the action cycle, properties of these representations affect the interactions that involve the artefact (Norman, 1991). In this context, the action cycle would be the overarching description of the relation between the world, our goals and the actions we take to achieve them.
When accomplishing a task thought the use of an artefact, Norman (1991) describes it as there being two points of view. One point of view is the one of the bystander, who sees a system consisting of an individual and an artefact. From this point of view, the individual seems to benefit from a cognition-enhancing artefact, as the system can do more than either the artefact or the individual on their own. The other point of view is the personal one, which tells the story of how the artefact affected the individual in performing the task. From this point of view, the individual’s cognitive abilities have not changed. However, in order to perform a task with any given artefact there are requirements now in need of fulfilment. These new tasks enable the usage of the artefact, which means that the initial task has changed. This brings about an important change; by deconstructing the task into sub-tasks, the old task is now subject to what Hutchins (2006) has named pre-computation.

3.2 Social Usage of Artefacts

As the previous section testifies, most of the work on cognitive artefacts is from the perspective of an individual. Moreover, the focus of this research is on how individuals’ process information stored in representational symbols such as artefacts. This, however, does not mean that the social angle is suffering from neglect. One example would be the Agents & Artefacts framework (Ricci, Violi & Omicini, 2007) which identifies three classes of interactions between agents and artefacts: internal actions, communicative actions and pragmatic actions. Internal actions are also known as epistemic actions, which are actions whose primary purpose is to modify mental tasks, communicative actions tellingly enough facilitate communication, and pragmatic actions are actions taken to fulfil a need to alter the world in order to fulfil a physical goal (Clark, 1997). Ricci et al., (2007) also claim that for agents to be able to reason about the usage of an artefact artefacts need three descriptions. Those three are a functional description (describing what the artefact does and in which state it is in), the usage interface description (which ought to be analogous to Norman’s surface representation), and the operating instructions (how to actually use it). All of this makes the usage of cognitive artefacts in social settings rather complex, as this means that artefacts must either inherently have these qualities or provide some means, transferring knowledge that is not inherent or afforded by the artefact, by which common ground can be established. This explains why we need extensive training to use basic cognitive tools. One way to acquire this training would be for an individual to study artefacts (which would cause the “three men and an elephant in a dark room” problem).

3 A community of blind men once heard that an extraordinary beast called an elephant had been brought into the country. Since they did not know what it looked like and had never heard its name, they resolved to obtain a picture, and the knowledge they desired, by feeling the beast ... and when they had found it, they felt its body. One touched its leg, the other a tusk, the third an ear, and in the belief that they now knew the elephant, they returned home. But when they were questioned by the other blind men, their answers differed. The one who had felt the leg maintained that the elephant was nothing other than a pillar, extremely rough to the touch, and yet strangely soft. The one who had caught hold of the tusk denied this and described the elephant as, hard and smooth, with nothing soft or rough about it, more over the beast was by no means as stout as a pillar, but rather had the shape of a post [amud]. The third, who had held the ear in his hands, spoke: "By my faith, it is both soft and rough." Thus
Another, perhaps better, alternative would be the research supporting the Vygotskian explanation of social mediation, in which a process of dialectical change causes internalisation of external tools (Rieber & Robinson, 2004). This would mean that learning is a socially interactive process where our understanding is challenged by our surroundings, forcing us to reconsider our understanding in order to reconcile our and the surrounding's understanding of the subject. "Learning", however, is a word with many interpretations, which is why Stahl (2002) describes "common knowledge building" as a preferable term as it, in his opinion, is more tangible, and he motivates his position as follows:

It cannot simply be applied everywhere, but refers to specific, identifiable occurrences. Cases in which new knowledge is actually constructed by groups – rather than reified facts being recycled – are actually relatively rare in classrooms.

With care and practice, one can directly and empirically observe the knowledge being built, because it necessarily takes place in observable media, like talk. Moreover, it produces knowledge objects or artifacts, which provide lasting traces and a basis for evaluating the knowledge building. (p. 63, Stahl, 2002)

Stahl (2002) further elaborates that it is under these circumstances participants build knowledge through discourse, allowing for mechanisms such as peer review and critique. The purpose of these mechanisms is to to cause conflicts, which Stahl stresses to be essential; conflicting perspectives serve as a ground for processes such as questions, arguments, negotiation and agreement about the artefacts, which are at the centre of the discourse. As the observant reader will have noticed, this definition is not far off the definition of Common Ground, and this will serve us as a pointer for where to look: in the social interaction, and what to look for in it: conflicts centred on artefacts.

With this (Stahl, 2002) approach, knowledge gained will be shared by the group and reflected in the shared understanding which is being centred on increasingly elaborate artefacts. However, individuals may still retain some of the group’s collective knowledge, which in itself is consistent with Vygotsky’s notion of a zone of proximal development, in which an individual’s developmental level is higher when, e.g., being able to, but not limited to, exploiting the scaffolding properties of the group. Scaffolding is an attribute commonly ascribed to cognitive artefacts, which describes our ability to exploit our social surroundings to support our cognitive abilities (Rieber & Robinson, 2004). Further, Stahl suggests that this should be studied using conversational analysis as:

...meaning of individual utterances is not given by some preconceived ideas represented in the speaker’s mind or from her personal perspective, which are then expressed and conveyed in verbal symbols. Rather, the meaning of the utterances is negotiated by the speaking and responding parties; it exists only in the group perspective that is formed by the intertwining of personal perspectives in the communicative interaction itself. The meaning of a specific utterance may be defined and affected by

he agreed with one of the others, but went on to say: "Nevertheless, it is neither like a post nor a pillar, but like a broad, thick piece of leather." Each was right in a certain sense, since each of them communicated that part of the elephant he had comprehended, but none was able describe the elephant as it really was; for all three of them were unable to comprehend the entire form of the elephant. (Muhammad al-Ghazālī “Theology Revived”, vol. 6)
subsequent utterances, responses, gestures, pauses, repairs, etc. (p. 65, Stahl, 2002)

It is important to point out that Stahl does not consider this from a Common Ground perspective, nor does he cite Clark or any of his more renowned associates, but rather researchers known for their studies of social groups or artefacts; such as Norman, Hutchins, Minsky, Streeck and Vygotsky. According to Stahl (2002), this, combined with study of situated actions like gestures, gaze, etc., has been done with small groups in face-to-face interactions, although without technical mediation, within the area of micro-ethnography, but the conclusion is still valid:

...in cases of collaborative knowledge building the unit of analysis for meanings should take into account the intertwining of personal and group perspectives by interpreting individual utterances as elements of the larger discourse and activity. (p. 69, Stahl, 2002)

Within the area of collaborative work, there have been studies that have encountered the very problems that would have been predictable, if artefacts used in production had have been considered as important as the produced artefacts. As an example of (a failure in) distributed collaborative work:

“The analysis of the impact of the computer system on co-authors’ communication and final documents indicates that they have communicated very little with each other, that their interest in the evaluation of others’ ideas and texts was low, and that they spent much more energy and time trying to coordinate their actions than co-elaborating the contents of the text. /.../. Even more important, co-authors indicate that the document written with computer-supported collaborative writing system support is not really the product of a true collaboration between thinkers. They said they had a hard time recognizing themselves as the authors of the document produced,” (p. 738, Pargman, 2003)

“Collaborating with writing tools” is a good example of where the problems are identified, but the analysis falls short by focusing on what participants do apart from the artefacts and what they implement with the artefacts, instead of the more important “How do they use the artefacts to implement stuff?” or “How do we deal with flawed artefacts?”. However, what Pargman does do is to both identify the fact that her participants do lack a shared understanding/Common Ground, and to identify assorted activities which are deemed necessary to reach a shared understanding. This is where further understanding of how artefacts are used in reaching Common Ground (within the specific activities) in social settings could have proved helpful.

