CAN RESEARCH ON PEER COLLABORATION IN MATHEMATICS
BE APPLIED TO EVERYDAY SCHOOL WORK?
Supervisors: Anders Nelson & Anders Persson
ABSTRACT

This paper integrates and applies established findings from previous research into peer collaboration to a realistic classroom situation in Swedish upper elementary school. The aim is to survey the research literature and to replicate some of the potentially beneficial effects of peer collaboration in an ‘ecologically valid’ setting, thus providing teachers with justifiable and readily adoptable techniques. The study investigated the effect of collaborative problem solving on students’ learning, where the conditions for collaboration were ‘optimised’ according to previous findings with regard to ability, gender, task characteristics, and collaboration strategy. Participants were 80 year 9 students (aged 15 years), who individually completed a pre- and post-test comprising moderately complex diagram interpretation tasks. During the experimental phase, students completed a similar task, either individually or collaboratively. Students who collaborated were assigned to mixed-gender pairs using a ‘weak-strong’ heuristic, based on pre-test results. Results indicated that lower-ability students collaborating with higher-ability peers improved from pre-test to post-test, while higher-ability students regressed significantly. Students working collaboratively did not perform significantly better than did students working alone. Discussion extends beyond these findings to implications of research on peer collaboration for teachers and students’ learning.
## CONTENTS

**INTRODUCTION** .......................................................................................................................... 5  
  Background ................................................................................................................................. 5  
  Aim of this paper .......................................................................................................................... 6  
**LITERATURE REVIEW** ............................................................................................................. 8  
  Piaget versus Vygotsky .................................................................................................................. 8  
  Previous research in summary ....................................................................................................... 8  
  Age ................................................................................................................................................ 9  
  Ability .......................................................................................................................................... 9  
  Motivation .................................................................................................................................... 10  
  Confidence .................................................................................................................................... 10  
  Gender .......................................................................................................................................... 10  
  Task characteristics ....................................................................................................................... 11  
  Collaboration strategy .................................................................................................................... 12  
  Social context ............................................................................................................................... 12  
  Incubation effects ........................................................................................................................ 13  
  Group-work versus individual learning and assessment .............................................................. 13  
**THEORETICAL FRAMEWORKS** .................................................................................................. 14  
  Positivism versus critical theory .................................................................................................. 14  
  Cognitivism ................................................................................................................................. 14  
**THE STUDY** ............................................................................................................................... 16  
**METHOD** .................................................................................................................................. 17  
  Participants ................................................................................................................................... 17  
  Materials ...................................................................................................................................... 17  
  Procedure ..................................................................................................................................... 17  
**RESULTS** .................................................................................................................................. 19  
**DISCUSSION** ............................................................................................................................... 20  
  Interpretation of results ................................................................................................................ 20  
  Design choices and ecological validity ........................................................................................ 20  
  Reliability, validity, and generalisability ...................................................................................... 23  
  Conclusion and further research .................................................................................................. 24  
**EXTENDED DISCUSSION** ......................................................................................................... 26  
  Ability and confidence .................................................................................................................. 26  
  Pre- to post-test change .............................................................................................................. 26  
  Gender ......................................................................................................................................... 27  
  Task characteristics ....................................................................................................................... 27  
  Collaboration strategy .................................................................................................................... 29  
  Incubation ..................................................................................................................................... 29  
  The role of talk ............................................................................................................................ 30  
  Interaction structure ...................................................................................................................... 32  
  Helping behaviour ...................................................................................................................... 32  
  Group-work versus individual assessment .................................................................................. 33  
  Interpersonal dynamics and the collaborative process ............................................................... 34  
  The status of the present study .................................................................................................... 35  
**CONCLUSION** ............................................................................................................................. 36  
  Theoretical frameworks in retrospect ...................................................................................... 36  
  The theories of educational psychology ...................................................................................... 36  
  Research on peer collaboration .................................................................................................. 37  
  Implications ................................................................................................................................. 38  
**ACKNOWLEDGEMENTS** ............................................................................................................ 39  
**REFERENCES** ............................................................................................................................ 40  
**APPENDICES** .............................................................................................................................. 44
INTRODUCTION

Background

Löwing (2006) has studied how Swedish mathematics teachers adopt constructivist ideals in their efforts to individualize each student’s learning experience within the non-differentiated Swedish school system. According to Löwing, most teachers cite Vygotskian influences, rather than Piagetian. Some teachers argue that, since students construct their own knowledge, they should be left to work on their own. However, this line of reasoning reveals a fundamental misconception of Vygotsky’s educational psychology, where the zone of proximal development consists of the learning possibilities available to a student, not working alone, but guided by a teacher or a more knowledgeable peer (Langford, 2005). Consequently, some teachers choose to let students work together in groups, in order for group members to help each other. Unfortunately, this, in itself, does not lead to improved learning for the individual student, as Löwing reports. One of the reasons is that students often perceive the goal of group-work as completing their assignments as fast as possible. Furthermore, it is not sufficient for some student to know something, in order for them to act as more knowledgeable peers. Students must also be able to offer helpful explanations – something even teachers struggle with at times.

Additionally, the idea of placing students in small groups to promote spontaneous helpful interactions does not seem to work even in cases where students work at their own pace, directed by a textbook, according to Löwing (2006). More often than not, students in such groupings soon diverge in terms of which material they have covered, rendering discussions meaningless. In fact, reliance on self-helping units may be detrimental to many students’ learning, as the following example illustrates. Löwing relates how a group of four students was given a two-page assignment. Two members of the group rapidly worked through the problems and wrote down the answers. When the teacher checked in on the group, she immediately received correct answers from those two students. The other two participated in the exchange, but generally gave incorrect answers. Despite this, the teacher soon moved on to another group. Afterwards, the two stronger group members resumed their rapid completion of the assignment, while the other two became passive, receiving no help from either teacher or peers.

Löwing (2006) suggests that much of the impetus for the introduction of various forms of group-work in Swedish schools in recent years originates from the tension between differentiation and individualization. Löwing defines individualization as efforts to adapt learning content and style to each student’s needs and abilities. Differentiation, on the other hand, Löwing takes to imply assigning students to groups according to some characteristic, such as age, gender, interests, or capabilities. This grouping may be of varying duration and groups may or may not be isolated from each other. Differentiation is a controversial issue, but regardless of opinions, the facts remain that:

- Differentiation (in various forms) is very common in mathematics teaching, both in Sweden and internationally.
- Differentiation is a way to organize education – not a form of individualization.

Viewed in this light, many of the efforts to individualize mathematics education espoused by the teachers in Löwing’s (2006) study borders on differentiation, based on work pace, rather than on individualization. In contrast, Löwing argues that a truly social-constructivist ap-
Aim of this paper

The aim of this paper is twofold: First, to survey the literature on collaborative learning in an effort to summarize current knowledge. Secondly, to apply some of these findings in a small-scale experiment in order to investigate their ecological validity, i.e., their immediate applicability to an everyday classroom situation. Implicit in this ecological approach is the supposition that most teachers have neither the time nor the opportunity to make detailed studies of research into this field. For that reason, findings – if they are to be of any real importance – need to be robust enough to apply in a heuristic fashion (C. Howe, personal communication, October 15, 2006).
LITERATURE REVIEW

Piaget versus Vygotsky

Most research into peer interaction is based on the theories of either Piaget or Vygotsky (Tudge, 1992). Piagetian scholars emphasise the need for differing perspectives among collaborating peers. These different viewpoints lead to socio-cognitive conflict, which in turn fosters fruitful arguments and discussion among equals (Garton, 2004). In theory, both (or all) partners stand to gain equally from such interactions. Vygotskian researchers, on the other hand, stress the importance of partners’ different levels of competence. Working together, partners form a shared understanding, primarily benefiting the child or student working with a more competent partner (Langford, 2005). Consequently, that partner may be either an adult or a more knowledgeable peer.

Previous research in summary

In general, research confirms that peer collaboration is a potentially beneficial form of education (cf. De Lisi & Golbeck, 1999). For instance, collaboration may engender active participation, and teach students to work together. According to Lazonder (2005), collaboration has proved successful in a variety of subject matters, such as mathematics, science, and language; and its benefits have been found to be relatively independent of group size, student age, and student ability.

However, some research (e.g. Dimant & Bearison, 1991; Kruger, 1993; Hogan & Tudge, 1999) has also questioned the assumption that peer collaboration is appropriate without exception. Some researchers (e.g., Tudge & Winterhoff, 1996) have even uncovered negative effects. Hogan and Tudge proposed that the value of peer collaboration might depend on the ability level and age of students and their partners, the kinds of tasks they work on, and their motivation. Phelps and Damon (1989) pointed to the character of the interaction itself as an important factor in collaborative learning.

Additionally, recent research (e.g., Howe, Tolmie, Greer, & Mackenzie, 1995) has found that the quality of collaboration improves when partners take different perspectives on the topic at hand, and how to approach it. Furthermore, researchers such as O’Donnell and King (1999) have suggested that, preferably, the topic itself should force groups of students to search for and make sense of new information; but also to necessitate discussion and consensus. Howe, McWilliam, and Cross (2005) advocates the need for opposing ideas among partners, while Webb (e.g., Webb, 2003b) endorses thoughtful and structured explanations and helping behaviour among group members.

How effective, then, is peer collaboration as a learning strategy? How should students be grouped? Which factors are crucial to the collaborative learning process? Moreover, which theoretical framework offers the best guidance? Reviewing the literature, cognitive benefits of peer collaboration seem to depend on a variety of factors, such as:

1. Age (e.g., Hogan & Tudge, 1999; Webb & Favier, 1999);
2. Ability (e.g., Forman & McPhail, 1993; Garton & Pratt, 2001; Harvey & Garton, 2003; Fawcett & Garton, 2005);
3. Motivation (e.g., Forman & Larreamendy-Joerns, 1995; Samaha & De Lisi, 2000; Ellison, Boykin, Tyler, & Dillihunt, 2005);
Confidence (e.g., Levin & Druyan, 1993; Tudge, Winterhoff, & Hogan, 1996);

Gender (e.g., Golbeck & Sinagra, 2000; Strough, Berg, & Meegan, 2001; Allard & Yates, 2001);

Task characteristics (e.g., Phelps & Damon, 1989; Kruger, 1993; Howe, Tolmie, Greer, & Mackenzie, 1995; King, 2002); and

Collaboration strategy (e.g., Ericsson & Simon, 1984) – including, among other things, the role of talk (e.g., Teasley, 1995; Saab, van Joolingen, & van Hout-Wolters, 2005); interaction structure (e.g., Webb & Palincsar, 1996; O’Donnell & King, 1999); and helping behaviour (e.g., Webb, Farivar, & Mastergeorge, 2002; Webb & Mastergeorge, 2003a, 2003b).

In addition, researchers have investigated the social context within which the collaboration takes place (e.g., Palincsar & Herrenkohl, 2002); short-term versus long-term effects and incubation times (e.g., Howe, McWilliam & Cross, 2005); and the relationship between groupwork and individual assessment (e.g., Webb, 1993, 1997; Lazonder, 2005). Furthermore, scholars like Tudge (1992, 2000), and Van Meeter and Stevens (2000) have examined the various theories that inform educational research and practice, and their implications.

**Age**

The question of age has mostly interested researchers working within a Piagetian paradigm. Even though no clear-cut boundaries can be determined, it has been suggested (Garton & Pratt, 2001) that, in general, children approximately from the age of eight years and upwards have achieved a level of collaboration that allows them to reap the potential benefits of groupwork. According to Dimant and Bearison (1991), once children enter into a formal operational phase, they communicate more effectively and are better able to take advantage of each other’s perspectives.

