The impact of goal focus, task type and group size on synchronous net-based collaborative learning discourses

H.-R. Pfister & M. Oehl
Institute of Experimental Industrial Psychology – LueneLab, University of Lüneburg, Lüneburg, Germany

Abstract
Net-based collaborative learning discourses often suffer from deficiencies such as lack of coherence and coordination. It is suggested that the provision of two functionalities, referencing and typing, which learners may optionally use to ground their contributions during a chat-based discourse, can improve collaborative learning. In particular, we examined if goal focus, type of task and group size affect learning outcomes and the use of these functionalities. A chat-based system, called a learning protocol, implements these functionalities and serves as a net-based collaborative learning environment. Results suggest that a learning protocol is more beneficial for knowledge-acquisition tasks than for problem-solving tasks, and that the use of supporting functionalities increases when goal focus is on the group rather than on the individual. Also, there is a tendency that learning outcomes improve as group size increases. We propose that learning protocols provide potentially valuable design features that can promote net-based collaborative learning.

Keywords
learning protocols, net-based learning, computer-supported collaborative learning, referencing, typing.

The present study examines the question of how participants of a distributed net-based learning group can be supported by a variant of scripted collaboration called a learning protocol. More specifically, we investigate whether synchronous learning discourses via text-based chat-tools can be improved by providing particular functionalities such as referencing and typing to enhance the coherence and meaningfulness of the ongoing discourse; and we ask under which conditions those functionalities are beneficial to improve learning outcomes.

A large body of research has demonstrated the beneficial effects of scripted collaboration on collaborative learning (Slavin 1996; O’Donnell & King 1999). This approach has been extended to net-based learning when groups of learners collaborate via the Internet (Jucks et al. 2003; Bromme et al. 2005; Fischer et al. 2007). However, in net-based collaborative learning it is unlikely that a productive discourse emerges spontaneously simply because sophisticated communication tools are available (Kreijns et al. 2003). A growing body of research demonstrates that typical deficits of net-based learning discourses are largely a result of problems of discourse coordination, such as incoherence of contributions and problems of maintaining thematically consistent threads of discussion (Herring 1999; Veerman et al. 2000; Lipponen et al. 2003; Soller 2004; Hewitt 2005). Participants are frequently not able to identify the relationships among individual contributions. That is, they fail to see the intended reference of a remark, they fail to identify connected
discourse segments, and they fail to join emerging threads – a phenomenon called *co-text loss* by Pimentel *et al.* (2003).

A variant of scripted collaboration, previously introduced as the *learning protocol approach* (Wessner *et al.* 1999; Pfister & Mühlpfordt 2002; Pfister *et al.* 2003; Pfister 2005), will be described in the next section. We then sketch the theory of grounding (Clark 1996) as a rationale linking learning protocol support to the objective of improving learning outcomes. Building on previous findings (Pfister & Mühlpfordt 2002; Mühlpfordt & Wessner 2005), the aim of this study is to further explore the specific conditions under which learning protocols might be helpful. The main section reports an experimental study testing the usage of referencing and typing in chat-based learning groups under variations of goal focus, task type and group size, and how they affect learning outcomes.

**The learning protocol approach**

Learning protocols are a variant of net-based collaboration scripts, using text-based chat as the primary communication device. Learning protocols provide two essential functionalities:

1. **Explicit referencing:** When submitting a contribution, the learner can indicate the contribution’s referent by means of a graphical external representation, that is, he or she can establish an explicit reference to a previous contribution, or to a segment of a contribution, or to some additional instructional material. The segment referred to is marked with the mouse, and by a mouse click, the relationship is established. This relationship is visualized by an arrow visible for all participants during the learning session. A learner can thus reread a discussion thread by tracking back the arrows. The referencing function has been shown to be able to promote learning outcomes (Pfister *et al.* 2003; Mühlpfordt & Wessner 2005; Stahl *et al.* 2006).

2. **Typing of contributions:** Each textual contribution, that is, a single chat message, can be classified and labelled by the contributing participant with respect to its communicative type, for example, as a question, explanation or critique. The learner deliberately attaches a label to her or his message, the label precedes the message on the screen and remains visible for all participants. Thus, the primary purpose of one’s contribution, indicating its speech act type (Searle 1975), is expressed unequivocally. Previous studies (Baker & Lund 1997; Soller 2004) have shown that typing can promote task focus and reflective communication.

Furthermore, a turn-taking rule (one-after-the-other) controls the sequencing of contributions among participants.

**Obligatory and optional learning protocols**

Dillenbourg & Jermann (2007; Dillenbourg 2002) have stressed the danger of over-scripting, that is, the risk of impairing collaboration by providing too strict and inflexible scripting rules. To avoid over-scripting, we distinguish between obligatory and optional learning protocols. With an *obligatory learning protocol*, the use of the supporting functionalities is mandatory, that is, learners are enforced by the system to type and to reference each single contribution. *Optional learning protocols*, in contrast, provide these functionalities via a menu, but learners are free to make use of them whenever they feel that it might be helpful. Hence, following Stahl (2007), optional learning protocols constitute situated resources rather than strictly enforced rules for collaboration.