Unfortunately, suggested solutions focus on increased annotation, separation of private and public work, alternative presentations (surprisingly enough in a manner which would further increase the connection between specific writers and specific portions of the text), and lastly a call for better support for sharing of text and ideas between co-authors (Pargman, 2003). Seeing as Pargman’s most obvious problem was the lack of shared understanding, which the following quotation displays:

Paul asks a question referring to the content of a paragraph written by Fernand. In fact Paul does not understand why Fernand had written “there is diversification[sic] of products”. The question about this phrase is repeated four times. It took a total of 25 min until the other party finally could identify and relate comments with the paragraph in question, that is, 11% of the time employed in writing the document. (p. 746, Pargman, 2003).
Pargman (2003) was able to identify a lack of shared understanding as a major problem and it possible to argue that only the last suggestion would have had any significant effect on the outcome, as increased annotation would mean more information about something, but probably not of the fragments, and not of the context. Separation between public and private work would only help if people were to predict that people might not understand something, and therefore add an explanation in their private notes – which is a rather counterproductive thing to do when collaborating. Alternative presentations would only have been helpful if the question was who had written what, which the participants knew in this case. As such, it stands to reason that the only thing which would have actually improved the quality would have been if the co-authors had had a higher degree of Common Ground.

As a contrast, Gutwin and Greenberg (1991) explored the possibilities of Common Ground within CMC-systems. While they did not focus on artefacts, they did have a much more faithful adherence towards the Common Ground theory. However, their workspace should not be confused with Clark’s workspace. The difference between the two is that Clark defines “Workspace” as the area immediately in front of the speaker’s torso (Clark & Krych, 2004) while Gutwin and Greenberg intends the whole of the space the speaker has available (Gutwin & Greenberg, 1999).

“Part of the problem with current systems is that they do not provide much information about other participants in the session. When people work together in a face-to-face setting, a wide variety of perceptual cues help them keep track of what others are doing. This awareness of others in the workspace is workspace awareness, the up-to-the-moment understanding of another person’s interaction with the shared space” (p. 244, Gutwin & Greenberg 1999)

They also validate the need for monitoring, which ought to indicate that the next logical step is to not only show what artefacts a speaker is interacting with, but also the “how” and “whys”, which becomes decidedly more important when dealing with the often less tangible cognitive artefacts.

An example of applied social artefacts is a case study performed by Paay, Sterling, Vetere, Howard and Boettcher (2009), they identified what they called “shared artefacts” while trying to capture social requirements as functional specifications. In their definition, artefacts two groups of people share an artefact, each group with their own collective understanding of the artefact. The meaning of “shared” is not as much about what data it represents, but as what it represents by virtue of its origin. The idea is that if one group produce and present an artefact, they will possess it, which will make others less inclined to interact with and modify it (Paay, et al., 2009). To circumvent this problem sharing of artefacts through joint creation was essential: “Both evolved as a participatory activity, with team conversations facilitating the cooperative identification of aspects of both artifacts.” (p. 451, Paay, et al., 2009) (it is possible to argue that it in fact was a joint activity...), as the two groups had to share the artefacts, the artefacts came to be both carriers of information and act as centrepieces of conversation. This is also an indication of the high degree of openness of the shared artefact, the more open an artefact is, the easier it becomes for other participants to modify it. This allowed both groups to negotiate a common understanding collaboratively of
what the data actually represented. In this context the artefacts primary purpose was to negotiate and cross boundaries, secondary they embodied and mediated work in collaborative settings (Paay, et al, 2009). Beyond pushing boundaries, the presence of artefacts also turned out to bring clarification to conversations through the usage of artefacts to extend the understanding (or as it would have been called in this thesis, common ground#, it is also worth noting that this turn of event is consistent with Clark’s levels of interaction).

3.3 Artefacts within Distributed Cognition

The purpose of distributed cognition, as expressed by Perry (1999), is to understand interactions at the system level through investigations of representations and the processes that transform them. This means it is still being somewhat true to the original theory of cognitivism, but rather than only considering computation on an individual level, the object of study in distributed cognition is the information processes taking place between the participants and artefacts in the system of which the interaction is part of. This is done in order to examine any emergent behaviour in the participants, which is illustrated by Nemeth, Cook, O’Connor and Klock (2004)

Hutchins and his colleagues contend that internal and external cognitive processes involve complex coordination at many different time scales between internal resources (memory, attention, and executive function) and external resources (the objects, artifacts, and materials). A cognitive process is bounded by the relationships among the elements that are part of it, instead of how close the elements are in physical space. ... Studies of group cognition in natural settings have discovered that cognitive processes may be distributed among the members of a social group. They may involve coordination between internal and external (material or environmental) structures. They may also be distributed through time so that the products of earlier events can transform later events (p. 727-728, Nemeth, et al., 2004)

Distributed cognition also includes cognitive artefacts, more precisely, artefacts as a part of a cognitive system which distribute its contents over both humans and their environment. Thus they can be described as:

Cognitive artifacts can be used to learn about the work that they have been designed to support. For Hutchins [12], “cognitive artifacts are involved in a process of organizing functional skills into functional systems.” This amounts to “a category of processes that produce cognitive effects by bringing functional skills into coordination with various kinds of structure. (p. 732, Nemeth, et al. 2004)

With such a holistic view of interaction, it seems to be a model of social interaction, which ought to be a good fit with Clark’s framework, as it is possible to abstract both to a system level. Another reason why a good fit should be possible is that, according to Hutchins (2006), Distributed Cognition shares the notion of social mediation developed by Vygotsky. Accordingly, all interactions between humans are thought to be embodied through interaction with artefacts situated in the world and that interaction itself is often situated in a social context in which usage of external tools becomes internalised through dialectical change.

Hutchins (2006) is of the opinion that there are two ways to interact with artefacts. The most obvious one is the body/world interaction, where interactions with artefacts are realised through physical interaction. The other
is participants talking about the world, which any given artefact is part of, and through the participants dialogue cognitive structures are transformed (Hutchins, 2006). As Hutchins describes it, the body itself can become a cognitive artefact, allowing for meaningful environmentally coupled gestures. This is achieved by cultural constructions of meaning of body motions, put into relation to the body’s spatial relation to the world. At this point, Perry (1999) describes it as actions bringing representations into co-ordination, which allow information to propagate, through an extended cognitive system which encompasses all parts of the functional system describing the activity, and continuously modifying it until an acceptable resolution has been found. The functional system consists of all representation carrying and transforming entities involved in the activity. This way, for an example, environmentally coupled gestures or direct manipulation of artefacts allow for an embodied approach to problem solving, also known as embodied reasoning (Hutchins, 2006). For mental structures and artefacts, Hutchins stresses the importance of multimodality for stability. The more aspects/modalities of an artefact we know and are able to describe, the more stability it gains as a cognitive artefact (for an example, it is easier to describe any pen-like object to a person familiar with a pen by describing which modalities differ between the pen and the pen-like object). This makes the artefact and its representation easier to manipulate and by integrating cognitive artefacts into actions performed in the world, it becomes easier to reach a shared understanding with other participants (Hutchins, 2006).

Perry (1999) provides definitions of two kinds of distributed cognition, of which the first is the individual distributed cognition (IDC), which is considered to be the most understood description of artefact usage.

...individual distributed cognition: this term is applied to cognitive systems involving a single actor, and one or more artefacts (cognitive artefacts, Norman, 1991). Good examples of such studies have included the use of a ‘flight bug’ (Hutchins, 1994, 1995b), and the solution to the Tower of Hanoi [sic] puzzle (Zheng and Norman, 1994).

These studies have investigated how human intellectual resources engage with, and cannot usefully be seen as separate from, the physical resources in the world in task performance... (p. 87, Perry, 1999).