**Ability**

When it comes to relative competence within a small group or pair, many researchers (e.g., Forman & Larreamendy-Joerns, 1995; O’Donnell & King, 1999; Fawcett & Garton, 2005) have come to the conclusion that only students of lower ability who collaborate with higher ability peers show significant improvement. This was also the case in a study reported by Tudge (1992). However, Tudge’s study further indicated that students were equally likely to come away from the interaction actually understanding less than they did before.

Levin and Dryan (1993) discussed the so-called optimal distance hypothesis. According to this hypothesis, students who collaborate with partners of higher ability stand to gain the most from collaborative work, provided that the difference in ability is not too great. Optimal distance refers to a difference in ability between partners that is large enough to create learning opportunities for the lower-ability student, but not so large as to merely confuse the lower-ability student, or even induce the higher-ability student to denigrate his or her partner. This line of thinking is clearly situated within a Vygotskian framework. According to Levin and Dryan, the hypothesis has repeatedly been born out.

Other studies, inspired by Piagetian thinking (e.g., De Lisi & Golbeck, 1999), have shown that even though interaction with more advanced peers is more effective, collaboration be-
Peer collaboration in mathematics teaching

between partners of similar ability can be beneficial if it entails some measure of conflict. In a recent study by Webb & Zuniga (2002), students benefited from collaboration with partners of both similar ability and partners of either higher or lower ability. The researchers concluded that the character of the interaction itself had a greater impact on later performance than did the ability level of group members.

Tudge (1992) raises an important point, namely that collaborating partners’ assessment of each other’s ability level may differ drastically from objective assessments made by the researcher studying the effects of collaboration. This could mean, for instance, that in the face of conflicting opinions the group accepts an (objectively) less competent partner’s solution as being the correct one. This means, as Tudge states, that “arriving at a situation of shared meaning or inter-subjectivity could have either adverse or beneficial consequences, depending on whose initial understanding is accepted as correct” (p. 1366). Harvey and Garton (2003), furthermore, have found that peer status may be a more important influence on group interactions than actual competence.

Motivation
Motivation may influence the results of collaborative learning, as has been demonstrated by, e.g., Forman & Larreamendy-Jones (1995). Motivation depends on a number of factors, two of them being socio-economic and ethnic background (Samaha & De Lisi, 2000). For example, a current study in an urban high school in the U.S. (Ellison, Boykin, Tyler, & Dillihunt, 2005) showed that even though students in general preferred cooperation to individual learning, Caucasian students’ attitudes toward collaborative learning were more negative than were those of African American students.

Confidence
In Tudge’s (1992) study, the outcome of collaboration seemed to depend more on confidence, than actual competence. Tudge speculated that the oft-reported correlation between collaboration with a higher-ability partner and subsequent improvement in understanding might actually be a consequence of the fact that competence and confidence often co-vary, so that the more knowledgeable partner is also the most confident. In his study, however, the task was chosen so that students’ degrees of competence and confidence were unrelated, which may have accounted for the deterioration of some students’ understanding.

Gender
Regarding group composition, much attention has been paid to gender. Some spurious findings notwithstanding, a majority of scholars seem to agree that the presence of male group members is beneficial to collaboration (see, e.g., Strough, Berg & Meegan, 2001). For instance, Phelps and Damon (1989) studied gender influences on reasoning and found significant gender differences, in favour of boys, in both mathematics and spatial reasoning. Tudge (1992) also reported gender differences. Contrary to Phelps and Damon, he reported no initial differences in reasoning levels. However, he found that girls were more likely to deteriorate in spatial reasoning than were boys, following collaboration with partners of the same gender. Tudge attributed this difference to girls having a greater interest in creating positive relations than in disagreeing with one another. In addition to problem solving, Samaha & De Lisi (2000) also investigated gender influences on judgment in a collaborative setting. In their
study, all groups except those consisting only of female students showed significant improvement in judgments following collaboration.

Two recent studies shed some additional light on this apparent imbalance. Golbeck and Sinagra (2000) investigated the effects of collaboration in physics. They expected that the opportunity to explore a physics problem together would enhance students’ understanding of that problem. A constructivist perspective also suggested that such benefits would pertain equally to all students. However, during post-tests, male students demonstrated a significantly better understanding of the problem than did female students. The researchers surmised that this might be partially explained by the fact that male and female students approached and discussed the problem in different ways. Supporting Tudge’s (1992) reasoning, Golbeck and Sinagra reported that male students were more task-oriented, and exchanged more task-related information. Male students were also more likely to offer alternative solutions and ideas. Female students, on the other hand, focused more on the social functioning of the group itself, and made more efforts to promote a pleasant atmosphere.

In a sociological study of an Australian high school, Allard and Yates (2001) reported that students found friendships between individuals of the same gender more challenging than friendships between girls and boys. Girls, in particular, valued male companionship as a refuge from the demands of female friendships, in terms of relationship management. It would seem, then, that the presence of male group members has the potential to increase the quality of collaborative work itself, as well as the outcome of that work, by maintaining focus on the task at hand. On the other hand, rivalry in all-male groupings – although perhaps more task-oriented – may also detract from the work. In Allard and Yates’ study, boys often experienced a high degree of competition amongst themselves. Nevertheless, based on the above account, all-female groups appear less productive than all-male groups, in producing judgments as well as in solving problems.

Task characteristics

Phelps and Damon (1989) investigated peer collaboration in mathematics, and compared different types of tasks. They found that collaboration increased performance on tasks that required formal reasoning. Tasks that only involved remembering and copying, however, did not increase performance. The researchers concluded, “Peer collaboration is a good method for promoting basic conceptual development but not for fostering rote learning” (p. 644). Howe, Tolmie, Greer, and Mackenzie (1995) also came to the conclusion that collaborative tasks should involve theoretical argumentation as opposed to the mere learning of model examples. To ensure such high-level thinking, Howe et al. proposed that tasks should require collaborating peers to build and test their own models – empirically as well as theoretically – and that those models should comprise rule-based systems.

King (2002) expounds the issue of higher-level thinking. In her view, even ‘simpler’ tasks, such as recall and repetition of information, may well be relevant to some form of group work. For instance, students may help each other to rehearse homework, or check each other’s spelling. However, King argues, to promote students’ construction of new knowledge, more elaborate forms of peer collaboration, such as structured group discussions, should employ equally more complex tasks. Tasks that demand a higher level of cognitive processing, according to King, “involves making inferences, drawing conclusions, synthesizing ideas, generating hypotheses, comparing and contrasting, finding and articulating problems, analysing
and evaluating alternatives, monitoring thinking, and so on” (p. 34). Thus, ideally, the tasks chosen for peer collaboration should involve more than application of routine procedures.

**Collaboration strategy**

Research (e.g., Ericsson & Simon, 1984; Kruger, 1993; Ogden, 2000) has found that groups whose members engage in active debate are more likely to benefit from collaboration. Forman & McPhail (1993) suggested that students learn when they listen to each other’s explanations and question their accuracy. This notion was born out in a recent study (Fawcett & Garton, 2005), where only those students who reflected on their partner’s ideas and were required to explain themselves improved significantly. King (2002) also stressed the importance of explanations, argumentation, and discussions of strategy for successful group collaboration.

Research focusing on the way group members talk to each other (e.g., Tudge, Winterhoff, & Hogan, 1996; Saab, van Joolingen, & van Hout-Wolters, 2005) further illustrates that active debate is a prerequisite for productive group work, but that it is not sufficient. For example, Saab et al. found that successful problem solvers, in their discussions, stated explicit rules for their reasoning, actively searched for critique, and continually ensured that partners understood and reciprocated their thoughts. Several investigators have also suggested that productive group work requires partners to work together towards a common goal (e.g., Levin & Dryan, 1993; Forman & Larreamendy-Jones, 1995).

In summary, the interactions that take place between group members direct the cognitive activity of each member, and that activity in turn accounts for both individual learning and collective problem solving (see, e.g., Webb & Palincsar, 1996; O’Donnell & King, 1999). Yet, research (e.g., Webb & Favier, 1999) has shown that productive interaction patterns do not generally occur by themselves within groups. For this reason, teachers need to structure group work in a way that forces students to elaborate on the issues at hand. “In fact, unless the teacher intercedes with explicit guidance in how to interact, students working in groups appear to be more focused on finding the right answers than on learning” (Webb & Favier, 1999, p. 141).

Recently, Webb and colleagues (Webb, Nemer & Zuniga, 2002; Webb, Farivar & Mastergeorge, 2002; Webb & Mastergeorge, 2003a, 2003b) have turned their attention specifically to the way group members seek and give help during collaboration. Their research has shown, among other things, that the amount of learning an individual student does during group work is strongly related to that student’s willingness to elicit, receive, and build on help from other members. Webb and Mastergeorge (2003a), for instance, found that in order to benefit from other group members’ help, students need to formulate precise questions about how to solve a problem rather than just asking for the correct answer. Further, help-givers need to provide detailed explanations and make sure that they are understood.

**Social context**

In addition to task characteristics and interaction structure, Palincsar and Herrenkohl (2002) discuss the social context of collaborative learning. According to these researchers, it is important to foster a sense of shared responsibility between group members. Shared responsibility implies that all members work actively on the problem at hand, and further that they focus on the same aspects of the problem at the same time. Equally important, according to Palincsar and Herrenkohl, is that all group members share the commitment to find a common
ground (cf. Grice, 1989) on which to base discussions and understanding. This commitment will translate into sharing and merging of ideas during collaboration.

**Incubation effects**

Benefits of peer collaboration may not be evident straight away. Howe, McWilliam, and Cross (2005) have investigated delayed effects of group work, and suggest that they may be thought of as incubation effects. These authors argue that individual benefits of collaboration are usually seen as an immediate consequence of collective insights during collaboration, in line with Vygotskian theory. However, individual progress sometimes transcends that of the group as a whole. According to Howe et al., this often co-occurs with delayed progress.

One possible explanation of incubation effects is ‘priming’. The term denotes the tendency for an individual, during the time following failure to solve a specific problem, to have a heightened sensitivity to information that may be relevant to the solution to that problem. Based on their study, Howe et al. (2005) contend that collaborative problem solving, when unsuccessful, may prime some members to make productive use of subsequent events, which in turn leads to increased individual understanding at a later time. In the title to their study, the authors paraphrase Pasteur in suggesting that, as regards the delayed effects of collaboration, “chance favours the prepared mind.”

**Group-work versus individual learning and assessment**

It is often the case that collaborating groups perform better than do individuals working alone (see, e.g., Samaha & De Lisi, 2000; Ellison et al., 2005). However, this does not mean that group work always leads to better learning on the individual level, as has been pointed out by Fawcett and Garton (2005). It also raises the question of how accurately group performance reflects each member’s individual performance.

A case in point is a study by Webb (1993) that compared group assessments to individual assessments. In that study, group assessments tended to overestimate the competence and performance of individual members. From this, Webb concluded that individual competence should be measured by individual assessments.
THEORETICAL FRAMEWORKS

Positivism versus critical theory

Tudge (2000) points to fundamental epistemological and methodological differences between two major scientific schools of thought, both of which are applied to research on peer collaboration: positivism, on one hand, and critical theory and constructivism on the other. As Tudge asserts, “the world view that one adopts has enormous implications for one’s notion of reality, the type of theory that one finds appealing, the methods one uses, the way in which one analyses and interprets the data, and so on” (p. 99).

Positivist ontology holds that reality is a constant that can be explored and unanimously agreed upon. This implies a dualistic and objective epistemology, which suggests that individual researchers each do their bit in a collective uncovering of a single, objective, and external phenomenon. The methodology used in that uncovering is one of falsification since, although theoretically knowable, the true nature of things may well elude the fallible human researcher. The constant effort to transcend the role of participating subject translates into an experimental methodology searching to establish causes and effects and to test hypotheses – the researcher’s best guesses so far as to a correct understanding of (some aspect of) the world. The goals of positivism are to explain, predict, and control the environment.