In the present study, we investigated optional learning protocols. Two questions were addressed: first, we ask if and under what conditions learning with optional learning protocols improves learning outcomes. Second, we investigate if the use of referencing and typing actually is associated with improved learning outcomes.

**Related approaches**

Related approaches have been proposed by, among others, Baker & Lund (1997; de Vries *et al.* 2002), by Hron *et al.* (2000), by Rummel and Spada (2005), by Soller (2004), and by Weinberger *et al.* (2005; see also Beers *et al.* 2005; Kollar *et al.* 2006). All these approaches are based on the assumption that improving the coordination of communication will lead to enhanced knowledge acquisition.

According to Dillenbourg and Jermann (2007), learning protocols can be analysed with respect to the degree of coercion as ‘follow-me scripts’ when they are obligatory, or as ‘prompted scripts’ when they are optional; referencing and typing are then considered as affordances as opposed to enforced activities. Second, with
respect to granularity, learning protocols usually concern the micro-level of scripting, scaffolding the basic communicative units of a discourse while the structuring of the macro-level is left to the learners (Dillenbourg & Jermann 2007).

**Grounding with learning protocols**

The theory of grounding in communication, as introduced by Clark & Brennan (1991; Clark 1996), has provided an important framework to understand computer-mediated communication in general and net-based collaborative learning in particular (Brennan 1998; Baker et al. 1999; Pfister 2005; Dillenbourg & Traum 2006). Learning in groups can be conceived of as a collaborative process of constructing *common ground*, that is, a mutual understanding of the subject matter under discussion (Beers et al. 2005, 2007).

According to Clark (1996), messages need to be *grounded* by feedback signals from the recipients of messages in order to establish a shared understanding. Positive feedback that a message has been understood proved to be particularly important. If a contribution has been sufficiently grounded, it becomes part of the shared knowledge of the group. The social activity of overcoming a grounding problem, for example, by supplying an extra explanation, can trigger cognitive activities that lead to improved understanding (Dillenbourg & Traum 2006; Dillenbourg & Jermann 2007).

We suggest that referencing and typing of contributions as implemented in learning protocols can serve as indicators of grounding efforts. First, assigning a reference from the current contribution to a previous contribution conveys an instantly recognizable feedback signal. Even if the previous contribution has not been perfectly understood, referencing can initiate a process of clarification among the participants in order to reach a common understanding (Brennan 1998). According to Stahl (2006), referencing is at the core of creating what he calls the ‘discourse fabric’, that is, the shared meaning of a narrative text.

Second, assigning a type to a contribution, for example, declaring it explicitly as an explanation or as a critique, also serves as feedback to a previous message. Especially, typing can increase the occurrence of adjacency pairs, such as question-answer, claim-critique or assumption-justification pairs. Adjacency pairs are viewed as a basic mechanism of discourse coordination (Cahn & Brennan 1999; Stahl 2006).

In sum, we argue that explicit referencing and typing both instigate cognitive processes that can help to ground a discourse segment.

According to Clark’s (1996) *principle of least effort*, speakers try to minimize the effort they expend for grounding. Hence, learners will not expend that effort routinely, but will make use of optional referencing and typing functionalities only as much as they feel that it is necessary to ground their current contribution. We assume that the amount of effort participants are willing to expend will depend on contextual characteristics of the learning situation. Three kinds of contextual factors are assumed to be of particular relevance: goal focus, task type and group size. We consider each in turn.

**Contextual factors**

**Goal focus**

Goal focus as used in this study refers to the distinction between individual- and group-oriented performance. Group performance is frequently affected by the degree of positive interdependence among its members. One type of positive interdependence that turned out to be crucial for the performance of collaborating groups is goal interdependence (Johnson et al. 1989; Ortiz et al. 1996). Johnson et al. (1989) define positive goal interdependence as a situation when a participant can achieve his or her goal if and only if all other members of the group also achieve their goals. In collaborative learning discourses, each individual’s goal is to achieve a satisfying learning outcome; the goal focus normally is on the individual learner. If, however, participants are awarded according to the average learning performance of the entire group, the goal focus is on the group, and goal interdependence is established.

In the present study we distinguish, accordingly, between two kinds of goal focus. If the relevant goal is for each individual participant to perform as well as possible, goal interdependence will be negative, constituting an *individual goal focus*. If, in contrast, the relevant goal is that the group as a whole performs as well as possible, goal interdependence will be positive, constituting a *group goal focus*. With a group goal focus, each participant is accountable not only for her or his individual success, but also for the performance of the other participants a well.
A focus on the group is assumed to motivate learners to make increased use of the referencing and typing functionalities in order to establish a higher degree of shared knowledge. On the other hand, with an individual goal focus, the individual learner is rewarded if he or she outperforms the other members of the group; it might then even be detrimental if an individual participant helps others to understand the common subject matter.