The other kind is the socially distributed cognition (SDC) (a definition which Perry (1999) explain to have been extended by him and Macredie), which also considers, and therefore in comparison to IDC comes to focus on, social coordination through communication between the participants of the system (Perry, 1999). Another difference is that in IDC, artefacts are merely resources, while in SDC artefacts are also meta-resources. This means that artefacts from this perspective also serve as tools to use other artefacts (resources).

socially distributed cognition is used specifically to investigate multi-person activities, often in concert with physical artefacts that act as cognitive resources (as in IDC), but also act as intermediaries in communication between individuals ... A further distinction within this has since arisen ... in which SDC systems can be divided into two forms:

a) well structured systems, in which all of the problem solving resources are initially known to the ‘functional system’.
b) ill-structured, or ‘messy’ systems, in which the participants, processes and artefacts involved are initially understood or unspecified (p. 87, Perry, 1999).

However, while these definitions show that artefacts have a use in accomplishing tasks, they still shed no light on how artefacts contribute to the process of achieving common ground, only how people use artefacts used in activities to solve problems computationally. The significance is that common ground, also known as shared understanding, is integral to collaborative problem solving while simple computation only explains individual problem solving, which might allow for cooperative problem solving, but not collaborative. This means that for the purpose of this thesis the “messy” systems definition seems to be a good starting point. Clark’s grounding theory (Monk, 2003) shows that conversations are lacking in structure, often requiring restarts, repairs and other joint actions to keep the conversation aligned. This causes linguistic artefacts such as structures and sentences to be reconstructed on the fly if they are deemed insufficient for the purpose of grounding. Additionally Monk (2003) also shows that conversations often consist of fragments that are assumed to be meaningful. On the subject of meaningful communication, Perry, Fruchter and Rosenberg (1999) observed something they called channels which is something Clark’s theory of Common Ground seem to be lacking.

There appears to be a simple principle at work here, with users minimising their effort - they do not tend to bring more channels into operation than they need, but add channels when they require them. Where communications channels are deemed to be unnecessary and hard to maintain, they are often closed down. For example, whilst the team found the video-conferencing technology useful in developing an awareness of the context of the remote setting and helping them to coordinate their ongoing interactions, it was not deemed to be a significant part of the collaborative process. When it proved to be hard to use (usually technical difficulties), it was abandoned. (p. 148, Perry, et al. 1999)

Given the circumstances there is a need to understand how artefacts and fragmental conversations fit into the system.

### 3.4 Unsolved Issues with Artefacts

There is no clear line between mental and social artefacts, physical patterns and opportunistic practices. Artefacts seem to be not as much objects as a process to coordinate functional skill sets with various structures (Hutchins, 1999).

A problem Stahl (2002) identifies in the application of this method is that participants divide projects into tasks, leading participants to cooperatively share results reached through non-collaborate work. Another (Stahl, 2002) caution is that when applied to social contexts, symbols and representations are only as meaningful for individual minds as their pre-existing knowledge allows them to infer.

### 3.5 Summary of Artefacts

This chapter starts with an exposition of cognitive artefacts; what they are, how individuals use them, what sets them apart from other artefacts and their primary attributes and components. After establishing some basic facts about
cognitive artefacts, the known social usage of artefacts in the learning and knowledge building domains are subject to exploration. Here we learn that two important aspects of artefacts in social settings are coordination and offloading, which allows for a passing connection to the zone of proximal development, which builds on dialectical change – which itself has more than a passing resemblance to Clark’s (1996) suggested grounding. Stahl (2002) also suggests that as artefacts become the centrepiece of interactions, interactions around the artefacts should be studied using conversation analysis, which the Common Ground theory utilises, in order to understand how shared understanding of the artefact comes into being. Additionally, this thesis explores distributed cognition in a bid to further that understanding. Distributed cognition contributes with an understanding of how humans interact with artefacts in a natural setting, and that man and artefact alike both are part of a larger functional system that is possible to study with ethno-methodological tools. This system view adds some of the computationally equivalent properties known to earlier models of human communication, which Clark aims to bridge. The following chapter will detail the subject of inquiry in this thesis, where and how this thesis seeks for answers.
4 Issue to be Addressed

It might be apparent but someone need to say it: In order to perform a collaborative task, defined as an iterative creative process where two or more people work toward a common goal – often with a leader, people need to coordinate their actions, which allows for a shared understanding of the task. This is different from cooperative tasks, here defined as one off occurrence where equals jointly solve a problem, where the tasks might not be concurrent but sequential. A collaborative task requires some coordination, and adds some constraints to the interaction. The most commonly used tool to achieve this coordination is language, and as such Common Ground (see chapter 2), which aims to formalise both joint activities and the grounding process where a shared understanding is built, seems to be a good fit.

Common Ground is understood as a shared understanding where all participants has roughly the same knowledge and reason to believe that other participants have the same knowledge, and how common ground is achieved - it becomes apparent that artefact usage has not been considered as an integral part of joint actions. However, in the criticism (see chapter 2.3) of Common Ground, it has been pointed out that artefact mediated communication is problematic for the grounding process. The problems identified relates to presentation, acceptance and closure (Koschmann & LeBaron, 2003). At this point, it is possible to hypothesise whether this is caused by the lack of apparent evidence or some other cause, absence of evidence should not be seen as evidence of absence, as there may be other explanations. Chapter 3 on the other hand describes what cognitive artefacts are, defined as “an artificial device designed to maintain, display, or operate upon information in order to serve a representational function” (p. 17, Norman, 1991), how they are used by individuals and how groups use cognitive artefacts. Chapter 3.2 displays how important artefacts are in social settings where a common understanding is essential. All this is in search for defining attributes that may explain how cognitive artefacts are easing the achievement of common ground in collaborative groups.

The issue addressed in this thesis is “How do cognitive artefacts contribute to the process of reaching common ground within collaborative groups?” This is a legitimate question by the fact that the bulk of research has been conducted on, e.g., distributed cooperative groups working synchronously, and how they reach common ground through usage of computer mediated technologies. There have been some studies in the area of asynchronous work among small cooperative groups, but seemingly only in the context of cooperative writing. At least one of those studies, although done from an activity theory perspective and not Common Ground, (Pargman, 2003) has shown that supposedly synchronous collaborative groups actually choose to work more or less asynchronously and very much individually (see chapter 3.2). However, there seems to have been little research into how cognitive artefacts contribute to the process of reaching common ground within cooperative groups that actually makes the work collaborative, which is the focus of this thesis. Given the previous chapters, there is a need to investigate the interaction between, on the one hand, grounding and Common Ground, and cognitive artefact usage in collaborative setting. On the other hand, grounding, based as it is on face-to-face conversation, is done through the exchange and assimilation of cognitive
artefacts (on several levels, firstly as individual words and later as concepts and structures) and therefore this thesis will study to what extent cognitive artefacts are used in the grounding process when reaching common ground#. This means that the primary areas of interest within the Common Ground theory will be within grounding and the first and third levels of description (see chapter 2.2.2) of interaction applied to joint actions. The intent is to apply the Common Ground framework to a naturalistic setting without preconceived notions and to enable the observation of usage of cognitive artefacts during any grounding processes, by means of an ethno-methodologically inspired approach. This should ideally allow for an extrapolation of how present cognitive artefacts contribute to common ground in a collaborative setting and which process they are an externalised representation of.

4.1 Areas of Interest

The focus of this thesis lies on two areas; the first area is that of Common Ground, which depends on of a series of co-constructed joint actions, relegated by repairs, restarts and alignment. The second area is distributed cognition, elaborating on problem solving utilising regular artefacts; this thesis is not about problem solving per se, as standard symbol-manipulation theory has proven that understanding is not a requirement for problem solving. Therefore, this thesis will extend on existing materials to and increase our understanding of how cognitive artefacts contribute to the process of establishing common ground#. Cognitive artefacts, in this context, refer to objects that are more or less tangible, not words nor culture. While it is possible to breakdown culture and all of its components to a large amount of cognitive artefacts, this thesis will maintain that there is a difference between physical cognitive artefacts and immaterial ones, for the purpose of analysing collaborative behaviour. The reason why physical cognitive artefacts are of particular interest is because we can fairly easy see how people use artefacts’ presence or absence to create, infer, process and represent information. As such, e.g., language is considered as a separate set of tools and only interesting for consideration when it appears as part of a structure.