Post-modernist or critical ontology, on the other hand, states that reality is relative and dependent on context. Critical theory and constructivist epistemology concedes that experiences are inevitably subjective and that any phenomenon is a product of the interactions between the investigator and that, which is investigated. Therefore, critical methodology is hermeneutical and dialectic in nature, acknowledging that descriptions of reality are negotiable and susceptible to perspective. The goal of research guided by critical theory is to arrive at an understanding that – for a particular audience – serves well in a given situation; rather than finding the ‘correct’ one.

Cognitivism

Although constructivism focuses on individual cognition, it is not analogous to the information-processing perspective of cognitive psychology. In principle, individual constructivism adheres to the tenets of cognitivism. However, educational psychologists often make an implicit connection between constructivism and hermeneutical approaches to the study of learning. Some scholars make explicit attempts to bring together constructivist approaches with a holistic view. Other approaches again, such as Marton’s adaptation of phenomenography (Marton & Booth, 2000), take non-dualism itself as their guiding principle, and consequently appear rather foreign to the essentially dualistic epistemology distinguishing cognitivism.

Cognitivism, or information processing theory, is a general model of the human information processing system, inspired by, but also applied to computers. The model is based on several more specific theories (see, e.g., Turing, 1950; Newell & Simon, 1976; Pylyshyn, 1986; Lakoff, 1987; Rumelhart, 1989; Harman, 1999; Hollan, Hutchins & Kirsh, 2000; Churchland, 2002). The general model is a system with a limited working memory. Knowledge is stored as inter-connected symbolic representations in long-term memory. The system receives sensory stimuli from the environment and transforms them through perception to internal symbols. Internal symbols are transformed into actions through motor processes. These two processes connect the system to its environment. Central to the cognitivist paradigm is the assumption
that all human behaviour can be explained in terms of internal mental representations, and processes which operate on these representations.¹

As Van Meeter and Stevens (2000) remark, the radically different philosophies applied to pedagogical research and educational psychology are not necessarily incompatible. Rather, they are – at best – complementary.

Although each focuses on a different aspect of the collaborative process, the theoretical premises do not contradict one another and, thus, do not prohibit simultaneous use of different theoretical frameworks in the same research study. Second, because each theory has a different focus, no single theory captures all that is important about peer collaboration.

(Van Meeter & Stevens, 2000, p. 120)

For instance, the dynamics and performance of groups, as opposed to individuals, may be successfully explored using a socio-cultural approach. Individual learning resulting from collaborative work, on the other hand, may be better understood through the lens of constructivism.

¹ This general characterisation includes sub-symbolic approaches, which do not require a direct correspondence between neural patterns and any high-level description of the behaviours to which they are associated.
THE STUDY

The information processing perspective of cognitivism heavily influences the study reported here. From an essentially positivistic point of departure, the aim of this study was to replicate some of the potentially beneficial effects of peer collaboration on students’ learning (as reported in the literature) by measuring quantitative change. Accordingly, two premises of vital importance were that individual cognitive gains result from (a) collaboration with a more competent partner, and (b) the active participation in such collaborations – none of which contradicts either Piagetian or Vygotskian theory.

Heeding the call of Van Meeter and Stevens (2000), the aim was further to study peer collaboration in an ‘ecologically valid’ setting, thus providing teachers with a justifiable and readily adoptable technique. Ecological validity was pursued in the choices made regarding the method of investigation as well as in the materials used, as indicated below. Additionally, the study tried to optimise the conditions for collaboration, with regard to individual gain. This implied selecting those conditions that have proven universally beneficial in previous research, while maintaining ecological validity, i.e., doing so under realistic circumstances. Those conditions included partner ability, gender, task characteristics, and collaboration strategy. Using the same experimental design utilized by many of the authors cited in this report (and described in more detail, below), the main hypothesis was:

1. Scores of lower ability students collaborating with peers of higher ability will improve from pre- to post-test more than will those of students working individually.

Secondary hypotheses were:

2. Pairs will score higher than will students working individually in the experimental phase.

3. Scores of higher ability students collaborating with peers of lower ability will not deteriorate more from pre- to post-test than will scores of students working individually.

The rationale behind the second hypothesis was the remark made by some scholars (e.g., Webb, 1993) about teachers’ reasons for using group-work. One might suspect that, on occasion, teachers employ group-work because the products of collaboration are usually of a higher quality than the average individual product. Further, one might speculate that, in some instances, the implicit assumption is that individual students’ learning is also of a correspondingly higher quality. This, however, is not always the case. The motivation behind the third hypothesis was that some findings (e.g., Tudge, 1992) have hinted at a tendency for regression in high-ability students, subsequent to collaboration. The phrasing of this hypothesis takes into account the fact that the test materials used were not identical. It also balances for possible effects of practice or fatigue.
METHOD

Participants

Participants were 80 year 9 students, from an upper elementary school outside a small semi-rural town (population 40,000) on the west coast of Sweden. Although no specific data were available, the school’s area of coverage represented a socio-economic stratum roughly characterized as middle-class. Permission from school principals and year 9 teachers, and student agreement were prerequisites to participation, in accordance with principles of research ethics (Swedish Research Council, 2002). The experiment was carried out in five different classes, consisting of between 15 and 20 students each (M 16.0, SD 2.5). Overall, there was a 96% participation rate (77 of 80 students completed the experiment), and no systematic subject mortality.

Students were assigned a rank order within their respective class, based on their score on a pre-test. Every third student from the ranked list was placed in a control group (23 students in total). Emulating a ‘weak-strong’ heuristic, remaining subjects were matched to form pairs in which the difference in ranking between members was the same across all pairs in a class (½ n, where n is the number of students in the class). This resulted in 27, predominantly mixed-gender, pairs (appendix A). Subsequent control measures confirmed that student capability (as measured by relative ranking) did not differ substantially between pairs (M 9.8, SD 2.7), although the mean ranks of student pairs in the experimental group were somewhat lower (i.e., indicating higher ability) than were those of students in the control group (M 11.4, SD 5.5).

Materials

The test materials used in pre-tests, experimental phase, and post-tests, were taken from recent national assessments. Tests focused on interpreting diagrams, statistics, and percentages (appendix B). Each of the three versions included 12 questions about a specific diagram, the answers to which were simple ‘yes’ or ‘no’ statements. These assessments are normally employed in a group setting at the end of year 9 to assess comprehension, clarity of communication, and participation. Diagrams, statistics, and percentages are generally not addressed in the year 9 mathematics curriculum, except in these national assessments (which had not yet taken place at the time of the experiment). The three versions used were selected to represent the same level of difficulty, and to be similar but yet different in content, in order to avoid practice effects while permitting comparison. Within each test, questions were of progressive difficulty, the last five questions giving students the opportunity to display extraordinary mathematical competence during discussion. Tests were administered in the same order to all students, regardless of group and condition.

Procedure

Students were told about the nature and purpose of the study (though in somewhat vague terms; see appendix C), and informed that their participation was welcome, but not mandatory; that they would remain anonymous in the analysis; and that results would not be disclosed (not even to their teacher) unless expressly permitted.

Students completed pre- and post-test components of the study individually in their regular place in the classroom, and were supervised by the experimenter. Also present in the classroom, or in close vicinity, was the regular teacher. The experimenter’s role was to instruct the
students, monitor session time and behaviour, and manually correct the tests. The experimenter played no active role during tests, either when children worked individually, or when they worked in pairs.

Students’ initial ability was assessed individually based on their pre-test scores. On average, students spent approximately 15 minutes completing the pre-test, and obtained a score of between 5 and 11 (M 8.3, SD 1.9). Students were ranked according to score, and then matched with a partner ranked approximately ½ n higher or lower (n being the number of students in the class), to form 27 weak-strong pairs (3 female-female; 5 male-male; and 19 mixed-gender). For every two students allocated to the experimental condition, one student was allocated to the control group (total 23).

Pairs in the experimental condition were then given precise instructions (appendix D). Instructions were based on the instructions given at the national assessment, which focuses (and grades) communication and collaboration skills, as well as comprehension. Thus, students were instructed to work together and to reach agreement; both members should be able to give the same answers and motivations to each of the questions. There was no time limit. The extent to which the instructions were obeyed was monitored but no experimenter intervention was required as there were no obvious deviations, nor any obvious variation between pairs.

The experimental task consisted of students working collaboratively in pairs (or individually in the control condition) to complete a task similar to that used in the pre-test. During this phase, students were free to choose where to work, provided they ensure some degree of privacy and a measure of peace and quiet. Most students (both individuals and pairs) relocated to nearby study rooms, while a few remained in the classroom. While students worked together, the experimenter informally observed their discussions and behaviour. On average, students working in pairs spent approximately 25 minutes completing the task, and obtained a score of between 5 and 11 (M 7.6, SD 1.7). Students in the control group spent, on average, 20 minutes completing the task, and obtained scores between 5 and 11 (M 7.0, SD 1.5).

Students’ ability to complete a similar task was reassessed individually in the classroom after an incubation period of 3-4 days following the experimental session to evaluate whether collaboration facilitated improved performance. On average, students spent approximately 15 minutes completing the post-test, and obtained a score of between 4 and 11 (M 6.7, SD 1.7).

In addition to test results, informal verbal reports from students and teachers were collected during and after each phase. Immediately following the post-test, students were debriefed and thanked for their participation.
RESULTS

Observation of the data revealed no significant skewness, kurtosis, or outliers; nor did the assumption of homogeneous variances seem to be seriously violated.

Table 1 shows the mean score (and the standard deviation) obtained for each condition in the pre-test, experimental task, and post-test; and the mean difference in score between pre- and post-tests.

<table>
<thead>
<tr>
<th>Group</th>
<th>Pre-test</th>
<th>Experiment</th>
<th>Post-test</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low ability</td>
<td>6.9 (1.3)</td>
<td>7.6 (1.7)</td>
<td>6.5 (1.7)</td>
<td>-0.5 (2.4)</td>
</tr>
<tr>
<td>Control</td>
<td>8.0 (1.7)</td>
<td>7.0 (1.5)</td>
<td>6.6 (1.7)</td>
<td>-1.4 (2.4)</td>
</tr>
<tr>
<td>High ability</td>
<td>10.0 (1.2)</td>
<td>7.6 (1.7)</td>
<td>7.3 (1.4)</td>
<td>-2.7 (1.8)</td>
</tr>
</tbody>
</table>

Hypothesis 1 and 3 concerned cognitive change. Since students’ initial ability (as measured by the pre-test) varied, the differences between pre- and post-test scores were used as the dependent variable, to allow comparison of the relative improvement in ability between the groups. A one-way ANOVA showed a significant main effect, $F(2, 71) = 10.2, p < .01$, with approximately 23% of the variability in scores attributable to collaboration. Tukey’s HSD post hoc test showed that, although lower ability students working with a higher ability peer improved from pre- to post-test (M -0.5, SD 2.4) compared with students working individually (M -1.4, SD 2.4), that improvement was not significant (with alpha set at .05). However, high-ability students regressed significantly (p < .05) from pre- to post-test (M -2.7, SD 1.8), compared with students working individually. Thus, the results support hypothesis 3 but not hypothesis 1, even though there was a strong trend in favour of the latter.

Students working together (M = 7.6, SD = 1.7) obtained, on average, a score 0.6 higher than those working independently (M = 7.0, SD = 1.5) in the experimental phase. A one-way ANOVA comparing the difference between the test scores showed no significant effect of collaboration. Thus, results did not provide support for hypothesis 2, although there was a trend in the hypothesised direction.

Informal observations and verbal reports did not provide sufficient information to warrant speculations about correlations between, e.g., individual and group characteristics, and outcomes.
DISCUSSION

Interpretation of results

The main hypothesis predicted that, following collaboration with a higher-ability peer, scores of lower ability students would improve more than would scores of students that had not collaborated. Although not significant, results revealed a strong trend along these lines. The ‘improvement’ actually turned out to be a deterioration in score, in all likelihood because different test materials were used. Students considered the post-test “harder” than the pre-test. However, the trend evident remains valid, since post-test scores of weaker collaborators did not deteriorate as much as either stronger collaborators or control group subjects.