Task type
Task type as used in this study refers to the distinction between learning by acquisition of factual knowledge and learning via problem solving. Problem-based learning is centred on the principle that having to solve a concrete problem leads to more reflective thinking and to deeper understanding than simply reading a presentation or an explanation of the relevant concepts and facts (Dochy et al. 2003).

In the present study, task type is manipulated via instruction: if the instruction explicitly asks to acquire knowledge about a domain, the focus of attention is directed towards information acquisition. If, in contrast, the instruction asks to solve a problem derived from the domain, the focus of attention is directed towards information application.

We assume that referencing and typing functionalities are especially appropriate to support a problem-based scenario, because learning discourses promote interactive processes of negotiating promising solutions (Johnson et al. 1989; Pfister et al. 1999). Specifically, we expect that, under a problem solving instruction, participants are more inclined to employ optional grounding functionalities, compared to participants under a knowledge-acquisition instruction.

In practice, goal focus and task type will most likely interact. Even if the task is to acquire knowledge, a focus on the group will increase the need for collaboration. And, conversely, even if the task is to solve a problem, an individual focus will reduce collaboration and grounding. Hence, the condition in which most grounding and learning should occur is when learners solve a problem and receive a group reward.

Group size
As is well known, group size influences group performance and group behaviour (Yetton & Bottger 1983). However, collaborative learning in net-based environments has mostly been studied only for dyads (Kollar et al. 2006). Severe coordination and communication problems in discourses will potentially arise with more than two participants (Clark & Brennan 1991). For example, the coordination of turn taking or the identification of the current speaker are minor issues when only two participants are involved, but considerably more difficult for more than two speakers. More generally, awareness problems will increase with an increasing number of participants (Carroll et al. 2003). Most important, it becomes especially effortful to relate incoming contributions to previous threads of the discussion. Preliminary findings indicate that group size is a determinant of learning outcomes in net-based learning groups (Pfister & Mühlpfordt 2002). In sum, we assume that participants will invest more grounding effort when group size increases.

Hypotheses
The objectives of the present study are threefold: First, we want to identify contextual conditions under which optional learning protocols are conducive for net-based learning discourses. Second, we want to investigate if the use of grounding activities is related to contextual factors. Third, we want to test if learning performance is related to the amount of grounding effort expended during learning. Thus, contextual factors serve as independent variables, and grounding effort as well as learning outcome serve as dependent variables.

We derive the following hypotheses:

1 An optional learning protocol is assumed to be more beneficial if the goal focus is on the group (group focus) as opposed to on the individual learner (individual focus) (Hypothesis 1a). Furthermore, learning outcomes are expected to be better if learners are involved in a problem-solving task compared to a knowledge-acquisition task (Hypothesis 1b). With respect to group size, learning protocols are expected to be more useful if group size increases (Hypothesis 1c).

2 It is hypothesized that more grounding activities as indicated by referencing and typing are instigated if the goal focus is on the group in contrast to an individual focus (Hypothesis 2a). Also, it is expected that more grounding effort will be expended when learners solve a problem in contrast to a knowledge
acquisition task (Hypothesis 2b). Additionally, we expect that more grounding effort will be observed in larger groups (Hypothesis 2c).

3 Learning outcomes are hypothesized to be positively correlated with the amount of grounding effort (referencing and typing) expended by learners during the learning discourse (Hypothesis 3).

Method
Participants
One hundred and forty students from different study programmes (psychology, social sciences and economic sciences) volunteered as participants and were randomly assigned to 40 learning groups, consisting of three or four members. Participants were paid 15 Euros for participation, plus an additional bonus depending on the experimental condition. As a result of technical (software failure) and organizational (exclusion of non-native speakers) dropouts, a sample of $N = 118$ participants (age: $M = 25.03$ years, $sd = 4.07$; gender: 70.10% female; years of study: $M = 2.55$ years, $sd = 1.66$) was included for further analyses.

Materials and procedure

The learning protocol environment
The experimental environment was based upon the CONCERT SUITE software, a platform for collaborative learning of distributed groups developed by the Fraunhofer Institute IPSI. The user interface was identical for all participants (Fig 1).

Referencing and typing functionalities
When it was a learner’s turn to make a contribution, he or she had the option to mark the segment to which the current contribution referred with the left mouse button. This referenced segment was then automatically connected with the learner’s current contribution, visualized by a black arrow originating at the current contribution and pointing to the selected segment. Additionally, participants were free to define the type of their contribution before sending it. Typing was possible on two levels. Participants could choose from five first-level types and, optionally, from several additional second-level subtypes: Question (with subtypes ‘. . . of a definition’, ‘. . . of a function’, ‘. . . of a system’, or ‘. . . of a reason’), Explanation (with subtypes ‘. . . about an assumption’ or ‘. . . about a prediction’), Comment (with subtypes ‘. . . as a justification’ or ‘. . . as a commentary’) and Critique (with subtypes ‘. . . as a complement’ or ‘. . . as an alternative’). According to the selected type and level (main type with or without subtype), a specific textual label such as ‘EXPLANATION – of a concept’ was inserted at the beginning of the learner’s message; the label was for all participants visible on the screen.