4.2 Scope of the Present Thesis

As the purpose is to reach an understanding of how people use cognitive artefacts to achieve grounding, areas as CSCW and CMC will not be an object of exploration, as those areas would benefit from the understanding sought in this thesis. That is to say, this is about theory, not implementation. Nor will questions which are pure social science, cognition as computation, or other areas of linguistics or common ground which are not explicitly mentioned (details are to be decided) in the thesis be pursued. Artefacts, as a matter of inquiry, will be limited cognitive artefacts to their usage in social interactions as a tool to reach a common understanding.

The next chapter will detail the methods, methodology and procedure used in search for answers to the question of how cognitive artefacts contribute to the process of reaching common ground within collaborative groups.
5 Methodology, Methods and Procedure

In this chapter the methods and methodology is explored, additionally selection of participants and validity of methods are discussed in the text. In short: the object of study was selected to allow observation related to a prediction grounded in Common Ground and distributed cognition. The study accomplishes this through an ethno-methodologically inspired study, inspired because it only covers almost an hour worth of interaction and not the month or years required by a proper ethno-methodological study. This documented and recorded artefact usage in a natural setting, coupled with a structured interview to measure consistency between what the participants are observed to do and how they themselves motivate their actions.

5.1 The Scientific Method

To explore the issues presented in earlier chapters this paper will use qualitative methods such as an ethno-methodologically inspired direct observation study coupled with structured interviews; these are subject to interpretation in accordance to the presented theories. This can be motivated in two ways. The obvious one is the prevalence of ethno-methodologically inspired methods used by proponents of both Clark’s theory of Common Ground and several schools of artefact studies, including but not limited to studies within distributed cognition. The other reason is the emphasis on interpretation of observations in accordance with subjects’ understanding (Bryman, 1995), the subjects’ understanding being key as common ground is all about understanding. Additionally to a focus on the understanding, ethno-methodology has a focus on the process it studies, “the process through which the perceived[sic] stable features of socially organized environments are constantly created and sustained” (p. 16, Coulon, 1995).

Every day collaborative tasks are arguably socially organised environments going through stages or construction, re-construction and at times, de-construction. Coulon (1995) further argues that an ethno-methodological approach is particularly well suited when the objective is to identify methods through which actors realise rules causing certain behaviours: “Therefore, is it crucial to observe how, in a commonsense manner, actors produce and treat information in their exchanges” (p. 17, Coulon, 1995). Coulon (1995) is also of the opinion that ethno-methodology allow for exposure of patterns in actions, which in turn allows an observer to identify what these patterns express. While expressing pretty much the same sentiments Patton (2002) argues that direct observation is especially suited when trying to understand the context and gain a holistic understanding and that firsthand experience allows the observer to draw on personal knowledge during the formal interpretation. He goes on to say, “Reflection and introspection are important parts of field research” (p. 264, Patton, 2002) noting that the observer becomes part of the data which is used to create an understanding.

Data from the ethno-methodologically inspired study was collected through a digital video camera recording while the participants were going about their tasks. Special attention was paid to their work areas and indexical gestures as Common Ground ascribe increased significance to these and other areas when communication occurs. By doing this, several patterns of action and work
practices could be identified. The video recording itself was made openly with consenting subjects; they were, however, not informed of the exact nature of the study. A total of 55 minutes of recorded data was collected. After the conclusion of the study, subjects participated in an interview using a structured method with formalised questions in order to ensure that common ground had occurred.

The unit of analysis was narrowed down to physical cognitive artefacts and how the participants used them in negotiating common ground#. The subjects were selected through purposeful sampling (Patton, 2002). The purpose was to lessen the impact of factors, such as language, which could have detracted from the importance of physical cognitive artefacts in the activity. Therefore, the sample was designed to be homogenous, as it is likely that people working together will develop a certain level of efficiency not found among strangers trying to collaborate. This vaguely defined level of efficiency represents the uncertainty caused by discrepancy in word usage and sentence construction, which may cause breakdowns at certain levels in accordance with Clark’s Common Ground theory. Thus, pair of brothers was thought to be homogenous enough, i.e., sharing a common vocabulary and mannerisms, to bring forth the contributions made by cognitive artefacts, without the added complexity of uncertainty through semantics.

The ethno-methodologically inspired study was open-ended and was implemented through the usage of two subjects who performed the everyday collaborative task of cooking in a rather small kitchen. No instructions, comments or advice were at any time provided by the observer. Thereby this ethno-methodologically inspired study ought to be comparable to other ethno-methodologically inspired studies.

5.2 Setting and Social Context

The ethno-methodologically inspired direct observation study was conducted in an everyday environment, which also fits the description of a messy system (see: end of section 3.3) where the properties of emerging artefacts are unknown, a kitchen. The two brothers performed an every-day task, cooking, and they the observation was an overt video recording. The chosen task of cooking was a two-course meal, roast beef with a vegetable gratin and panna cotta, as it is a collaborative task where a common understanding is subject to continuous negotiation during the time of preparation. The preparation time is also quite long, around an hour, which would give plenty of instances of collaborative behaviour that would require negotiation of the common ground. Additionally it is something that allows for the usage of a wide range of cognitive artefacts: from physical objects such as cookbooks to physical structures (as shown by Kirsh in his “Intelligent use of space”).

In addition, it adds constraints to the study such as timing for task completion, which serves to ensure that it stays a collaborative task and does not turn into a cooperative task. Therefore, it was considered suitable to provide information rich situations. The selected method had the objective of capturing how common ground# is achieved through usage of artefacts in the natural setting of a kitchen, between two persons preparing a meal; or more specifically, finding a pattern of artefact usage in the individuals’ workspace while performing a collaborative task requiring common ground#. Also, as
cooking task requires a well defined work area which is thought to ease the capture of “work”, “signals” and possibly other forms of interaction as described in earlier chapters.

The scene of the observation was a medium sized kitchen in a three room apartment in the outskirts of Skövde, Sweden, with a westwards facing window, overlooking the access balcony, allowing the setting sun to bask the kitchen in light as the study commenced (see: figure 1). The kitchen itself is equipped with a full size refrigerator and a half size freezer, and an oven with a stove, two sinks with a free swivel faucet, a non-functioning clock in the shape of some kind of tableware, painted to resemble a practice target – the hands are, deliberate or by chance, set to five minutes to twelve. Apart from that the kitchen is sparsely furnished, a brown wooden table with three matching chairs evenly spaced around it. As for the walls, the large window with its curtains dominates the room, only the impressive collection of spices lording on its perch above the stove, opposes it. On the spice rack was also a non-functional timer, mute and functionally replaced by a mobile phone on the dining table.

5.3 Structured Interviews

Open-ended structured interview was selected as a supporting method in order to find out if common ground# was actually achieved, the theory being that if common ground# is not achieved, participants will answer the questions differently. According to Patton (2002) open-ended structured interviews have four major advantages. Out of these four, three applies: anyone reading this paper can replicate the interviews, structured interviews are focused and short which may not facilitate the collection of a lot of data (but collected data will be relevant), lastly it makes the analysis easier as answers to the very same questions are easier to compare than answers to questions which are somewhat alike.