The second hypothesis, stating that pairs would score higher than would students working individually, did not receive significant support in this study. However, results were not contradictory to this conjecture.

The third hypothesis held that, as opposed to the intriguing findings by, e.g., Tudge (1992), scores of higher ability students collaborating with peers of lower ability would not deteriorate from pre- to post-test more than would those of students working individually. The results did not support this hypothesis, offering compelling corroboration of these earlier results. Contrary to intuition, there did seem to be some detrimental effects of group work for higher ability students. Moreover, this tendency was stronger (2.7 - 1.4 = 1.3) than the relative ‘improvement’ of weaker students (1.4 - 0.5 = 0.9).

Design choices and ecological validity

The purpose of this study was to replicate beneficial effects of peer collaboration while maintaining ecological validity, i.e., optimising relevant factors under realistic circumstances.

The students in this study were 15 years old. Presumably, this means that they had reached an age where they, would be in a position to reap all potential benefits of group work, in terms of cognitive capabilities. As for selection, it may be argued that in general, studies concerning cognitive processes are less susceptible to selection biases compared to, e.g., studies that attempt to establish correlations involving socio-economic or cultural strata.

The assignment of students to pairs where the rank difference between members were ½ n, was motivated by previous research indicating the advantages of heterogeneous groups, especially for low-ability students. It also emulated a ‘weak-strong’ heuristic thought to reflect teachers’ reasoning and to operate within the constraints of everyday school contexts.

In terms of motivation, the assumption, in agreement with the literature, was that high motivation promotes both the processes and outcomes of collaboration (as well as of individual work). Efforts to motivate students centred on the choice of task, and the instructions given. The researcher reasoned that students would welcome the opportunity to practice for the national assessment in mathematics, scheduled to take place a few months after the experiment. Furthermore, the test material (and, by implication, the upcoming assessment) treated topics not addressed during year 9. In addition, the format of the impending assessment – a group setting where comprehension, clarity of communication, and participation is evaluated by an observer during student discussion – is rarely employed in mathematics teaching. This was believed to further motivate students to apply themselves. Informal observation indicated that
students’ motivation varied considerably, but revealed no apparent correlation between motivation and ability.

Informal observation and verbal reports gave no reason to question the interdependence of individual ability and confidence in the current circumstances. Despite the results being consistent with Tudge’s (1992) findings, explaining any improvement or regression subsequent to collaboration as the result of more knowledgeable peers yielding to less knowledgeable partners would seem remote.

Previous research strongly suggests that mixed-gender pairs promote the quality of both group-work and later individual achievement, with male-male pairings being almost equally favourable. In this study, 24 of 27 pairs were either mixed-gender or all male. To enforce mixed-gender or all-male pairings, however, would infringe upon the principle ecological validity: In applying a ‘weak-strong’ heuristic, it is unlikely that a teacher would be able to apply this restriction. On the other hand, it might be more important to encourage mixed-gender cooperation than to adhere strictly to a ‘weak-strong’ heuristic.

Tasks focused on interpreting diagrams, descriptive statistics and percentages (see appendix B), and were deemed to be authentic and relevant in at least three respects. First, they were recently used in national assessments (2002 and 2006). Second, the statistical information they present is adapted from recent surveys made by the Swedish Census Agency. Third, surveys chosen for these assessments deal with topics relevant to Swedish teenagers (e.g., the number of male and female teenage smokers during the last thirty years). To function in the national assessments, they are specifically designed to generate discussion and to invite alternative interpretations. Additionally, to cater to a wide range of students, the assessments, in their entirety, represent moderately complex tasks, containing questions of progressive difficulty.

The specific instructions accompanying these assessments (appendix E) reflect the collaborative strategy encouraged by the experimenter and state, in summary:

This assessment is carried out in groups of 3-4 students, seated together with their teacher.

You will receive a diagram, which you are allowed to study for a few minutes.

You will then be assigned a number of propositions concerning the diagram.

Students now take turns presenting their assigned propositions to the rest of the group. During presentation, describe to the group how you arrived at your conclusion as to whether the proposition is true or false, based on the diagram.

Following your presentation, group members are able to ask questions, make additions, and argue in favour of, or in opposition to, your conclusion.

Your efforts will be evaluated in three respects:

- Understanding: The degree to which you display understanding of the assignment, the topics it deals with, and how they are related.

- Language: The clarity and lucidity of your presentation, and the manner in which you utilize mathematical language.
Participation: The extent to which you participate in discussions, make a case for your ideas, and respond to others’ explanations.

Bear in mind that this is an opportunity to show your ability, during your own presentations as well as in discussing your classmates’ presentations, and in concluding discussions.

The experimenter read these instructions aloud to students prior to the experimental phase. In addition, students were told to treat the experiment as a dress rehearsal, and that the experimenter, together with their teacher, would listen in on their discussions. Further, students were instructed to work together and to reach agreement; both members of each pair should be able to give the same answers and motivations to every question, if later prompted.

In keeping with earlier findings, the promotion of active participation and a variety of – possibly conflicting – views were hypothesized to increase the individual gains of collaboration. Instructions expressly require students to verbalise their thought, as well as listen intently to, and respond to, other students’ reasoning, also found to be crucial to cognitive gains.

There is also a measure of (generative) structure to the question-and-answering process, although not nearly as detailed (empowering, even) as those proposed by King (2002) or Palincsar and Herrenkohl (2002). Finally, as is the case during actual national assessments, there were no time constraints. This, in itself, may not encourage deliberation. On the other hand, it does not impose a restriction on the collaborative process.

Effects of incubation were difficult to replicate in this study, due to time constraints. Post-tests were conducted 4-5 days subsequent to the experiment. This time span is clearly insignificant in light of the fact that Howe, Tolmie, Greer, and McKenzie (1995) reported considerable improvements up to eleven weeks following collaboration (but none after merely four weeks). It is doubtful, however, that even a period of several weeks would have made any difference in this case, since – as remarked previously – the year 9 curriculum does not address the mathematical concepts assessed by the tests. Thus, the likelihood that students would experience relevant events for which they were now primed seems small.

In addition to incubation effects, the literature strongly suggests that the individual benefits of collaborative work accrue with continual practice. While the students participating in this study were quite used to various forms of collaborative work, it is not clear whether that familiarity is transferable to this particular task, topic, and setting.

In keeping with Palincsar and Herrenkohl’s (2002) suggestions, this study took several measures to create a context that promoted productive group work: Instructions emphasized that pair members owned shared responsibility; and the pairing scheme ensured that expertise was distributed. Arguably, these students’ familiarity with group work ensured that the experimental setting seemed natural. Participation was voluntary, and pains were taken to avoid any form of implicit coercion on the part of classmates or teachers. As was the intention, students appreciated and seemed highly motivated by the chance to practice for the upcoming national assessments. In terms of physical context, students had a (limited) choice of work place in a well-known environment. Pair members were classmates of several years.

This study did not assess the extent to which group work reflected respective pair members’ individual competence, although the data gathered is amenable to further analysis.
Reliability, validity, and generalisability

Overall, the study attained ecological validity on a wide range of criteria. In order to generalize its results it is crucial to consider internal validity, as it rests on the absence of potential confounds. At least two issues deserve attention: the equivalence of experimental and control groups, and the reliability of the test materials.

The mean rank of students in the experimental group (M 9.8, SD 2.7) was somewhat lower than that of students in the control group (M 11.4, SD 5.5). This means that students working individually in the experimental phase demonstrated a slightly lower ability during pre-test than did those students who worked together. The implications for analysis are that – although not significant – the benefits of having previously collaborated with a more knowledgeable peer, or of collaboration in itself (corresponding to hypotheses 1 and 2, respectively), may be partially attributable to differences in student ability. On the other hand, the fact that high-ability students regressed significantly compared to control-group students during post-test lends further support to hypothesis 3.

The national assessments meet a number of reliability criteria. However, their use in this study entailed a trade-off between reliability and efficiency of analysis. By limiting students’ responses to ‘yes’ or ‘no’ statements, scores could be computed immediately following pre-test. This made it possible to conduct the major part of the experiment in a class during a single session. Consequently, the analysis did not take into account the communication- and process aspects that the national assessments are specifically designed to probe and which account for two thirds of the maximum score. Furthermore, the assumption was made that concept comprehension could be estimated by calculating the number of correct ‘yes’ and ‘no’ responses – as opposed to observing and listening to students’ presentations and discussions. This assumption was based on the experimenter’s previous experience of administering national assessments.

Even if these restrictions can be accepted, there are additional limitations inherent in the materials used, and the manner in which they were analysed. Each assessment consists of twelve statements about a specific diagram, the responses to which were, ultimately, simple ‘yes’ or ‘no’ expressions. Statistically, the chance of giving the correct response to any one statement is 50 %. There is an obvious risk for regression toward the mean. Moreover, although students’ scores ranged from 4 to 11, it is quite possible that the tests introduced subtle floor- and ceiling effects. This obvious threat to internal validity is further exacerbated by the fact that the different phases of the study each utilized a different test material. For instance, the seemingly robust pattern of regression in high-ability students may be partially attributable to post-test ceiling effects.

In order to exclude task and subject interactions from the analysis, much research into peer collaboration use ‘neutral’ and repeatable tasks, e.g., such as sorting blocks. In this study, the insistence on ecological validity implied using different authentic (school) tasks, which in turn bears with it the added complexity of such interactions in the analysis. Inter-reliability tests of the test materials were not feasible within the confines of this study. Their use in national assessments warrants some confidence in the assumption that they are interchangeable. However, there are no guarantees, especially since they were used and graded differently. In addi-

2 Higher scores resulted in lower ranks, the student ranked 1 being the top scorer.
tion, results of the 2006 national assessment (from which two of the three tests were taken) have yet to be publicised, which precludes comparison.

Even if an inter-reliability-test were feasible (based either on national results or on a pilot study), (groups of) students’ varying perceptions of the difficulty of the different tests would still be a factor, since individual schools interpret and schedule the goal-oriented Swedish mathematics curriculum differently. However, since this study compared subjects working alone to students working in groups, there was no need for the different versions to be identical, even though, to demonstrate individual learning on a specific topic, pre- and post-test would need to be interchangeable (A. F. Garton, personal communication, October 20, 2006). Pre-test and post-test were purported to represent the same degree of difficulty. The experimental phase utilized a version supposed to be slightly more difficult. According to the reasoning above, this does not affect any of the three hypotheses.

A final caveat about the materials is that the test used in the collaboration phase did not look exactly like the other two. It showed a bar diagram instead of a line diagram, although the information presented, and the questions asked, were similar (to a mathematics teacher, all three tests deal first and foremost with percentages). This difference in outward appearance may have made it harder for students to apply their learning to the post-test. In summary, the dependent variable is a rather coarse estimate of cognitive gain and learning. This reduces internal validity. On the other hand – to the extent that internal validity may be asserted – the procedure taken as a whole provide for a high degree of external validity which, in conjunction with the design choices made, in turn assures ecological validity.

Conclusion and further research

This study replicates previous findings in the sense that a majority of related research has substantiated two of the trends indicated by the current results. Firstly, students of lower ability benefit from working together with students of higher ability. Secondly, groups generally perform at a higher level than individuals working alone do. This seems to hold true even when no member of the group exceeds the individual working alone, in terms of ability. The possibility that high-ability students may be negatively influenced by working with lower-ability peers has yet to be explored fully, although similar findings have occasionally been reported in the literature.

These three phenomena may be well suited to exemplify the discussion of complementary philosophies in educational research. The fact that a ‘weak-strong’ heuristic seems to favour lesser-able students would fit especially well within a Vygotskian framework, stressing the benefits of working with a more capable partner and, in so doing, utilising a zone of proximal development to develop skills and knowledge that may be retained for later individual use.