The learning domain
The learning domain was the catastrophe at the nuclear power plant at Chernobyl in 1986. An initial introductory text was compiled, including general information about the plant and the disaster, facts about the workings of the power plant, as well as special information about the events leading to the disaster. Any further information was to be contributed by the learners themselves, relying on their prior knowledge, or was to be constructed in a self-directed manner during the discourse. The initial text was visible on the screen during the entire learning session.

Independent variables
Task type was manipulated via instruction. The learning task was explained as either to learn as much about the disaster as possible (knowledge acquisition condition), or to try to solve a problem (problem solving condition). The concrete problem to solve was to arrange a predefined set of 13 critical events into the correct causal order that ultimately led to the disaster (e.g. ‘Power drops to 1%’, . . . ‘All pumps of the primary circuit are activated’, . . . ‘The emergency cooling system is activated’). The critical events were listed in the material pane of the interface (lower left pane in Fig 1) and numbered by lower-case letters, so that participants could refer to them. No special tool was provided to arrange or represent these events.

Goal focus was also manipulated via instruction. Participants were either informed that they could earn an additional amount of five Euros if they achieved the best individual test score within the group (individual goal focus condition), or that they would receive an additional amount of 5 Euros if all members of the group achieved at least a 75% score on the post-test (group focus condition).
goal focus condition). Note that this manipulation tries to change goal interdependence via reward interdependence. An individual reward should establish competition in so far as participants are motivated to maximize their own monetary rewards. A group reward should establish group interdependence in so far as participants are motivated to support and help other members in maximizing their monetary reward. A different way to establish goal interdependence would have been via task specialization, but Slavin (1996) points out that without group reward team specialization is usually not effective in improving achievement.

Group size was varied by randomly assembling three (triad) or four (quartet) participants into one learning group.

Procedure

The learners worked in groups of three or four participants, depending on group size condition, at standard PCs in isolated cubicles in a computer laboratory, simulating a distributed scenario; communication was only possible via the PC-network. First, a questionnaire about participants’ experience with computers, e-learning, chat and the Web was administered, followed by a pre-test about subjects’ knowledge concerning the Chernobyl disaster. Participants were then introduced to the purpose, rules and handling of the learning protocol in a detailed tutorial session. Then, the experimental session started and the participants were instructed depending on the experimental task type and goal focus condition. Then, the learning discourse
started, with one session taking 40 min. Immediately
after the learning session, tests to assess learning out-
comes were administered.

Measurements and design

A knowledge pre-test was administered to control for
possible a priori differences, consisting of one open
question which asked participants to produce a short
description of what they know about the Chernobyl
disaster. Each participant’s answer was rated by two
raters on a six-point scale from ‘knows nothing’ to
‘knows a lot’. Learning outcome was measured after
the learning session in two ways: (1) with a knowledge
test for factual knowledge, consisting of 15 four-option
multiple choice items, tapping a representative sample
of concepts and events from the domain; and (2) with
a problem-solving test, where participants were to
arrange a set of 13 critical events that lead to the disas-
ter in the correct temporal sequence. The multiple-
choice test yielded scores from 0 to 15. For the
problem-solving test, the rank correlation between a
participant’s generated event sequence and the correct
causal sequence of the 13 events was computed and
used as test score. The average number of contributions
per participant and session that involved either a refer-
cencing and/or a typing activity was computed as an
indicator of grounding effort.

A 2 × 2 × 2 three-factorial design was applied with
two levels for each factor: goal focus (group versus indi-
vidual focus), task type (knowledge acquisition versus
problem solving) and group size (three versus four
learners). All factors were treated as between subjects
factors, yielding eight independent experimental
conditions. As a result of experimental dropouts, the
design was slightly unbalanced. The total of 118 partici-
pants was distributed across 33 groups, yielding 14
triads and 19 quartets, 15 individual-focus and 18
group-focus groups, and 17 knowledge-acquisition and
16 problem-solving groups.

Results

No differences between conditions were found concern-
ing previous experiences with computers, nor concern-
ing prior knowledge about the Chernobyl disaster.

Generally, the referencing and typing functionalities
were used quite frequently in all conditions: 73% of all
messages involved at least one grounding activity, 58%
involved both referencing and typing, 10% relied only
on referencing and 5% only on typing. These high fre-
frequencies should not be taken for granted, because par-
ticipants were not enforced or encouraged in any way to
use these functionalities, although they might have felt
invited to use them as a result of mere availability. This
confirms, however, that an optional learning protocol is
a viable way to provide supporting discourse function-
ality as a learning resource that is accepted and used
frequently and voluntarily.