During the interview, both questions and answers were in Swedish and a full list of interview questions is in the appendix. A note to be made about the questions is that question number two had to be clarified to the participants, as it caused the respondents to ask if verbal or non-verbal communication was sought after, which was clarified to them as both being valid forms of expression. The objective of this method was to gather data describing how the participants themselves experienced their artefact usage, i.e. measuring the degree of consistency between objectives. As such, they were created after the object of study and unit of analysis was decided, but before the study was implemented. The structured interviews also allowed for some degree of respondent validation as the questions pertain to the recorded collaborative task.

5.4 Procedure and Overview

In the kitchen, two twenty-something year old brothers set out to prepare a medium complexity two-course meal (medium complexity since there was no need for finesse, however, there were certain stages that required the use of a thermometer). This was studied through direct observation which was video recorded with a simple digital camera capable of motion picture capture, equipped with a microphone. That which happened; two brothers engaged in a
naturalistic collaborative task, cooking a two-course meal that was subject to a
direct observation ethno-methodologically inspired study. After the study was
completed, they were asked to participate in complementary structured
interviews, consisting of eight questions repeated for three selected sequences
of the observation with a finishing question that asked them to detail any
thoughts concerning how “objects” were used in their communication. This
was aimed to measure primarily if the brothers had achieved common ground
during their collaborative task, and if so, had they done it through usage of
artefacts? Additionally the questions attempted to measure the participants’
awareness of the various roles artefacts can play in communication.

It took the brothers almost precisely 50 minutes to prepare a piece of roast
beef with a vegetable gratin as the main dish and panna cotta with strawberry
sauce for dessert. Collaborative behaviour could be observed throughout these
50 minutes, but it was especially tangible at different stages, two involving
the vegetable gratin and one the recipe itself.

The first instance was the interaction with the recipe, the older brother reading
key aspects of the recipe aloud, of which the younger brother repeated parts,
which would prove meaningful later when they relied on this shared mental
structure instead of consulting the recipe once again. The second instance was
during the vegetable processing stage, where the older brother placed
vegetables at the left side area of the younger brother working area (see:
figure 1). This allowed younger brother to process them (in the left area of
the common work area which they shared) as he saw fit, and both brothers jointly
put the processed vegetables into the gratin form. The third instance was at the
stage where it was time to season the gratin. At this stage the older brother first
made a selection of spices which he put on an empty area on the left side of the
stove, returning one kind of spice as he changed his mind about including it. Then he
moved the remaining spices further into the common workspace where the
younger brother in turn was literally prodded to comment on them. With a short
utterance and a gesture, the garlic powder was banished from the selection and the
older brother returned it to the spice rack.

After the direct observation the brothers were asked to participate in an open-ended structured interview which they agreed to. Each was interviewed in the
comfort of their homes and took about 75 minutes to conduct, the older
brother was interviewed the day after the study and the younger brother two
days after the study was conducted. The older brother was interviewed in the
evening in front of his computer in the room next to the kitchen in which the
study had been conducted - and the younger brother was interviewed during the afternoon in front of his computer in his student apartment, located at Plutonen, Skövde, Sweden. During the interview, they were shown a couple of clips from the collected materials to which they were asked to answer questions. The questions were designed to elaborate on their actions in the current scene, in order to establish if they did it in response to something, did it to cause a response, if they consciously understood what to do next based on others interaction with cognitive artefacts etc. The questions also included the collaborative task they had performed and they were asked to review some of the recordings to which they were asked to relate their intentions of their actions, and their inferred intentions to the other subject’s actions in the same clip. This was done to match subjects’ interpretations to the observation of their interaction.

Examples of questions are “What you are doing in this scene?” followed by “How did you understand the goal of this activity?”, “What did you intend to communicate?” and “What makes you think you were understood?” as a final question they were asked to recount any thoughts they had on the importance of artefacts in their communication. Their responses were then compared, if one of the brothers reacted on what the other did, and his action was what the first one expected, it was deemed that common ground had been established and that the previous actions had been grounding and the last question was used to assess their awareness of artefact usage.

The following chapter presents the analysis and results.
6 Analysis and Results

This chapter details what data was collected and how it was interpreted.

6.1 Analytic Overview

The usage for data collected through the recording of the direct observation was twofold, of which the repeated observations it allowed was the most important. The other application of the collected data was during the interviews where the interviewees were shown sequences of the collected data and asked to answer questions related to the sequence. The repeated viewings of the recording allowed for a continued search for patterns of how artefacts were used beyond what could have been noticed during a simple direct observation. These patterns were divided into categories depending on their intention, which were as follows:

- Structures for public use, e.g., arranging of tools with the purpose of adding options to contextual interactions.
- Structures for private use, e.g., arranging items to suit the workflow.
- Indirect usage, e.g., using the recipe to direct work involving other artefacts.
- Direct usage, e.g., using tools to affect change and the usage of artefacts to create common ground through glances.

All categories had nine to twelve instances associated with them; however, direct usage was not exhaustively investigated, as that category’s potential for collaborative was deemed insignificant. Then after structuring the data from the study into categories, the answers given in the structured interviews were matched to these four categories in order to gain an understanding of awareness and if there was intent to the collaboration or only “happy accidents”.

6.2 Artefact Usage

As previously mentioned the artefact usage was divided into four categories which of three were deemed interesting. The categories in the study that were most common were structures for public and opportunistic usage followed by direct usage, the latter will not be pursued further as its collaborative qualities are lacking.

6.2.1 Structures for Public use

The study shows that common ground seems to be commonly created through the creation of structures, which all participants can exploit, and given that they can fit any of the three classes of interaction identified in section 3.2, are probably the most open artefacts in this study. Structures can accommodate internal actions, as Kirsh (1995) identify the simplification of internal computation as one category of reasons for structures, structures can also be a communicative act (further elaborated later) and structures fits pragmatic actions as they simplify choice (Kirsh, 1995). A common way of achieving these public structures is by placing physical objects into other person’s workspace in order to either propose joint projects or prompt a joint action involving the object. This behaviour is in line with both Kirsh “intelligent use
of space” and the part of Clark’s theory of common ground. There Clark asserts that when creating common ground we invest the least possible amount of work and that we are opportunistic, both in the sense that we are using assumptions of knowledge and in the sense that we construct actions as being joint if it fits our individual understanding of the activity, or suit our domain goals. This could be observed in the study as the older brother arranged artefacts in the form of vegetables and tools in the workspaces. To an observer it seems as if he is preparing the workspaces, but during the interview, it was made apparent that the purpose was not merely to enable actions or create a smooth workflow, but also to implicit influence which actions were to be taken. It was substantiated by the older brother who claimed that he “seeded” the environment to provide “cues” of what would be needed next, enabling the younger brother to continue working without interrupting the workflow. This is in line with Kirsh categories of spatial arrangements, as these preparations were deliberately made to simplify perception and thereby simplify choice. In the same way, artefacts were placed in the sink to remove them from the workspace and to give them a new meaning: these items are dirty, do not use without washing, which is grounded in the pre-existing knowledge that items in the sink should not be used for hygienic reasons, which is yet another example of simplified perception. Likewise, the sink was cleaned out to afford the putting of dirty dishes into it.

In this context, it should be safe to assume that everyone is aware of all artefacts that are within plain sight, and their affordances. For an example: dirty unwashed vegetables to left of the sink, pretty far from the work area and washed vegetables on the other side of the sink, near the chopping board, this is a structure which is an artefact (see: Normans definition of artefacts, section 3.1). It is an example of a structure where prepared and unprepared vegetables are separated creating a structure that allows for both faster computation: how much do we get, how much do we need, and simplify perception: which vegetables are clean - the ones located near the chopping board. This reduces the need for communication between participants to identify which vegetables belong to which category. That is, instead of initiating a joint project sorting the vegetables one by one, the artefact structure is exploited, allowing a participant to assume knowledge about the structure based on previous knowledge and act accordingly. If the assumptions are wrong, then the assumptions will lead to incorrect actions, which given Clark’s theory of Common Ground, will be discovered by other participants, as we monitor other participants work spaces, and be interpreted as a symptom which will give rise to joint actions in an attempt to repair the common ground. On a brighter note, this also allows for implicit communication where people act on artefacts to the best of the knowledge provided in the shared common ground, as repairs will result in further grounding allowing correct actions that will still mean less spent effort compared to explaining from scratch.