The potential of collaboration to yield a combined outcome exceeding that of individual efforts appears to be a case of ‘one plus one equals three’, and seems rather intuitive. However – to continue with the metaphor – it might well be the case that in some instances, one plus one equals four, or even five. Conversely, it seems likely that in some situations, one plus one could actually equal less than one. This variability is a consequence of the complex interplay between partners, task, and context. Not least important is the collaborative process itself. Many of the variables responsible for predicting the outcome of group work are related to social context and to the socio-psychological interactions among participants. This would suggest that socio-cultural approaches to the study of collaborative learning are vital.
Perhaps the most interesting phenomenon exhibited in this study is the apparent decline in performance of high-ability students, subsequent to cooperation with a peer of lower ability. This could be explained within a Vygotskian framework, but might be better understood from a constructivist perspective. If cognitive conflict is a prerequisite for development, it would seem natural to assume that those conflicts predominantly arise in the lower-ability student when confronted with new information supplied by a more knowledgeable partner. Conversely, if the less knowledgeable student presents few or unconvincing suggestions, then no cognitive conflicts are induced in the higher-ability partner. This would account for the lack of improvement on the part of the stronger member of the pair.

To explain the decline in performance demonstrated by high-ability students, constructivist (or Vygotskian) theories may need to be supplemented by additional explanations. One possible account is offered later on in this text (on page 30).

As this study demonstrates, the potential power of collaborative learning is apparently robust enough to generate at least hints of effectiveness even when no great effort is put into detailed design of group characteristics, task, context, collaboration strategy, etc. However, as is equally clear from the results, applying peer collaboration in an unsophisticated manner with regard to those same features considerably lessens the optimal result of group work. Moreover, it may lead to undesirable side effects of equal or even greater magnitude than the desired benefits.

From the outset, this study was designed with the following question in mind: From an ecological perspective, given limited resources and a complex environment, which heuristics are (universally) applicable and which are not? The results partially verify the efficacy of some general rules of thumb, mainly concerning group composition and task characteristics. However, the intricacies of the collaboration process itself needs to be supported by similar guidelines, in order for peer collaboration to be successfully and safely applied in everyday teaching.

What seems clear in retrospect is that this and other studies of collaborative learning should be extended to include structured observation and interviews to probe individual and group attitudes and characteristics, and – perhaps most importantly – interaction patterns. It is also of interest to correlate such data to the actual test results. Regardless of methodology and epistemology, the focus of interest should rest firmly on the dynamics of the collaboration phase, and their impact on outcomes.
EXTENDED DISCUSSION

The literature review that opened this paper discussed a number of interesting and relatively consistent findings made by prominent scholars in the field of peer collaboration. To bring out the didactical implications of that research, this section relates those scholars’ discussions of their designs and results. Interspersed in the narrative are personal remarks and parallels to the current study. It is this author’s intention that the literature review, the study presented above, and this discussion, together comprise a toolbox that into which pedagogues may delve. It is up to the individual educator to locate those strategies that he or she deems credible, and to assess which didactical implications may be drawn from this corpus and applied to any specific educational context.

Ability and confidence

In Tudge’s (1992) study, students’ individual performances following collaboration deteriorated as often as they improved. To Tudge, this finding suggested that neither Piagetian nor Vygotskian theories suffice to explain how collaboration affects understanding. In Tudge’s study, lower-ability students did benefit from collaboration with a higher-ability peer, but only under certain conditions. Firstly, for learning to occur, the more knowledgeable partner had to explain relevant concepts at an appropriate level. Secondly – and perhaps more importantly – those explanations had to be accepted as credible by the less well-informed partner. As Tudge pointed out, this suggests that the relative confidence of collaborating partners may be as important as their actual knowledge.

Pre- to post-test change

Forman and Larreamendy-Joerns (1995), too, argued that the relationships between students – built up during a considerable time prior to collaboration – have a strong influence on the nature of their collaboration and on its outcomes. In their analysis, Forman and Larreamendy-Joerns also remarked that many students whose performances deteriorated during post-testing appeared to be solving a different problem during post-tests than during pre-tests. These researchers proposed that because students negotiate the goals of learning during collaboration, what is actually learned might differ considerably from that which is subsequently tested by the experimenter. This led them to the conclusion that “studies need to take account of the plurality of goals, interests, and expertise and of the process of negotiation in order to effectively evaluate peer collaboration” (p. 561).

Forman and Larreamendy-Joerns (1995) further reasoned that perhaps peer collaboration should not be evaluated primarily by looking at isolated individual performances, but rather by examining features of the group activities themselves. Webb, Nemer, and Zuniga (2002) came to a similar conclusion. In their study, high-ability students’ performances following collaboration with low-ability partners fell into two distinct categories. Some high-ability students’ performances improved after collaboration, as demonstrated by post-testing, while others deteriorated. Webb, Nemer et al. found that the strongest predictor of subsequent performances was the nature of collaboration, rather than the relative competence of group members. Some ‘weak-strong’ pairs worked well together, while others did not.

Webb, Nemer et al. (2002) suggested that conflicting results concerning the effects of collaboration for high- and low-ability students in various combinations might be accounted for by acknowledging that, in most cases, those effects have been measured by pre-determined tests,
administered to each student individually following collaboration. The quality of group work is partially determined by group composition, but the relative competence of members may not be the most important factor in that composition. Individual pre- to post-test change seems an unreliable measure of the appropriateness of collaborative learning – at least in cases where relative competence is the only basis for group composition.

In this study, collaborating students came from the same classroom. They had been classmates for at least nine years. Pairs were put together solely based on a ‘weak-strong’ heuristic. This was thought to enhance the learning opportunities for the weaker partner. It also served as a way to isolate the independent variable, and reduce the possibility that some students might feel more comfortable working together than others. No in-depth study was made of students’ interactions during collaboration. In retrospect, allowing students to form pairs freely, or assigning pairs based on knowledge of students’ personalities and ‘chemistry’ (as a teacher might well do) may be more conducive to successful group-work than enforcing differences in ability.

It is interesting to note that instructions for the national assessments simply state that teachers should divide students into “appropriate” groups. This author’s personal experience seems to indicate that both homogeneous and heterogeneous groups, with respect to ability, work equally well for the purpose of individual assessment within a group setting – a fact that may be explained by the teacher acting as a moderator. How this relates to posterior individual cognitive gains remains an open question.

**Gender**

Consistent findings indicate that groups containing at least one male member perform better than all-female groups. This was not tested in the current study but, rather, it was assumed based on previous reports. Possible explanations offered in the literature are that men are more task-oriented and willing to incite and engage in conflict (Van Meeter and Stevens, 2000) while women, in contrast, are more concerned with fostering a pleasant working climate and avoiding conflict (Golbeck & Sinagra, 2000).

These explanations are closely related to Piagetian thinking, in assuming that conflict is a pre-requisite of learning. The impact of differences in male and female behaviour could, however, also be encompassed by Vygotskian theory as, even in the absence of conflict, task-focus seems necessary for productive discussion. In addition, it is probably safe to say that gender composition affects not only group performances, but also subsequent individual performance.

**Task characteristics**

Samaha and De Lisi (2000) found an interesting difference between, on the one hand, giving correct answers (judgment), and on the other hand, motivating those answers (reasoning). In their study, groups performed at a higher level than individual members had done prior to collaboration. This applied both to the number of correct answers produced, and to the quality of explanations to those answers. The beneficial effects of group-work did not carry over to subsequent individual testing, in terms of the number of correct answers. However, following collaboration, individuals reasoned about problems in a more sophisticated fashion than they had done before.
Samaha and De Lisi (2000) speculated that the decline in the number of correct answers during post-testing could be partially explained by granting that the materials used in post-tests were somewhat harder than those used in the group setting – as may also have been the case in this study. However, their findings highlight an important distinction, with ramifications for both the study of peer learning (including the present one) and for the application of group work in classrooms.

The task used in the current study was selected based on, e.g., King’s (2002) assertion that (moderately) complex reasoning tasks require high-level cognitive processing. Reasoning was also encouraged by the instructions. However, results were measured solely in terms of judgment – the number of correct answers. As informal observations illustrated, the boundaries between reasoning and judgment are not distinct. Specifically, statements made in test materials were anticipated to yield unambiguous responses. This turned out to be questionable. Since no tests were made of students’ ability to reason about the problems, it may well be that some benefits of group work remained undetected.

In Samaha and De Lisi’s (2000) study, groups reasoned beyond the ability of individual members. However, as the discussion of the independence of ability and confidence has revealed, the consensus answer is not always the correct one. This may account for the decline in score from pre-test to post-test (and, indeed, also to the experimental phase) of some higher-ability students, evident in this study. Webb, Nemer et al. (2002) also used a task with (what was believed to be) unambiguously correct answers. They found that a further problem with using such tasks – both in research and in practice – is that they may have a negative influence on the collaborative process.

Webb, Nemer et al. (2002) focused specifically on high-ability students, working with less-able partners. In their study, some high-ability students benefited from collaboration, while others regressed (as did their partners). There was a clear pattern indicating that those high-ability students who contributed to the group in a positive way also benefited from collaboration (as did their partners). The opposite was also true. High-ability students whose results declined following collaboration were the same ones that had disrupted or neglected the preceding group session.

What is interesting in this context is the researchers’ suggestion that the nature of task itself might induce some high-ability students to act in ways that are not conducive to productive collaboration. Specifically, when there is only one correct solution or answer to a problem, students who believe they already know that solution or answer may simply ignore other group members. In light of this, Webb, Nemer et al. (2002) gave two recommendations as to the nature of tasks.

First, according to Webb, Nemer et al. (2002), successful completion of the task should require input from multiple perspectives, so that each group member is recognised as an important part of the process. This might be accomplished by asking groups to generate as many alternatives as possible, or to gather as much information as possible. In the case of the current study, it was believed that tasks were sufficiently complex to encourage groups to explore many different alternatives in order to find a solution.

The choice of task appeared to be appropriate in this respect, as observations confirmed that few, if any students – regardless of ability – had a clear idea of the correct answer initially. However, this does not mean that every group member actually contributed with suggestions.
Nor does it guarantee that every group member was interested in contributions from the partner.

Second, Webb, Nemer et al. (2002) suggest, tasks should be ‘open’ in the sense that no single answer is the only one admissible. In this respect, the current study was clearly deficient. For reasons of simplicity, brief and unambiguous answers were expressly desired in the design. Moreover, only the answers were analysed – not the reasoning behind them. This would have been a shortcoming even if answers were left open to interpretation.

Collaboration strategy

As noted above, Webb, Nemer et al. (2002) focused mainly on high-ability students, and their contributions to and profits from collaboration with peers of either equal or lower ability. Implicit in their analysis seems to be the assumption that – even though lower-ability students may well make crucial contributions – productive group work hinges on the presence of at least one high-ability student (indicating a Vygotskian inclination). This reasoning applies both to the group-work itself and its impact on individual members.

Regardless of task, Webb, Nemer et al. (2002) hypothesised that the contradictory results of groups and individuals in their study might have to do with students’ attitudes toward that task, and to the process of solving it. Specifically, they distinguished between two main types of individual orientation. A focus on performance, with the primary goal of displaying accomplishment, might prove detrimental to collaboration. On the other hand, a focus on mastering the concepts involved – on maximising learning – would presumably benefit every member of the group. It seems plausible that this applies to all students – not just those of high ability.

In this study, high-ability students performed at a consistently low level during post-tests. Following Webb, Nemer et al.’s (2002) logic, this might be partially due to a widespread focus on performance rather than learning. It is likely that the experimental setting itself contributed to fostering such attitudes, despite efforts to motivate students. A ‘natural’ setting, in which the teacher assigns group work as part of everyday work would possibly increase the likelihood of a focus on learning.