Because of the hierarchical structure of the data, we
employed a multi-level approach in the following analy-
ses (Hox 2002; De Wever et al. 2007), thus controlling
for the dependencies of learners within groups. We
applied a simple random intercept model; because
explanatory variables are associated only with the group
level, a random slope model would be pointless (Hox
2002). Although the small number of learners per group
(three or four) and the number of groups (33) limit the
accuracy of parameter estimates, the data permit to test
the fixed effects of interest with sufficient precision
(Maas & Hox 2004). Note that the small number of indi-
viduals within a learning group constitutes a defining
characteristic of the domain of small group research.

First, we report results concerning learning outcomes
as a function of experimental conditions (Hypothesis 1).
Then, findings concerning grounding effort are reported
(Hypothesis 2). Finally, we report results on the rela-
tion between grounding effort and learning outcome
(Hypothesis 3).

Learning outcome as a function of goal focus, task
type and group size

To assess learning outcome, the 15-item multiple-
choice knowledge test and the one-item measure of
problem-solving achievement were used. Cronbach’s
α of the multiple-choice test is rather low (0.451); how-
ever, heterogeneity is intended in order to incorpo-
rate the diversity of relevant concepts. We regard the
multiple-choice test as comprising a representative
sample of items from the population of knowledge
items concerning the Chernobyl domain. Mean item
difficulty (P = 0.64) is adequate, and the variation of the
overall score is fairly large (M = 9.63, sd = 2.16,
min = 4, max = 14). For the following statistical analy-
ses, both outcome scores were normalized to a range
from zero to 100, and a combined overall learning outcome score was computed as their unweighted sum. This overall score does not significantly deviate from a normal distribution according to a Shapiro-Wilk test \(P = 0.128\).

First, we tested if learning outcome as measured by the overall score is affected by manipulations of goal focus, task type and group size. The intra-class correlation ICC1 of 0.487 (Table 1) indicates substantial variation across learning groups. Group reliability as indicated by an ICC2 index of 0.769 indicates that learning groups are fairly homogeneous. A full multi-level model with task type, goal focus, and group size as level-two explanatory variables, including a random intercept term, yields a significantly improved fit compared to the Null model according to a log-likelihood ratio test \(P < 0.001\). The reduction of intercept variance from 489.43 to 343.48, thus accounting for 29.8% of the variation in the random intercept, is substantial (Table 1).

We find a significant fixed effect of task type \(P = 0.009\), indicating that the overall learning outcome is greater when the task is to acquire knowledge \((M = 131.80, \text{SD} = 26.70)\) than when the task is to solve a problem \((M = 107.00, \text{SD} = 30.50)\), contrary to the hypothesized direction (Hypothesis 1b). The interaction between task type and group size turns out to be marginally significant \(P = 0.073\), suggesting that the difference between task types is more pronounced for groups with three members (Fig 2). This pattern lends support to Hypothesis 1c that larger groups benefit more from a learning protocol. The most pronounced improvement in learning outcome is observed when problem-solving groups increase in size from three to four members.

Group size fails to reach significance \(P = 0.12\), even though groups with four members perform slightly

![Fig 2 Mean overall learning score as a function of group size and task type. Error bars indicate 95% confidence intervals.](image-url)
better ($M = 120.60$, $sd = 26.88$) than do groups with three members ($M = 117.80$, $sd = 37.88$), as was expected.

The manipulation of goal focus turns out to have no effect on learning outcomes at all, contrary to Hypothesis 1a.

In sum, it turns out that with an optional learning protocol, substantially improved learning outcomes are observed when the task is to acquire knowledge instead of solving a problem, and that groups of larger size tend to benefit slightly more when the task is to solve a problem. Including the amount of referencing and typing as covariates yields virtually identical results. This suggests that the effects of task type and group size do not necessarily depend on a particular learning protocol.

## Grounding effort as a function of goal focus, task type and group size

A grounding activity has been defined as the deliberate utilization of either referencing, typing, or the simultaneous use of both functionalities with respect to one’s current contribution. Two numerical indicators of grounding effort were computed: First, a value of 1 was assigned to each single contribution if the learner used the referencing function, otherwise, a value of 0; this constituted the referencing indicator. A second indicator was computed by assigning a value of 1 to a contribution if it was typed on level 1 (main type), a value of 2 if it was also typed on level 2 (including a sub-type), and a value of zero otherwise; this constituted the typing indicator. A joint index for each participant representing his or her average grounding effort was then created: for each participant the reference indicator and the typing indicator were each averaged across all contributions of this participant. Then, a weighted sum of the reference indicator and the typing indicator was computed with a weight of 0.5 for the typing indicator in order to compensate for its larger range (from 0 to 2). This joint index served as a proxy for the average grounding effort expended by a participant per learning session.

A multi-level random intercept model was applied with grounding effort as the dependent variable, learning group as a random effect, and task type, goal focus and group size as fixed-effect explanatory variables (Table 2). The intra-class correlation coefficient ICC1 of 0.336 indicates that one-third of the variance is to be attributed to the group level ($\sigma^2_u = 0.09$; $\sigma^2_e = 0.178$); also, group homogeneity is fairly large (ICC2 = 0.641).