Likewise, use of structures can express objection as well as acceptance. While placing a peeled carrot next to someone slicing vegetables is an implicit proposal of a joint action, in which, the person slicing vegetables will notice the carrot, consider using it and either present evidence of acceptance by slicing it along with the others or reject it by not using it. In this position, it would be able to exploit the structure by returning any unsatisfactory artefacts to their previous location, which in Common Ground would constitute a
signal that would initiate a joint project, in this case a repair. The purpose of such a repair is to initiate new joint actions striving to reach common ground describing why it rejection occurred and what can be done to rectify it. This was observed when the brothers coordinated the selection of spices, the older brother (as illustrated by figures two to five) moved the spices into the younger brothers work area, at which point he caught his brothers attention by asking, “Do we want to use garlic and basil?” (Figure 3), then they negotiated through some quick joint actions in which the younger brother looks at the spices and asks “don’t we have any fresh garlic available?” (Figure 4, due to less than impressive camera handling it is hard to see, but the older brother to the right was watching the younger brother, the younger brother was, during the exchange, gazing at the presented spice) which results in a reply from the older brother to the effect that there is fresh garlic, and the garlic powder is returned to the rack.

However, it does give us that if someone is not following the plan envisioned through a cognitive structure, it will require further joint actions to establish the cause of rejection and to negotiate a new common ground.

Another example of public structure use is the establishment of mental structures such as a verbalised list of items, which participants use to establish public goals. By distributing the structure between participants, it is possible to avoid interruptions of the workflow in the shape of leaving the workspace to check the cookbook, which was situated on the table behind both participants, by effectively perform an alignment action by exploiting the
shared cognitive artefact created through the verbalisation of the recipe. This was observed in the study when the older brother read the recipe out aloud, effectively turning the book artefact into a mental structure (see: section 3.1) which served as a distributed cognitive artefact (see: section 3.3). In the interview he motivated this turn of events by saying that he wanted his younger brother to know what they needed to do, and what they would need to do in the near future, which could be formulated as an intent to create a cognitive structure which simplified choice of actions. Both brothers later exploited this as they asked each other about different aspects of the recipe.

Additionally Clark lists “workspace”, i.e., the area in front of a person’s torso, as one area that is monitored. This monitoring is something that is exploitable using artefacts, which is done by the arrangement of artefacts in their workspace, thus it ought to be safe, for the purpose of communication, to assume that people will make assumptions about any new artefact in their workspace and act up on its affordances. Symptoms displayed in the handling of said artefacts would alert other participants to any significant discrepancy between the shared common ground and the displayed understanding of the actor. That being said, the action of establishing structures can itself be viewed as a signal, proposing joint actions, e.g., presenting, which in itself adds the possibility for grounding actions if the presentation expose a lacking common ground. Presenting artefacts like a cutting board and peeled vegetables is a proposal for someone else to start cutting, if it meets with rejection there are two options: renegotiate a common ground or utilise the structure for oneself, which leads us onto the next section.

6.2.2 Structures for Private use

Structures for private use, that is arrangement of artefacts by an individual to offload cognitive effort onto the environment for their own use, are hard to tell from the public structures. The reasons why is because both of the fact that intent of creation is non-obvious, and that once a structure is created someone else might decide to use it. This is probably due to the open nature of structures, once again, the interactions with these structures can be internal or pragmatic (see: section 3.2), it is when they also happen to be communicative there is an overlap with public structures. However, the interviewees corroborated most of the instances identified in the study. From the materials collected during the study there seems to be two ways to interpret the usage of “structures for private use” as 1) symptoms and 2) as general-purpose.
information carrying structures. During the study it was observed that participants created structures of artefacts in their work areas to ease the work flow and direct their work.

This was displayed as shown by figure six, where a bowl with gelatine, serving bowls and, just barely, the saucepan in which the cream is brought to boil. Figure seven depicts the cream and gelatine mixture being poured from the saucepan into the strategically placed serving bowls, which constituted a cognitive structure which simplified computation by adding physical representations of the remaining steps (add gelatine, pour) and calculation (just fill the bowls, no need to keep track of how many portions were to be served). As an individual, for itself, makes this kind of structure this category of activity is almost beyond the scope of this thesis, but because it would be able for someone else to find and exploit the cognitive structure as if it was a public structure it must have some collaborative qualities.

How would structures fit with Clark’s Common Ground theory? This is one way to look at it: Clark (1996) does not elaborate on the purposes or advantages to talking to ourselves, but at least within distributed cognition we can attribute at least one explanation to Vygotsky. In talking to ourselves, we create external structures of knowledge, which is not yet fully internalised, and the result of talking to oneself is clearly cognitive artefacts. In the same way, cognitive structures are the same as interpersonal communication, often arranging artefacts in a meaningful way. Considering the collaborative nature of conversations (Clark, 1996) perhaps interpersonal communication is more than a way of rearranging existing knowledge in the hope of gleaning new insights, that is: structures are not random but purposeful. Therefore, given this perhaps structures in Common Ground are rather symptoms that social entities in the environment could interpret as a signal of someone being stuck and could use some help. Structures for private use are the kinds of structures a person create in order to distribute a cognitive task across the environment in order to help themselves. Due to the openness of distributed tasks, other persons than the creator sometimes exploit these structures, even if the intent at creation time may not have considered them, which is the same as Kirsh (1995) found.

Thanks to affordances of artefacts and cues provided by the artefacts relations to other artefacts in the environment, the purpose of the structure may still be open to participants to discern and exploit. This means that it is possible to view these structures as externalised representations of knowledge, which means that even if the purpose of their creation is cognitive offloading, as a "happy accident" other humans can also use the resulting structure for other purposes, such as learning through internalisation of those representations. In terms of Common Ground, to study an existing structure, relate it to pre-existing knowledge and incorporate any knowledge gleaned from the structure, would constitute a form of grounding. That would also be consistent with distributed cognition, which states that participants can gain knowledge of a process through the study of the involved artefacts and how functional skills fit into the functional system. Both are examples of dialectic change gained through exposure and feedback from the social interactions (assuming that the creator of the structure is nearby, ready to offer feedback).
Additionally it is possible to view these private structures as a way to signal one's role in the activity, such as the fictional and not related to the study: “I’ve arranged vegetables and knives in my work area, therefore my role in this activity is slicing things, if something needs to be sliced, give it to me”. This is something people generally do not bother to say, but is implicated by the presentation of arranged artefacts. We assume that the guy in the kitchen with the knife is both ready and able to cut things, or he would not handle a knife, and share our common goal of preparing a meal we both want. Speaking of presentation, the next section describes a form of interaction that relies solely on presentation.

6.2.3 Indirect Usage

Indirect usage of artefacts is a category that at first is hard to tell from structures, as they mainly come into play for internal or communicative actions, the main difference being in the presentation layer. While structures can be likened to object symbols, the indirect usage of artefacts relies on surface representations. These artefact interactions are almost passive in their nature, as most observed interactions were glances, i.e., a glance is a very quick look where specific information is sought, figure eight depicts the older brother quickly searching, and finding, the amount of liquid to use. Artefacts, which are indirectly used, are at the periphery of the functional system, in some way, they generate helpful but non-essential output for the system and accepts no input. Which would explain why this was the least common usage. The most obvious examples identified in the study was the cookbook and timer, as both were placed on the dining table behind the work area, which allowed both brothers equal access. This usage would probably mean that they would fall into the category of shared artefacts (see: end of section 3.2) as both timer and cookbook was used to extend the participants understanding of their activity.