However, even during ‘normal’ conditions, it is important to emphasise learning, as opposed to performance. In addition, teachers need to encourage a positive and supportive atmosphere in groups, and to reinforce productive collaboration strategies – to the benefit of all students. Integral to a productive strategy is that each member shares and takes part of explanations, and that all contributions are taken seriously. Furthermore, a positive atmosphere entails a social dimension. As pointed out by Webb, Nemer et al. (2002), some individuals and groups spontaneously adopt a behavioural code that emphasises both serious cooperation and a welcoming environment. Others require teacher guidance to create such a climate.

Incubation

As pointed out by Howe et al. (2005), learning from collaboration may materialise as individual learning following collaboration. This marks a departure from a traditional Vygotskian analysis of peer learning, but may be encompassed within Piagetian theory. Howe et al., however, favour a third explanation, namely the semantic-network approach of cognitive psychology. Specifically, Howe et al. view subsequent cognitive changes in terms of mental schemas...
Schemas represent concepts by slots, and corresponding ‘slot fillers’. A slot may be empty, or it may take a value that is either correct or incorrect, depending on the context. The standard example is the concept of ‘table’, containing the slot ‘number of legs’. The default value is ‘four’, but in the presence of contradictory evidence, this value might be changed to ‘three or more’.

Howe et al. (2005) reason that incubation effects of peer learning may be construed as schema changes, made over time. In this view, the collaboration itself may present conflicting evidence, that is, reveal that a slot is blank or currently filled with a value that is not universally appropriate. It may not be until later, however, that a suitable replacement value is found. The incentive to search for new slot fillers that account for both older experiences and those made during collaboration is, in effect, what Howe et al. label ‘priming’. As regards the study reported here, it may be conjectured that both lower- and higher-ability students would obtain higher scores, were further post-tests to be performed at a later stage.

As remarked by Tudge (1992), knowledge and confidence do not necessarily correlate. In a group setting, this may translate into an inability of even highly knowledgeable students to articulate their understanding. It may also cause a group – including its more knowledgeable members – to adopt inferior theories or answers. At best, a supportive group climate may help members tease out the most successful ideas, and also to solidify those ideas – not least in the mind of the originator.

One of the teachers in this study suggested an alternative hypothesis that, although no distinctions were made between groups with respect to internal dynamics, might account for the almost universal regression of high-ability students during post-tests. According to this theory – which we might call ‘revision prompting’ – in articulating their understanding, higher-ability peers in heterogeneous pairs do indeed inform their group-mates in a way that these lower-ability students can put to good use in a subsequent post-test. On the other hand, for the high-ability students, the process of articulation disrupts previously constructed schemas, and brings them into question, rather than stabilising them – at least in the short term. The revision of schemas prompted by explicit articulation may, even when those articulations reflect a correct understanding, lead to temporary confusion and ambivalence in the immediate aftermath (i.e., during post-test).

This idea fits readily into related discussions of incubation. Although compatible with Tudge’s (1992) thesis, the latter would seem to indicate that, predicted by the high-ability students’ regression, all groups in the current study functioned poorly. However, this is cast in serious doubt by the fact that, over all, the lower-ability group members improved following collaboration. The apparent paradox may be resolved by conceding that the post-test exerted a ceiling effect, particularly for high-ability students.

The role of talk

Teasley (1995) investigated how learning relates to talk and how group work influences the way students talk about a problem. Focusing mainly on individual cognition, she started by establishing that talking while learning, even if only to one self, may help an individual to reinforce concepts. She also confirmed that talking with others in a group setting may further

---

3 Incidentally, it also fits very nicely with the intentions underlying some of the teaching- and assessment regimens used in teacher education (see Bengtsson, 2006).
facilitate individual learning. However, Teasley questioned conventional wisdom in explaining the added benefit of talking to others. Instead of emphasising the value of input from other group members (whether scaffolding or conflicting), she argued that one of the primary effects of talking to others is that it forces individuals to clarify themselves.

Teasley (1995) noted that the likelihood of students expressing themselves in a deliberate, lucid, and explanatory way when articulating their own thoughts on a problem (what Teasley referred to as ‘interpretive talk’) increased when they worked together. She argued that Grice’s (1989) maxims of conversation are inherently at work in collaborative learning, so that there is a mutual obligation, not only to make sense of what others are saying, but also to present a coherent argumentation – to the benefit of both parties. In Teasley’s words, “the answer to the question ‘Why are two heads better than one?’ may fundamentally rest on the interdependency of cognition and social relations in communication” (Teasley, 1995, p. 215).

In the current study, informal observations revealed notable variations in the way partners talked to each other across pairs. Variations included the amount of talk produced, the kind of utterances most commonly used, and the actual subject matter. These observations were not included in the analysis, but it seems safe to say that although group work may inspire interpretive talk, it does not guarantee it. On the other hand, Teasley’s study indicates the significance of thoughtful articulation in both individual and collaborative learning. This may turn out to be at least as important as, say, the respective ability and prior knowledge of partners.

Based on her findings, Teasley (1995) also raised the question of whether the benefit of thoughtful articulation actually requires that a conversational partner be present. She surmised that the anticipation of communication at a later stage might serve equally well. In discussing her study, Teasley speculated that the positive results of some individuals working alone might have resulted from the fact that subjects knew that they would eventually be accountable to the experimenter.

A similar assumption underlies the current study. It is reflected in the choice of national assessments as test materials, and in the instructions that accompanied them. As well as motivating students to participate in the study, the nature of the actual assessments would also lead students to anticipate future partners and supervising teachers.

If the sheer expectation of later review may prompt individuals working alone to adopt the same strategy of deliberate articulation which may account for much of the benefits of collaborative learning – does this mean that one could simply dispense of group work altogether? Probably not. A certain measure of ability and prior knowledge would seem necessary requirements, regardless of strategy. This, in turn, implies that the anticipation of accountability – verbal or not – is efficacious only for moderate-to high ability students. Nevertheless, it may well serve to motivate or ‘prime’ lesser able students.

Conversely, the importance of input from partners during collaboration – a factor that is often heralded as the most crucial aspect of collaborative learning – might apply primarily to students of lesser ability or knowledge. This said, the literature clearly indicates that – in favourable circumstances – external input and exposure to multiple perspectives are advantageous to all students, including the already well informed.
Interaction structure

King (2002), too, emphasises that deliberate articulation reinforces learning in the individual actually doing the explaining. Further, she argues that teachers can influence the way in which students think and talk about problems by providing them with a set of well-designed general-purpose questions. The thinking behind this is that “asking and answering thought-provoking questions forces students to actually think deeply about the material, integrating it with prior knowledge and constructing new knowledge” (p. 36). Accordingly, King has devised checklists designed to help students, working in groups, to reason about any given topic, in science as well as in social studies.

Apart from the instructions, this study made few explicit attempts to structure student interactions in detail, or to guide students in dealing with the material. Whereas the instructions established a desired format of interaction, they did not specify the content of discussions.

It could be argued that the assessments themselves, given their statement-and-response design, offered a certain measure of structure to students’ thinking about the diagrams. However, during informal observation it became evident that many students would have benefited greatly from considering, e.g., contra-factual examples, hypothetical arguments, and attempts to confirm conjectures (i.e., scaffolding along the lines of Socrates’ dialectic midwife). This suspicion was confirmed in debriefing sessions, during which several students demanded private sessions with the experimenter in order to resolve conflicts and uncertainties. It also appeared unlikely that generic questions could have provided a sufficient scaffold in the absence of a teacher, at least without prior practice.

Helping behaviour

Informally observed in this study, some students listened intently to their partner, took great pains to ensure that understanding was mutual, and actively worked to ensure a positive and constructive working climate. Conversely, some students appeared frustrated with their partners, either because they believed them unable to contribute to the collaborative effort, or because they did not value or understand their input.

In order to ensure productive and mutually beneficial collaboration, group members need to make good use of one another, especially in situations of confusion. As in the present study, observations by Webb and colleagues of students working in groups have demonstrated important differences in the way members seek, give, and apply help (Webb, Farivar et al., 2002; Webb & Mastergeorge, 2003a, 2003b).

Some students may be aware that they need help, but refrain from asking questions. Others may ask for answers, rather than for help with actually understanding a question or solution. When it comes to giving help, some students may simply neglect questions or signs of uncertainty from other group members, while others may not be able to offer satisfying explanations. Webb and Mastergeorge (2003a) observed that specific questions and answers are more likely to help both help-seekers and help-givers in attaining a greater shared understanding.

Previous findings indicating that several features of the collaborative context (such as behavioural norms and time constraints) influence the frequency and quality of mutual help, led Webb, Farivar, et al. (2002) to suggest that the teacher is instrumental in ensuring productive group work. More specifically, these researchers argued, teachers need to make sure that students understand and appreciate the importance of supportive behaviour. Further, teachers
must select tasks and structure work so that productive helping behaviour is encouraged. Lastly, collaborating groups should be kept under supervision. Yet, these researchers note:

> Despite the overarching goal of developing students’ understanding, our analyses of group work show that the behavior of individual students, of groups, and even of the teacher often served to undermine this goal. More often than not, students focused on obtaining or giving answers or procedures, and saw their task as finished if the group had correct work written down. The teacher did not always check for understanding, often focused on procedures instead of concepts, and seemed satisfied when a student or group gave the correct answer.

(Webb, Farivar, & Mastergeorge, 2002, p. 18)

These observations are reminiscent of Löwing’s (2006) analysis of group activities currently employed in Swedish schools.

**Group-work versus individual assessment**

Regarding the application of group activities in educational settings, Webb (1993, 1997) offers a wealth of insights. First, she establishes that the results of group assessments – such as the national assessment used in this study⁴ – are not a reliable indicator of individual knowledge or ability. Group assessments measure what individuals accomplish within groups, and that is another matter. However, as Webb (1993) points out, there may still be several good reasons for using group assessments: (1) to assess what an individual achieves as part of a group; (2) to measure an individual’s ability to function within a group; (3) to observe how an individual approaches and solves problems; (4) to allow for large-scale problems; and (5) to make efficient use of limited resources.

The national assessments are designed with all of these goals in mind. The group assessment is only one part of a comprehensive assessment regime, which also includes several individual tests. Taken together, these tests make up a solid foundation on which to grade students in mathematics. As far as the national assessments go, there is little risk of confusing what a student can accomplish individually with what the student achieves within a group. Nevertheless, such confusion may well arise during ordinary day-to-day schoolwork.

Webb (1997) also discusses the importance of defining clear goals for group work, and of designing group activities to match those goals. Two common – but radically different – goals of group work are individual learning and group output. Webb suggests that if the goal is learning, then group work should focus on active participation by all students, and structured helping behaviour. If, on the other hand, the goal of a particular group assignment is the pooled output of a group of students, such considerations may be less important, according to Webb. A similar argument is made concerning the assessment itself:

> Also important is to ensure that the design of the assessment and what will be evaluated are consistent with the goals of the assessment. We cannot expect students to focus on learning and engage in extended explanations of the material if the quality of the group’s solution or product is to be evaluated and if the group is under time pressure to complete the task. Conversely, we cannot expect students to concentrate fully on creat-

⁴ When deployed under normal circumstances, i.e., not as they were utilised in the current study.
ing a high quality group product if individual students are being held accountable for their own contributions to group collaboration.

(Webb, 1997, p. 212)

**Interpersonal dynamics and the collaborative process**

This extended discussion has mainly discussed the didactical implications of a number of recurring issues in research on collaborative learning. Forman and Larreamendy-Joems (1995) apply these same concerns to the research itself. Specifically, they argue that researchers interested in understanding what constitutes truly successful collaborative work must look beyond measures of average performance. Studies designed to establish causal links between characteristics of general designs and corresponding effects on some parameter of output, must be supplemented by studies that examine the particulars of individual instances of collaboration. This is also one of the major conclusions of the current study.