A multi-level model including the full set of explanatory variables increases model fit significantly ($P = 0.048$), and reduces intercept variation by 26.7%. A significant fixed effect was obtained for goal focus ($P = 0.038$), indicating greater grounding effort under a group focus compared to an individual focus. Marginal

### Table 2. Model estimates for the multi-level analysis with grounding effort as dependent variable.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Null model</th>
<th>Full model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Estimate (SE)</td>
</tr>
<tr>
<td><strong>Fixed effects</strong>&lt;sup&gt;1&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1.218</td>
<td>0.955</td>
</tr>
<tr>
<td>Goal focus</td>
<td>0.464 (0.221)</td>
<td>0.038*</td>
</tr>
<tr>
<td>Group size</td>
<td>0.397 (0.216)</td>
<td>0.068</td>
</tr>
<tr>
<td>Goal focus × Group size</td>
<td>−0.469 (0.250)</td>
<td>0.063</td>
</tr>
<tr>
<td><strong>Random effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept $\sigma^2_u$</td>
<td>0.090</td>
<td>0.066</td>
</tr>
<tr>
<td>Residual $\sigma^2_e$</td>
<td>0.178</td>
<td>0.178</td>
</tr>
<tr>
<td>ICC1</td>
<td>0.336</td>
<td></td>
</tr>
<tr>
<td>ICC2</td>
<td>0.641</td>
<td></td>
</tr>
<tr>
<td><strong>Model fit</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta$: log Likelihood-Ratio</td>
<td>12.70</td>
<td>0.048*</td>
</tr>
<tr>
<td>Proportional reduction in intercept variance</td>
<td></td>
<td>0.267</td>
</tr>
</tbody>
</table>

<sup>1</sup> Only fixed effects with $P < 0.10$ are tabulated.

*, $P < 0.1$. 

© 2008 The Authors. Journal compilation © 2008 Blackwell Publishing Ltd
effects were found for group size ($P = 0.068$), indicating greater grounding effort in groups with four participants, as well as for the interaction of group size with goal focus ($P = 0.063$). Fig 3 reveals that the increase in grounding effort is especially pronounced under an individual goal focus when groups increase from three to four participants.

In sum, grounding effort is significantly greater when the goal focus is on the group ($M = 1.28$, $sd = 0.50$) compared to on the individual ($M = 1.17$, $sd = 0.53$), thus confirming Hypothesis 2a. Grounding effort tends to be slightly greater in groups with four participants ($M = 1.34$, $sd = 0.47$) than in groups with three participants ($M = 1.03$, $sd = 0.53$), providing some support for Hypothesis 2c. However, contrary to expectation, grounding effort is not affected by task type.

**Relations between grounding effort and learning performance**

We assume that increased grounding effort leads to better performance. As a straightforward test, we correlated the grounding effort variable with the three learning outcome measures (knowledge test, problem solving and combined overall score). We find a significant correlation of $r = 0.28$ ($t(116) = 3.142$, $P = 0.002$) only between grounding effort and knowledge test score: Participants who showed more grounding effort achieved higher scores on the knowledge test.

A multi-level analysis with knowledge test score as the dependent variable (ICC = 0.438), learning group as a random effect, and amount of referencing and amount of typing as two explanatory variables, reveals that it is only the amount of referencing ($P = 0.023$) which significantly increases learning outcome. Using specific types such as explanation or comment as predictors does not yield significant relationships with learning. Thus, Hypothesis 3 could only partially be confirmed, although the general relationship is as expected: more grounding effort promotes better learning outcomes.

As a more focused test, we carried out a series of mediation analyses to test for a mediating role of grounding effort. Taking group focus, task type and group size, respectively, as the independent variable, test score as the dependent variable, and grounding effort as the intervening mediator variable, it turned out that grounding effort significantly mediates the effect of group size on learning outcome (indirect effect = 0.468, Sobel’s $z = 2.018$, $P = 0.022$). We conclude that larger groups learn better if they expend more effort on grounding, that is, on referencing and typing.

**The content of referencing and typing**

A closer look at the grounding components reveals that the typing categories are used with different frequencies across conditions. Across all contributions, conjectures are applied most frequently (20.4%), followed by questions (16.4%), comments (10.2%), explanations (9.8%) and critiques (6.4%); 36.8% of all contributions are not typed. A test on independence shows that types are used differently depending on experimental conditions ($\chi^2(15) = 64.3$, $P < 0.001$): In particular, questions are used much more frequently in the knowledge acquisition conditions, and comments as well as explanations are used more frequently in the group focus conditions. Overall, typing occurs more often under a goal focus on the group than under an individual focus.

Also, referencing and typing mostly go together. 26.8% of all contributions are neither typed nor referenced, but 58% are typed as well as referenced. References can either point to other previous contributions, or to the introductory material made available on the screen. Contribution references are significantly more
often used under the group focus conditions, whereas material references are evenly distributed between group and individual focus conditions ($\chi^2(6) = 33.3$, $P < 0.001$).