Also in terms of Common Ground it can be said that anyone entering a scene has to go through all the same levels of interaction (see: section 1.2) as they would have to if they stepped into a conversation. Recognising what is done (level 1), recognise the artefacts for what they are in the given context (level 2) and their purpose (level 3) and then considering a reply (level 4). Although a reply involving artefacts would mean that the reply could be in the form of an action, not necessarily a verbal one, which Clark’s framework assumes, and as these artefact interactions in themselves doesn’t really require the fourth level, their major contribution to the collaborative work in this study was to provide topics for conversations. However, it is possible to describe conversations themselves as cognitive artefacts, which is something to be aware of. As for an example where a conversation was triggered by indirect usage, a glance at the timer did trigger the following citation: “the steak has
been roasting for ten minutes now” “then the gratin needs to be put in the oven within the next ten minutes right?” “Yeah”. This conversation represents the more complex arithmetic operation where it is known that the steak will need about an hour in the oven, and the gratin will need somewhere around 45-50 minutes, and so the elapsed time for the steak is subtracted to give the time the gratin would spend in the oven if it was put in right then. The first statement shows that the younger brother has done this calculation and figured that there is something like a ten-minute gap need, which the older brother confirms with his “yeah”.

6.2.4 Direct Usage

The subjects seldom engaged in face-to-face conversations, rather they seemed to prefer to monitor their own or any other participants work areas, and they used mainly hybrid signals (consisting of both a verbal component and a gesture) to communicate and coordinate (what, how and when) and are mainly pragmatic in nature. This may be considered to have been done demonstratively and thus a joint action, in order to elicit shared attention, i.e., “I’m shredding these” combine with a stare at the potatoes, meaning “I’m going to shred the potatoes, as I’m looking at them” but as it would only be joint if someone objected this feels rather weak. There was a lot of what could be called “talking-about-while-looking-at” which is consistent with both Stahl’s (2002) view that artefacts are centrepieces and Clark’s (Monk, 2003) monitoring, we pay attention to artefacts in our work area. However, these are commonly not indicators of joint actions or collaborative in any other sense of the word, their only purpose seems to be grounding by forcing an alignment check to ensure that all participants still share a common ground. Other direct usages involved artefact manipulation in order to determine best fit. Figure nine serves as an illustration of this, where the older brother is trying to decide which pincers to use, and to compare shapes etc. Suitable artefacts were negotiated through this physical comparison, which was the one artefact usage the participants themselves identified in the last question of the interview. The observant reader will have noticed direct usage of artefacts does not have anything to do with common ground other than that activity draws attention which may cause another participant to initiate a joint project regarding the artefacts in question.

Other observed examples were tools, such as knives, were applied to artefacts to affect change like new shapes, string was used to tie a piece of meat; measuring units were sued to measure liquids. While not in focus it is possible to add gestures here, as gestures - especially gestures with artefacts deserve consideration as symbols, as they consist of both verbal and non-verbal components. These were used to give direction, such as “open *that* door” which opened up to expose new artefacts within the cupboard.

Figure 9. The elder brother deciding on a pincer
6.3 Summary of Results

In summary, the study allowed the discovery/construction of:

- Structure usage is unavoidable
- Structures are complex and context dependent
- Coordination is achieved through open artefacts
- Artefacts wants to be shared
- Simple artefacts are at the centre of attention

The issue addressed in this thesis is “How do cognitive artefacts contribute to the process of reaching common ground within collaborative groups?” The simple answer to this question provided by this study is: through enabling us to create cognitive structures which we and our peers exploit in search of easy answers and by acting as focal points around which we can centre both the imagined common ground and further mental cognitive structures.

The long answer is a tad more complicated: the results of the conducted study and interviews indicates that by and large interactions involving artefacts follow the same basic rules as face to face conversations. Examples of this are monitoring, which supports how participants exploit both private and public structures, and how all interactions are subject to mechanisms such as repair, restart and alignment. For the purpose of collaboration, two especially important connections are possible. Firstly, simple cognitive artefacts serve as both proposals for joint projects and as tools to align the conversation, and thus provide coordination. They are also purposeful and support the distribution of activities, as observed with the timer, which was only a mobile with stopwatch functionality. Second, participants will create cognitive structures; Kirsh (1995) has already pretty much said that structures are a natural consequence of human nature. However, not only will we naturally create structures to exploit, we ground our common ground in the present structures and opportunistically construct its meanings and possibilities to reach our own goals, which is what happens in “hey, I was going to use that” situations. This does not have to be the result of a lacking common ground, rather it might be the natural result of private goals having higher priority than domain goals. Nevertheless, Clark (1996) is of the opinion that we, when talking face to face, seize any opportunity to, from our point of view, complete a joint action; it should not be surprising if we are equally eager to find the intentionality we seek in artefacts as well.

Through the observations made in the study it seems that structures are either constructed ahead of time – in case the entire structure is treated as a joint project, or they are co-constructed during the grounding. There seems to be two prerequisites for this though, the artefact must possess a certain degree of openness – enabling all participants take part and influence it, secondly fit at least one of Kirsh (1995) categories. Once these criteria’s are fulfilled the structure can be understood at a levels of interaction level (see: section 2.1.2), observations in the study supports that structures are a major factor for common ground. The study also indicated that while artefacts which are directly interacted with does not provide grounding by themselves, they do bring attention to themselves or an issue associated with them, which means that their function is to initiate joint projects, which seems to be aligned with the findings of Paay, et. al. (1995), which seems to indicate that collaboration
benefit when it is possible to utilise easily modifiable artefacts without a clear owner, which was thought to be more open and inviting as a subject of discussion.

When using artefacts to establish common ground# evidence of rejection, or acceptance, is visible through *monitoring*, as participants are bound to object if they see someone doing something wrong, or give their silent implicit acceptance. Speaking of implicitness, considering artefacts as a layer in the conversation explains implicitness. It also explains why they should be part of the consideration when analysing an interaction through Clark’s four levels. By being vague, they serve collaboration as they leave room for interpretation for both parties, giving equal possibility to judge “good enough” instead of leaving that to be judge by the more knowledgeable participant. This is preferable as the break in common ground is only subject to discovery when it is tested. This is supported by the study and interviews, as the next paragraph shows, which shows that in most cases a good-enough understanding is sufficient, which is in line with Clark (Monk, 2003) who considers that participants only need to develop sufficient common ground to fulfil their purpose. Tangible cognitive artefacts represent the current state of the system, no mere surface representations.

As gleaned through interviews, participants created most structures for their own benefit, but when the participants made the public structures, they also intended to bring the other participant up to speed and to *align* the work process. Many activities aim to communicate or fulfil *procedural goals*, working toward the *domain goal*. As for communication, they considered most of it to be intended to one-way verbal communication, but when describing it they admitted to welcoming feedback, which means it was not really a one way communication, it was proposals of joint projects which in many cases were silently accepted. They also used artefacts in order to seek attention and feedback, by positioning the artefacts in work areas or modifying them. As for the most important part: Common Ground, through the participants answers it could be established that even if they could not say exactly what the other had meant at any of the three selected scenes, they had grasped a good enough understanding of the general intentions that any differences could quickly be negotiated. Their most cited evidence, which they based their assumption of shared ground on, was “he completed the task correctly”, “he didn’t need help”, “he didn’t object” and “he did what I wanted him to do”.
7 Discussion

While it looked grim at the start of the project, when the issue of interest invariably slide into pedagogy, which is similar, but dissimilar enough to force a change of focus. This ultimately meant that a good two weeks were wasted on trying to figure out a new issue to look at to contribute something new, and hopefully significant, to the growing body of knowledge within cognitive science. With help from the advisors a new issue was found and work could, with some issues relating to overestimation of reading capability, be initiated. One of the great joys with this project is that despite a couple of weeks of sickness the project almost followed the schedule made at beginning of the project – regrettably with a constant two week delay which was incurred at the start (at least the final paper will be on time!).