One example of such an examination is Kumpulainen and Kaartinen’s (2003) in-depth study of the inter-personal dynamics of collaborating pairs. Specifically, these researchers carried out a qualitative, case-based investigation of how features of the interaction between two students affected their work and its results.

Kumpulainen and Kaartinen (2003) managed to correlate a number of characteristics of specific interactive episodes to positive outcomes. Most salient of these were (1) equal participation; (2) joint negotiation; and (3) active conceptualisation. Such strategies were further found to be supported by interaction patterns characterised by the researchers as ‘reasoning’, ‘organising’, ‘agreeing’, and ‘arguing’. When looking at what made these interactions successful, Kumpulainen and Kaartinen found that creative pairs explored problems from many different angles, using a variety of approaches, such as construction, visualisation, testing, and measuring.

The researchers concluded:

Collaboration between peers seemed to be supported by reciprocal attempts to create a joint meaning when making problem solving visible via explanation and demonstration. The students’ appreciation of each other’s contribution to problem solving and their explicit communication of appreciation to their partner were also seen to support joint reasoning and to promote the students’ sense of being legitimate participants in their collaborative endeavour. From the viewpoint of mathematical reasoning, the interactions that supported collaborative problem solving were those that operated on the partner’s reasoning and those that coordinated perspectives and explanations into a global view to apply and test them. In addition, symmetric interaction was evidenced in representational interactions that paraphrased the partner’s reasoning and in interactions in which ideas and suggestions were challenged with opposing views.

(Kumpulainen & Kaartinen, 2003, p. 366)

Successful peer collaboration appears to be characterized by symmetric interaction, during which conceptualizing the situation is a joint effort, constructed with coherent problem-solving strategies and mathematical language.

(Kumpulainen & Kaartinen, 2003, p. 367)
The status of the present study

According to Kumpulainen and Kaartinen (2003), their study “demonstrates the power of micro-level process analyses in revealing the interpersonal dynamics of collaborative reasoning and shows how those dynamics mediate the learning opportunities in peer interactive dyads” (p. 333). Kumpulainen and Kaartinen stage their work within a purely qualitative paradigm and, accordingly, their findings offer no clues as to cause and effect. Although their study sheds some light on the intricate details of peer interaction and its outcomes, the authors hesitate to offer any concrete advice on how to implement group-work in educational settings – other than the noncommittal remark that “attention has to be paid to the situated dynamics of student interactions and collaboration to ensure that social activity supports collaborative reasoning and the construction of joint meaning” (p. 367).

The question underlying this paper is whether research into peer collaboration can be applied to everyday schoolwork. The natural corollary of this question is, if so, how? With respect to the case-based work of Kumpulainen and Kaartinen (2003), the answer to the first question is no. Consequently, the second question is left without a reply. How can educators create learning contexts conducive to successful peer interaction? In Kumpulainen and Kaartinen’s words: “Students’ communication and problem-solving strategies, as well as their command of mathematical language are likely to have a mediating role in shaping the nature of collaborative activity in peer dyads” (p. 368). On the surface, this might be interpreted as stating the tautology that in order for peer collaboration to be successful, students must be able to collaborate successfully. How, then, can teachers make sure that all students seize the opportunities presented by collaboration? All students – not just students who happen to ‘do it right’ from the start.

Kumpulainen and Kaartinen’s (2003) study clearly makes a welcome and necessary contribution to the understanding and application of group-work in educational settings. As has been pointed out previously in this paper, quantitative studies need to be complemented with selective, local, intimate, and in-depth investigations of individual characteristics, attitudes, experiences, processes, and outcomes. Such investigations are essential to constructing a comprehensive theory of peer learning – indeed, of all learning – regardless of whether the research is informed by constructivist thinking, socio-cultural approaches, (distributed) cognitive psychology, or some other perspective. In addition, such case-based studies may be quantitative or (as is more often the case) primarily qualitative.

However, to arrive at general models that can be generalized into universally applicable guidelines – to the extent that one believes this is possible – it is this author’s firm belief that cognitive psychology must ultimately be responsible for confirming or refuting any hypotheses about human interaction by means of quantitative research. Moreover, contrary to the view predominantly advocated in educational psychology – as in the social sciences as a whole – a cognitivist approach implies that any ‘true’ understanding will eventually be framed in positivistic terms.
CONCLUSION

**Theoretical frameworks in retrospect**

Recently, Tudge (2000) performed a meta-analysis of current research in educational psychology. In so doing, he identified three different theoretical perspectives and corresponding methodological ‘toolboxes’.

In keeping with the goals of this commentary, I must consider what would count as an analytic technique relevant to one theory or another. Goldhaber (2000) discussed this issue well, linking analysis of variance and regression analysis (useful for assessing the effects of independent variables) to theories stemming from a positivist or mechanistic worldview; clinical interviewing and pattern analyses to theories (such as Piaget’s) that fit within an organismic worldview; and ethnography and discourse analysis to contextualist theories such as Vygotsky’s.

(Tudge, 2000, p. 107)

As Tudge (2000) points out, the theories one adopts, and the toolboxes they offer, have a fundamental impact on which phenomena one chooses to study, and how one goes about it. Tudge urges educational researchers – regardless of their choice of theory – to carefully consider the tools of their trade.

Our field needs more theoretically driven research in which care is taken to match methods and analyses to the theoretical demands. In other words, I think that we need to spend more time undertaking what Winegar (1997) called the ‘methodological’ work of the research process. To do otherwise is to fall prey to an empiricist stance in which variations in findings, no matter how interesting, cannot be related to each other in a coherent whole. What makes findings interesting is not simply whether they support or call into question results reported by other researchers. A whole variety of reasons could be held responsible – from variations of task to differences of age of participants. What surely matters is that patterns of similarities or differences in findings can be related to each other in some coherent manner.

(Tudge, 2000, p. 111)

**The theories of educational psychology**

Although counted in the ranks of behavioural and social sciences, psychology as an academic discipline straddles – sometimes uncomfortably – between positivism and hermeneutics, as is definitely the case in educational psychology. However, the overwhelming majority of research in the discipline as a whole – fundamental as well as applied – is carried out within a firmly positivistic framework. Examples of the latter include psychometrics in traffic safety, social psychology in urban planning and crowd control, and clinical psychotherapy. As a case in point, the Swedish Board of Social Health recently regulated that only treatments based on cognitive behavioural therapies – as opposed to, e.g., psychodynamic therapy – will receive accreditation and subsidy. The Board motivated their decision with the need for scientific rigor in assessing the long-term effects of treatment.

These examples may be taken to imply that where human health and safety is concerned, society tends to rely on a mechanistic construction of the environment. More often than not, theories informing educational practices, on the other hand, tend to interpret the surrounding world in organismic terms. One may wonder, in a field that occupies close to a quarter of the
population and that lies at the very core of future generations’ health and safety, why this is so.

Although contemporary thinkers such as Marton shun dualism and reject constructivist accounts of the learning mind, their epistemology seems best understood – and applied – from a constructivist perspective (cf. Bengtsson, 2006). The author of this report sees no irresolvable philosophical problems with dualism. The criticism levelled at the information-processing paradigm since its inception by, e.g., Ryle (1948), Searle (1980), and most recently by Marton (Marton & Booth, 2000) has consistently been countered. Dualism is far from expelled from the core of cognitivism.

Dualism and methodological solipsism do not imply behaviourism, nor does quantitative accounts of behavioural change – to which cognitive psychology stands as a testament and archetypal example. Learning, even if described in positivistic terms, involves more than simply copying some behaviour or attitude. It is a process whereby information is internalised and connected to prior personal experiences.

Mechanistic and organismic world-views do not necessarily stand in opposition. However, there seems to be an implicit notion in much of educational psychology and philosophy that mechanistic models are to crude and barren to be able capture human thinking. Nevertheless, disregarding mechanistic models on aesthetic grounds is not scientific. In addition, relevant models of human behaviour are, per definition, general and generalisability can only be attained through quantitative measures.

Research on peer collaboration

As the literature review has shown, research on peer collaboration has yielded a number of converging indications that may be taken as guidelines by educators. These include, but are not limited to, the following:

- Students of relatively lower ability seem to benefit more from collaborative work than do students of higher ability. At least, the gains of lower ability students are more robust in the sense that their occurrence is less dependent on contextual factors such as individual and task characteristics, and collaboration strategy.
- Positive effects of group-work – collective as well as individual – are more pronounced when assignments are relatively complex, entailing non-routine tasks and open-ended questions, and requiring discussion from multiple perspectives. On the other hand, task should not exceed a moderate degree of complexity as this stifles creativity.
- Positive effects of collaborative work increase when such work is practiced continually, over extended periods. This is especially true for individual gains.
- Individual gains from any specific instance of collaborative problem solving may appear after some considerable incubation period, during which relevant incidents receive heightened attention and deeper elaboration, due to ‘priming’.
- The organization of group-work, including collaboration strategies and task characteristics, should promote active discussion and provoke (benign) cognitive conflicts. An essential factor for cognitive growth is the need (as experienced by the learner) to understand and be understood by partners.
• Gender composition of groups affects the results of collaboration. This is definitely true of the product of collaboration itself but by implication also applies to (long-term) individual gains. Gender effects vary across different tasks and contexts, but the most salient pattern is that the presence of at least one male group member is favourable in generating such conflicts that presumably lead to more advanced thinking about a given subject.

• Good collaboration practices do not appear by themselves. They must be designed and exercised. There is no one panacea; rather, collaboration strategies are highly context-sensitive. However, detailed programmes are available for a variety of such contexts.

Implications

The study reported in this paper was somewhat confounded, most notably by the choice and use of materials. This, in turn, was a result of limited time and resources. As it happens, limited resources are consistent with the ecological constraint imposed by the research question at hand: Can research be applied to everyday work? Even so, the study succeeded to a degree in replicating some of the previous research findings. It does appear that positive outcomes of collaborative work are robust enough to materialize even in the face of makeshift implementations. Carried out essentially as a pilot study, the most important lesson learned, however, is that despite this robustness, educators should not be content with applying any strategy, or with achieving any result of collaborative learning. There are myriad concerns to attend to, and they all have profound impact on learning.

Hence, the answer to the fundamental question of this paper – can research on peer collaboration in mathematics be applied to everyday schoolwork – seems to be yes, but it should be done with care and deliberation.

• Unqualified group-work, or differentiation, yields little or no results. Instead, it is fraught with risks and drawbacks.

• Well-structured group-work is a complex undertaking that requires a lot of effort on the part of both teachers and students.

• Even well thought-out collaborative learning experiences do not ensure specific or even beneficial outcomes.

• Moreover, the outcomes of group-work are not necessarily the ones planned by the educator. They may not even be obvious.

• Finally, one has to be very clear about the goals that are being set for collaborative learning.

As Webb (1997) puts it:

Group work may have an important place in assessment of achievement. But educators must carefully consider the purpose and design of the assessment, the goal of group collaboration in the assessment, and the interpretation of the scores that result.

(Webb, 1997, p. 212)
ACKNOWLEDGEMENTS

The author wishes to extend his heartfelt gratitude for the indispensable assistance in completing this report provided by:

Professor Timo Mäntylää, Department of Psychology, Umeå University.

Professor Alison F. Garton, School of Psychology, Edith Cowan University.

Professor Christine Howe, Department of Psychology, University of Strathclyde.

Professor Noreen M. Webb, Graduate School of Education, University of California.