These findings suggest that task type and goal focus not only influence overall grounding effort, but also determine which particular types are selected and which particular kind of referencing is used. A full-fledged content analysis is beyond the scope of this paper, but we present two examples to provide some anecdotal evidence on the differences between high- and low-grounding discourses. Discourse Fragment A (Table 3) lists a discourse sequence from the learning group with the highest grounding index, which is contrasted with a discourse sequence from the group which showed a minimum amount of grounding activity (Discourse Fragment B). Both fragments illustrate a phase during which learners struggle to get the sequence of critical events in the proper order (events are symbolized by lower-case letters). In Fragment A, all contributions are referenced as well as typed, whereas in Fragment B neither referencing nor typing occurs.

Fragment A reveals that most references refer to the previous contribution. Making this explicit might instigate the use of anaphoric references (‘That’s right’, learner 2; ‘That’s correct’, learner 3). It can also be seen that the actual content largely corresponds to the declared type, for example, a critique as alternative starts with an alternative proposal on how to deal with the events (‘Shouldn’t we first . . .’, learner 3). Two complement types then actually extend the proposed sequence in a step-wise manner. Fragment B contains

---

**Table 3.** Examples of discourse fragments from a group with high degree of grounding (Fragment A) and a group with low degree of grounding (Fragment B).

**Discourse Fragment A (degree of grounding: 1.92). Condition: problem solving, group focus, group size = 4:**

<table>
<thead>
<tr>
<th>Learner 2:</th>
<th>REFERENCE→material</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘EXPLANATION OF PURPOSE: That’s right, the points are not in the text I think we have to include them somehow. In my opinion first point (g) (j) (a) then (l) turning off safety measures. I’m not sure when (f) comes in play’</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Learner 3:</th>
<th>REFERENCE→previous contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘CRITIQUE AS ALTERNATIVE: Shouldn’t we first bring all points in the text in some sequence and then try to make an ordering, but well I think that after g comes e j and maybe d. That is correct, the points are not in the text.’</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Learner 4:</th>
<th>REFERENCE→previous contribution, →Material</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘CRITIQUE AS COMPLEMENT: I would also say it makes sense to first get the points from the text in the right order. I think the order is g, a, j, d, b, and then maybe m.’</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Learner 1:</th>
<th>REFERENCE→previous contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘CRITIQUE AS COMPLEMENT: sequence g-j-a-d-i-b-h-m; I must be somewhere at the beginning, if russians consider the safety of humankind is a question’</td>
<td></td>
</tr>
</tbody>
</table>

**Discourse Fragment B (degree of grounding: 0.27). Condition: problem solving, individual focus, group size = 3:**

| Learner 2: | ‘we shouldn’t think so long and write nothing. who say after 20 s knows nothing should click and hand over to the next one. what about if we try to go from top to bottom I think f comes later, -g -i -j?’ |

| Learner 1: | ‘yep faster – g j i l think , a – d – i l can’t figure out h’ |

| Learner 3: | ‘I’m not sure about i, could be. Then b. I think everybody should make just one step, otherwise it becomes confusing. I cannot so fast, ok?’ |

| Learner 2: | ‘after b e?’ |

| Learner 3: | ‘again one after the other g-j-a-d-i-b, now k I would say, but we have to built in h – between a and d?’ |
more meta-utterances (‘we shouldn’t think so long . . .’, learner 2; ‘again one after the other . . .’, learner 3) about how to organize the discourse. Also, participants tend to repeat the evolving sequence of events before they make a complementing proposal, which might serve as a kind of implicit reference.

These examples give a hint on possible differences between high- and low-grounding effort. In fact, learners who do not make use of the explicit functionalities might still invest as much effort, for example, by having to repeat much of what has been contributed before, or by meta-comments that coordinate the discourse. Although the overall structure seems not to be very different, there is a certain elegance when referencing and typing is used. For the creator of a contribution, the effort might be as great with as without explicit grounding. For the addressees, however, the effort to decode might be lower because of the explicit representation of references and types.

Discussion

We studied if and when learners use discourse supporting functions such as referencing and typing during a chat-based collaborative learning session. It was suggested that if learners put more effort into grounding this would lead to improved learning outcomes. A learning environment which makes referencing and typing functionalities available for optional use, called an optional learning protocol, was implemented and run with 33 groups of learners. Three context factors, which we hypothesized to systematically influence grounding and learning, were examined: goal focus, task type and group size.

In sum, it turned out that with respect to learning outcomes optional learning protocols are more conducive to support collaborative learning framed as a knowledge acquisition task than a task framed as problem solving. Goal focus showed no effect on learning outcome. Furthermore, we found a significant propensity of participants to apply more referencing and typing under a group focus compared to an individual focus. Task type, in contrast, did not influence referencing and typing activities. Generally, there was a slight tendency that learning protocols are more beneficial, in terms of learning outcome as well as in terms of grounding, for larger groups of four participants than for groups of three participants; it seems justified to assume that with larger groups, say of five or six participants, this effect might be more pronounced.