Now, there are some issues with this study, and thereby its results, while the choice of a ethno-methodologically inspired study at the time was the correct one, based on this study’s results it seems safe to say that better results could most certainly have been gained from doing a different kind of study. Now, after the fact, the kind of study which probably would have generated most useful data would have been an obtrusive study in which the participants would be asked to explain what they are going to do, and define the problem space as the distance between what they said they were doing and what they actually were doing. Another regrettable issue is that for this study was the unit of analysis, as it was determined to physical cognitive artefacts it had the unfortunate consequence that the camerawork was focused on the participants hands. This caused most of the footage to be of headless people, which meant that for some scenes not all details were preserved. This leads on to another issue, if someone want to conduct a similar study, a recommendation is to use several cameras, if more cameras had been used collecting data from several angels, such as, left, right and a wide angle from the back, would not have been a problem.

While structured interviews them self seemed, and still do, very suitable given the task, questions could have been better formulated, as both participants asked for clarification on question number two. There might have been other questions that they did not understand which would be ironic given the subject. Another problem with the interviews is that the participants were interviewed in Swedish and their answers were then transliterated in order to provide citations in English, this opens up for slight bias issues as the chosen words in English might have a slightly different meaning, which may result in a citation that does not properly reflect what was said. Another issue is that one of the interview subjects was very well acquainted with the interviewer, which may have caused said object to be more outspoken while the other subject remained somewhat reserved.

In addition to this, conducting an ethno-methodologically inspired study was a learning experience, which without a doubt means there are errors both in how it was conducted and in how the results were presented; hopefully none of them are major. However, what was major was the confusion of how to infuse the recorded material with material gathered during the interviews, this means that not all data were included and it is questionable if the included data is as emphasised as it deserves to. However, as respondents answered questions,
which answered the underlying question “Did common ground occur at all?” in a manner, which indicated that it had, the structured interview can still be considered successful even if some of the potential was squandered.

Additionally answers indicated that arrangement of artefacts were done in order to save cognitive burden and streamline the workflow, which was reportedly done because the younger brother was unfamiliar with the kitchen, were also in line with Kirs’s (1995) findings in “Intelligent use of Space”, which none of them had prior knowledge of. There were also answers to, for an example 2b, which strongly indicated that communication intentionally was initiated as a form of negotiation, which is pretty much what Clark (1996) describes – the major difference was that in some of the replies it was stressed that no form of answer was expected, but would have been welcome.

Given the collected data it can be said with some confidence that collaborative work in natural settings like a kitchen relies heavily on cognitive structures, but also use other kinds of artefacts, both to perform the various tasks and coordinate them. Common ground in a kitchen is negotiated through mental structures, which are artefacts, and symbols that are, essentially, composite artefacts.

However, there are still some issues to address and two that needs to be addressed somewhere is:

- There is no clear-cut definition of mental and physical structures (see Perry, 1999),
- Mental structures are open, which is good from a perspective of Common Ground, but they are also fragile, which is not so good from a collaborative perspective, which was evident from the frequent repairs regarding the recipes.

7.1 Trustworthiness

Truth, and nothing but the truth – it is such a harsh requirement, really, so let us settle for trustworthy. To establish some degree of trustworthiness three things taken place: Firstly, the description of the ethno-methodologically inspired study is hopefully in enough detail to reproduce it. Secondly the study was recorded and analysed with an understanding in Clark’s (1996) theory of Common Ground and Norman’s (1991) understanding of artefacts and Kirs (1995) understanding of social usage of artefacts, which means there was a method to the madness. Thirdly the observed subjects were interviewed after the study was concluded, and in that interview they were shown recorded sequences to which they were asked to relate the questions, this means that they had a better opportunity to recollect what they had done and why. All three methods indicated that common ground occurred. While one does not count, two might be a coincidence, three indicators is a good sign that there is something to it.

7.2 Transferability

Transferability to other kitchens and other pairs of brothers should be relatively high, as neither had any special qualities about them, except for their nationality and location. In addition, the reasoning behind the selection is fairly well explained so it should not be too difficult to assemble a similar test
group. The study does for an example not try to answer how sisters collaborate in a workshop.

7.3 Dependability

While it is impossible to re-create the exact same circumstances, as the study is of an ordinary task in an ordinary setting with ordinary people with results which were close to what it should have been according to the applied theories, any attempt to recreate the study have a high probability of reproducing the same results. The one influencing factor here would be if they are more skilled data collectors.

7.4 Confirmability

This is a weak point, as most of the interview data is not shared in the thesis and only a few, cropped, frames of the 55-minute recording is provided. However at least a creative reader can see which question triggered which quotations. However, as the procedure is documented and described it should be trivial to remake the study, and non-trivial to analyse the data, to see if this data which is concurrent with the associated research, is wrong.

7.5 Future Considerations

While writing this thesis some problems were identified: As Koschmann and LeBaron (2003), noticed about interaction mediated through a monitor, there was a severe lack of connection between presentation and any corresponding signals of acceptance (or indeed, signals of not-acceptance). Participants seem to act on the current common ground# grounded in the present artefacts in a direction of their interpretation of the goal state in the future common ground, forcing other participants to after-the-fact either accept the new state of the common ground or demand a do-over (in essence a repair). A parallel that might require further research is the phenomena of lack of connection between feedback on presentation and the unwillingness people in groups show towards speaking up if they do not understand.

The latter is often mistaken to be an attribute of shyness and other social factors, when perhaps it is rather an indication of people in general withholding acceptance awaiting further elaborations. In addition, in the absence of further elaborations they would abstain from asking for elaborations because they have “some” understanding, which they hope will be “good enough” – and in most cases, the presentation of this “good enough” understanding will result in positive or negative feedback, which allows for a dialectical change in the participants. A major problem with this would be that in many real life cases, there are few opportunities to practice practical negotiation of common ground. An example would be formal education, the first time a student is asked to present his understanding is also the time the student will be graded on the very same understanding, or as in the analysed task, once a given ingredient is added, it’s difficult to undo and any chance of a repair lies in a re-do.

If this is the case, it is possible to argue that while the responsibility of understanding may lie in sufficient signalling and show of symptoms by the listener, it is possible to ease the grounding processed by the usage of techniques, which allows the listener to demonstrate their understanding,
thereby exposing the common ground#. One such technique for an example could be the collaborative shaping of an artefact. If one member of the group is doing the shaping wrong, the group will notice this symptom of deviation from the common ground, they will initiate joint actions for a short term joint project with the aim of getting the failing member back up to speed, fulfilling the domain goal. This approach would move the burden of alignment from where it should not be: the weakest member of the group; to where it should be – the members with the better domain knowledge who are capable of judging if the other participant's knowledge is “good enough”. This is preferable to a model where signals of non-acceptance are weak or overlooked and where symptoms of non-acceptance are ignored or discovered too late in a display of massive lack of common ground. By revealing lack of common ground earlier, the cost of repairs and realignments becomes lower.

Considering the vast amount of additional tools we stand to gain from technologies such as virtual reality and augmented reality, it ought to be a very interesting area of research within, e.g., CSCL. In the future computer mediated interactions will provide us with more artefacts and even richer environments in which we will have to solve collaborative problems, which means that effective communication is but the first problem to solve.
References


Appendix

1a) Beskriv vad ni gjorde i den här scenen
1b) Hur förstod du målet med aktiviten?

2a) Kommunicerade ni?
2b) Hur yttrade sig kommunikationen?
2c) Vad kommunicerades?

3a) Hur uppfattade du det?
3b) Tror du att ni förstod varandra?
3c) Hur vet du att ni förstod varandra?

4) Har du några tankar kring hur ni använder föremål för att kommunicera?