Berit Ljunggren and her colleagues, together with their pupils, who participated in this study.
REFERENCES


APPENDICES
APPENDIX A1: Participants and raw data

Low ranking collaborators

<table>
<thead>
<tr>
<th>Class</th>
<th>Subject</th>
<th>Pre-test</th>
<th>Experiment</th>
<th>Post-test</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Score</td>
<td>Rank</td>
<td>Score</td>
<td>Score</td>
</tr>
<tr>
<td>E1</td>
<td>E101</td>
<td>9</td>
<td>14</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>E102</td>
<td>6</td>
<td>24</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>E103</td>
<td>8</td>
<td>18</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>E107</td>
<td>7</td>
<td>21</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>E108</td>
<td>8</td>
<td>17</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>E2</td>
<td>E201</td>
<td>7</td>
<td>20</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>E212</td>
<td>5</td>
<td>25</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>E215</td>
<td>8</td>
<td>15</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>E216</td>
<td>6</td>
<td>23</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>A</td>
<td>A01</td>
<td>7</td>
<td>12</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>A03</td>
<td>7</td>
<td>11</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>A04</td>
<td>7</td>
<td>10</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>A08</td>
<td>6</td>
<td>15</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>A10</td>
<td>6</td>
<td>14</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>A19</td>
<td>7</td>
<td>8</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>B</td>
<td>B03</td>
<td>9</td>
<td>8</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>B08</td>
<td>8</td>
<td>11</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>B09</td>
<td>8</td>
<td>10</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>B11</td>
<td>6</td>
<td>14</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>B15</td>
<td>6</td>
<td>13</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>C</td>
<td>C01</td>
<td>4</td>
<td>20</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C04</td>
<td>4</td>
<td>19</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>C07</td>
<td>7</td>
<td>17</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>C08</td>
<td>7</td>
<td>16</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>C11</td>
<td>8</td>
<td>14</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>C15</td>
<td>8</td>
<td>13</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>C17</td>
<td>8</td>
<td>11</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>27</td>
<td>27</td>
<td>24</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>187</td>
<td>204</td>
<td>157</td>
<td>-11</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>6,9</td>
<td>7,6</td>
<td>6,5</td>
<td>-0,5</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>1,3</td>
<td>1,7</td>
<td>1,7</td>
<td>2,4</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX A1: Participants and raw data (continued)

Control group

<table>
<thead>
<tr>
<th>Class</th>
<th>Subject</th>
<th>Pre-test</th>
<th>Experiment</th>
<th>Post-test</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Score</td>
<td>Rank</td>
<td>Score</td>
<td>Score</td>
</tr>
<tr>
<td>E1</td>
<td>E104</td>
<td>10</td>
<td>9</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>E1</td>
<td>E105</td>
<td>7</td>
<td>22</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>E1</td>
<td>E106</td>
<td>9</td>
<td></td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>E1</td>
<td>E112</td>
<td>9</td>
<td></td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>E1</td>
<td>E116</td>
<td>8</td>
<td>16</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>E2</td>
<td>E203</td>
<td>9</td>
<td>13</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>E2</td>
<td>E204</td>
<td>7</td>
<td>19</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>E2</td>
<td>E205</td>
<td>11</td>
<td>6</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>E2</td>
<td>E218</td>
<td>8</td>
<td></td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>E2</td>
<td>E219</td>
<td>11</td>
<td>3</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>A</td>
<td>A02</td>
<td>9</td>
<td>5</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>A</td>
<td>A05</td>
<td>5</td>
<td>16</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>A</td>
<td>A09</td>
<td>7</td>
<td>9</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>A</td>
<td>A14</td>
<td>6</td>
<td>13</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>B</td>
<td>B06</td>
<td>6</td>
<td>15</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>B</td>
<td>B10</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>B</td>
<td>B12</td>
<td>10</td>
<td>3</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>B</td>
<td>B14</td>
<td>9</td>
<td>6</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>B</td>
<td>B16</td>
<td>6</td>
<td>12</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>C10</td>
<td>9</td>
<td>6</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>C13</td>
<td>5</td>
<td>18</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>C16</td>
<td>8</td>
<td>12</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>C</td>
<td>C18</td>
<td>7</td>
<td>15</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>23</td>
<td></td>
<td></td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>184</td>
<td></td>
<td>153</td>
<td>151</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>8</td>
<td></td>
<td>7</td>
<td>6,6</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1,7</td>
<td></td>
<td>1,5</td>
<td>1,7</td>
</tr>
</tbody>
</table>
APPENDIX A1: Participants and raw data (continued)

High ranking collaborators

<table>
<thead>
<tr>
<th>Class</th>
<th>Subject</th>
<th>Pre-test</th>
<th>Experiment</th>
<th>Post-test</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Score</td>
<td>Rank</td>
<td>Score</td>
<td>Score</td>
</tr>
<tr>
<td>E1</td>
<td>E111</td>
<td>10</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>E114</td>
<td>11</td>
<td>7</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>E115</td>
<td>12</td>
<td>2</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>E2</td>
<td>E206</td>
<td>11</td>
<td>5</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>E208</td>
<td>9</td>
<td>12</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>E209</td>
<td>11</td>
<td>4</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>E213</td>
<td>9</td>
<td>11</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>E214</td>
<td>9</td>
<td>10</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>E217</td>
<td>12</td>
<td>1</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>A</td>
<td>A06</td>
<td>12</td>
<td>1</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>A07</td>
<td>8</td>
<td>7</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>A13</td>
<td>10</td>
<td>3</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>A16</td>
<td>8</td>
<td>6</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>A17</td>
<td>10</td>
<td>2</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>A18</td>
<td>9</td>
<td>4</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>B</td>
<td>B01</td>
<td>11</td>
<td>2</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>B02</td>
<td>10</td>
<td>5</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>B04</td>
<td>10</td>
<td>4</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>B05</td>
<td>11</td>
<td>1</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>B07</td>
<td>9</td>
<td>7</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>C</td>
<td>C03</td>
<td>9</td>
<td>10</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>C05</td>
<td>9</td>
<td>9</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>C06</td>
<td>9</td>
<td>8</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>C14</td>
<td>11</td>
<td>2</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>C19</td>
<td>11</td>
<td>1</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>C20</td>
<td>10</td>
<td>3</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>C21</td>
<td>9</td>
<td>5</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>27</td>
<td></td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>270</td>
<td></td>
<td>204</td>
<td>196</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>10</td>
<td></td>
<td>7,6</td>
<td>7,3</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1,2</td>
<td></td>
<td>1,7</td>
<td>1,4</td>
</tr>
</tbody>
</table>
APPENDIX A2: Calculations (according to Schweigert\textsuperscript{5})

ANOVA summary table for pre- to post-test difference (hypothesis 1 and 3).

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>2</td>
<td>92.4</td>
<td>46.2</td>
<td></td>
</tr>
<tr>
<td>Within groups</td>
<td>71</td>
<td>321.4</td>
<td>4.5</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>73</td>
<td>413.8</td>
<td></td>
<td>10.2</td>
</tr>
</tbody>
</table>

ANOVA summary table for individual work vs. collaboration (hypothesis 2).

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between groups</td>
<td>1</td>
<td>4.4</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>Within groups</td>
<td>47</td>
<td>121.6</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td>126.0</td>
<td></td>
<td>1.7</td>
</tr>
</tbody>
</table>

APPENDIX B1: Pre-test material

Version A – Materiell standard

![Diagram of material standard 1975-2003](image-url)
Diagrammet visar att

1. en femtedel av de tillfrågade personerna hade tillgång till husvagn år 2003.

2. cirka 2 % av de tillfrågade personerna hade inte tillgång till TV år 2003.


4. under 1984 och 1985 ingick inte husvagnar i undersökningen.


11. 70 % av befolkningen 1998 hade varken tillgång till båt eller husvagn.

Flickor vana rökare
Varje år dör 8 000 personer av rökning i Sverige. Det betyder att en person i timmen dör av tobak. 25 000 ungdomar börjar röka varje år. Det betyder att 70 unga börjar röka varje dag.

Andel av elever i årskurs 9 som röker varje eller nästan varje dag.

Källa: Världshälsoorganisationen (WHO), Centralförbundet för alkohol- och narkotikaupplysning (CAN).
Diagrammet visar att

1. andelen rökande flickor har varierat mellan 14 % och 20 %.

2. avståndet mellan två skalstreck på x-axeln är två år.

3. 1992 var andelen pojkar i årskurs 9 som inte rökte 84 %.

4. 1994 rökte en femtedel av alla flickor.

5. under en period på 80-talet rökte inga pojkar.


7. andelen rökande bland flickor respektive pojkar ökar samtidigt.


10. under 1994 rökte 32 % av eleverna i årskurs 9.


12. under perioden 1992–94 minskade andelen rökande pojkar med 25 %.
APPENDIX B3: Post-test materials
APPENDIX B3: Post-test materials

Diagrammet visar att

1. dubbelt så stor andel pojkar som flickor ”Spelade dataspel”.

2. de tre vanligaste sysselsättningarna bland pojkar var ”Såg på TV eller video”, ”Läste läxor”, ”Spelade dataspel”.

3. bok/tidningsläsning är hälften så vanligt som läxläsning.

4. en av fyra flickor gjorde ”Annat”.

5. nästan 30 % av alla barnen ”Tog hand om husdjur”.

6. 88 % av pojkarna hjälpte inte till i hushållet.

7. tre av 20 pojkar ”Höll på med dator”.

8. nästan 70 % av alla barnen ”Var med kompisar”.

9. ungefär hälften av barnen ”Spelade dataspel”.

10. mer än 100 % fler pojkar än flickor ”Spelade dataspel”.

11. minst 30 % av pojkarna både ”Såg på TV eller video” och ”Läste läxor”.

12. varje barn höll på med högst tre olika aktiviteter.
1. Uppgift och villkor

Jag kommer att undersöka om man blir bättre i matematik av att arbeta i grupp. Kanske borde grupparbete i matematik vara vanligare i skolan?

Uppgifterna är sådana som ni skulle kunna stöta på vid ett vanligt prov. Räkna med att varje uppgift tar 10-15 minuter att lösa.


2. Samtycke
Om du inte vill vara med i undersökningen, kan du gå ut i studiehallen och arbeta vidare själv. Om du stannar kvar i klassrummet, tolkar jag det så att du vill vara med.

3. Vad händer sen?
Det spelar ingen roll hur det går för er när ni löser de här uppgifterna – vare sig enskilt eller i par: mattelektionerna kommer att vara precis som vanligt i fortsättningen.

Om ni vill läsa min rapport när den är klar, kan ni be er lärare om hjälp att skaffa den. Rapporten kommer att vara klar efter jul. Om ni har åsikter om hur ni vill arbeta med matematik, eller om ni har funderingar på det som står i rapporten, kan ni diskutera detta med er lärare.
APPENDIX D: Instructions

Information till eleverna

Denna övning ingår som muntligt delprov i det nationella provet. Provet genomförs i grupper med 3–4 elever som sitter tillsammans med läraren runt ett bord.

- Var och en av er får ett papper med ett diagram eller en tabell som ni under några minuter får studera och sätta er in i. Sedan kommer ni att få ett papper med en rad påståenden som handlar om diagrammet eller tabellen. Din lärare talar om i vilken ordning ni ska redovisa.


- När alla redovisat sina påståenden diskutrar ni i gruppen några frågor som läraren ger er.

- Dina insatser under det muntliga delprovet bedöms ur tre aspekter nämligen förståelse, språk och delaktighet.

  I vilken grad du visar att du förstått uppgiften, de begrepp som ingår och sambanden mellan dessa.

  Hur klar och tydlig din redovisning är och hur väl du använder det matematiska språket.

  I vilken grad du deltar i diskussionen, kan argumentera för dina idéer och ge respons på andras förklaringar.

Tänk på att det är ett tillfälle att visa vad du kan både vid din egen redovisning, i diskussionen efter kamraternas redovisningar och i den avslutande diskussionen. Dina insatser vid detta delprov sammanställs och ger ett antal g- och vg-poäng och du kan även visa MVG-kvaliteter. Resultatet på det muntliga delprovet räknas samman med övriga delprovsresultat.