Finally, it could be shown that participants who use the referencing function more frequently are more likely to perform better in a final multiple choice knowledge test. In particular, it turned out that referencing plays a mediating role: when group size increases, participants are more likely to use the referencing function, and thereby to improve their learning outcomes.

Figure 4 provides an overview of the conceptual structure of the study and the pattern of results.

We proposed that learning protocols would be more conducive for problem solving than for knowledge

---

**Fig 4** Conceptual structure of grounding effort and learning performance, and major causal links of the contextual factors under study.
acquisition tasks, but just the converse turned out to be the case. In their meta-analysis Dochy et al. (2003) found a positive effect of problem-based learning on skill acquisition, but a small negative effect on knowledge acquisition. In the present study, primarily, knowledge gain was measured, and skill acquisition was neither aimed at nor assessed.

Furthermore, one of the shortcomings of our study is its brief duration, that is, a single learning session of only 45 min. The findings of Dochy et al. (2003) indicate that retention period is a significant moderator in problem-based learning: following problem-based learning, after a longer retention interval students remember more of what they learned in contrast to students learning in a more conventional setting. It is quite conceivable that learning by means of problem solving takes more time initially, but because of deeper processing, leads to better recollection in the long run.

A focus on the group increased the implementation of grounding mechanisms, but did not improve learning outcomes. Again, a possible explanation is the time scale investigated. Ortiz et al. (1996) compared achievement of groups with high positive goal interdependence and groups with no interdependence over a 5-week period. Interestingly, after 1 week achievement scores of groups with no interdependence were higher than the scores of high interdependence groups. After 5 weeks, however, high-interdependence groups significantly outperformed no-interdependence groups. As with problem solving, one might speculate that the effect of a focus on group performance might reveal itself only after a longer time interval, though the mechanisms of this process are not obvious (see also Dillenbourg & Traum 2006).

We found that participants actually make use of optional grounding functionalities to a large extent. This might partly be because of demand characteristics of the experimental session, and might be less pronounced in more natural settings. We also found that grounding effort is correlated with a modest gain in learning (Stahl et al. 2006). Note, however, that grounding as defined in this study is entirely formal and does not take the content of a message into account. A detailed content analysis of each single contribution might possibly highlight different aspects of collaborative learning and grounding (Beers et al. 2007), but is beyond the scope of this study.

Stahl (2006) also emphasizes the importance of references as conversational devices that constitute group cognition by connecting individual messages. Stahl’s distinction between elliptical, indexical and projective references, inferred from a detailed chat-log analysis, might be helpful to further clarify the significance of referencing on the group level. Likewise, Dillenbourg and Traum (2006) analyzed 18 pairs of participants of a MOO problem-solving task with respect to the ratio of acknowledged messages, which they take as an indicator of grounding. They found no relationship with problem-solving success, but a weak relationship with the efficiency of the discourse. Dyads with a high acknowledgement ratio showed less redundancy in their dialogues, hence higher efficiency.

Our results suggest that with an optional learning protocol and sufficient instruction, learners are quite willing to make use of the available functionality. One might conjecture that in the long run learners will adapt to the special conditions, constraints and possibilities of a learning protocol; or, more generally, to any kind of net-based learning environment, and employ the available functions in a natural way. Newlands et al. (2003) showed that learners adapt quickly to the constraints of text-based computer-mediated communication, and modify their communication behaviour accordingly. The notion of ‘fading’ suggests that it might even be an objective and a design feature of scripted collaboration to support a smooth transition from the external to an internal script (Carmien et al. 2007).

Instead of providing learning environments with fixed collaboration scripts, it might turn out to be more beneficial to make flexible environments available, which are equipped with a set of functionalities from which learners can select what seems helpful, and which learners or tutors might even employ in creative ways not anticipated by the designer (Dillenbourg & Tchounikine 2007; Haake & Pfister 2007). This view corresponds to the probabilistic view of design or the ‘designing for interaction’ approach as advanced by Strijbos et al. (2004). They showed that providing critical prerequisites for interaction and participation increases the likelihood of knowledge acquisition through collaborative learning. The provision of optional learning protocols, we suggest, is one design feature, among others, that might increase the usefulness of chat-based learning environments for collaborative net-based learning.
Acknowledgements

This research was supported by the German Research Council (Deutsche Forschungsgemeinschaft DFG) under research grant PF330/1-2 to Hans-Rüdiger Pfister as part of the special priority programme ‘Net-based Knowledge Communication in Groups’ (SPP 1081). We would like to thank Nicole Brandenburg and Werner Müller for their help in conducting the experiment. We also greatly appreciate the valuable comments of two anonymous reviewers.

Notes
1See http://www.ipsi.fraunhofer.de/concert and http://sourceforge.net/projects/concertchat/ for further developments. All other materials can be obtained from the first author on request.
2Statistical tests of parameter estimates are based on Markov Chain Monte Carlo (MCMC) simulation, which is recommended to obtain robust estimates especially for small samples (Hox 2002). Computations were carried out with the R Statistical Software (R Development Core Team 2007).

References


H.-R. Pfister & M. Oehl


