Since May 1998 Scandinavian group researchers and social psychologists have met bi-annually for what has to be called the GRASP conference. GRASP originally stood for "Group as Paradox". The first five conferences were held in Linköping, Lund, Stockholm and Skövde and, again, Linköping.

The sixth conference was organized at Lund University May 15-16, 2008 under the auspices of the network Organisational and Social Applications of Psychology (POST) at the Department of Psychology, Lund University, in cooperation with the School of Social Work, Lund University, and Forum for field research in organisations and groups (FOG), Linköping University.

Generous assistance was given by the the School of Social Work and the Swedish Council for Working Life and Social Research (FAS). This year’s theme was “Dynamics within and outside the Lab” and a special emphasis was put on inter group processes. 60 researchers and students from Sweden and Norway took part and listened to twenty-seven presentations and two key-notes by Susan Wheelan, Ph.D. Eleven contributions have been chosen to represent the conference in this volume.
GRASP is an interdisciplinary conference, which aims to provide a platform for researchers, practitioners, and graduate students from the Nordic countries within the fields of psychology, sociology, behavioural science and social work to share, exchange, learn, and develop preliminary results, new concepts, ideas, principles, and methodologies, aiming to bridge the gaps between paradigms, encourage interdisciplinary collaborations and advance and deepen our understanding of group and social psychology.

GRASP 2008 was the sixth Nordic conference in a series that seeks to develop a better understanding of group and social psychology. This biannual conference celebrated its tenth year 2008 since the first conference in Linköping 1998. The following conferences were held in Lund 2000, Stockholm 2002, Skövde 2004, and again in Linköping 2006. The focus of the earlier conferences and their proceedings (See list below) have been Small group studies; Studies of group and change; The group as a paradox; Building our theories better and Interaction on the Edge.

The 2008 conference was held in Lund with 60 attending participants and hosted by the network Organisational and Social Applications of Psychology (POST) at the Department of Psychology, Lund University, in cooperation with the School of Social Work, Lund University, and Forum for group and organisation research (FOG), Linköping University.

The main focus of this year’s conference was on how proponents of different research approaches can cooperate and cross-fertilize. Specifically, we were interested to bring together researchers working in the two main traditions of group and social psychology: the experimentalists and the naturalists. We did this out of our firm conviction that a research forum encompassing different methods of data gathering and differing traditions in design may fruitfully inspire researchers in the field.

27 papers were accepted and two key note speeches were given by Susan Wheelan, Ph.D.: Researching Work Groups in Natural Settings and Helping Work Groups Be More Effective: A Research-based Approach.

It may be of interest to see which methods were represented in the contributions. A summary is reproduced in Table 1.
Table 1

*Paper categories, number of experimental studies and data collection methods*

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>Exp</th>
<th>Quest</th>
<th>Obs</th>
<th>Int</th>
<th>F-gr</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attribution, social influence</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Justice</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Destructive group processes</td>
<td>2</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross cultural aspects</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interdisciplinary teams</td>
<td>4</td>
<td></td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interaction in habilitation teams</td>
<td>3</td>
<td></td>
<td>3</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group work in schools</td>
<td>3</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organization, learning, identity</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sum:</td>
<td>27</td>
<td>4</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>3</td>
<td>8</td>
</tr>
</tbody>
</table>

*Exp = Experimental study; Quest = Questionnaire(s); Obs = Observation; Int = Interview(s); F-gr = Focus group(s); Other = E.g. theoretical study, literature review, case study*

From Table 1 it can be seen that only four out of the 27 studies were experimental in nature, so the expected meeting between the “experimentalists” and the “naturalists” did not occur as expected. One reason for this may well be that e.g. researchers working in the area of social cognition, which is mainly experimental in nature, prefer conferences with a narrower focus.

Looking at designs and methods, however, it is possible to conclude that the majority of the contributors do naturalistic research and that the dominating methods are questionnaires, observation and interviews.

Of the 27 presentations given, 16 were invited to contribute to this volume, but some had already submitted to journals or been published elsewhere. Consequently, these proceedings bring you 11 papers. We hope that they, together with abstracts of all papers, will give a fair picture of on-going naturalistic research in group and social psychology in Sweden today.

We thank all who generously contributed to the sixth GRASP conference and look forward to the seventh, which is planned to be held at Göteborg University in May 2010.

*Stefan Jern and Johan Näslund*
PROCESSES MEDIATING MAJORITY AND MINORITY HERD INFLUENCES ON PREDICTIONS OF SIMULATED STOCK PRICES

Maria Andersson, Ted Martin Hedesström & Tommy Gårling

Abstract
We investigate the degree to which participants’ predictions of a simulated stock price are influenced by predictions made by a herd of five other (fictitious) participants when they consist of a majority or minority. The results of Experiment 1 showed that the participants followed a majority herd independently of whether its predictions were accurate or random. In Experiment 2, the majority influence was reduced by requesting participants to focus their attention on the accuracy of the others’ predictions. In Experiment 3, a minority herd was found to have an influence only when its predictions were accurate and the participants focused their attention on the accuracy of the others’ predictions. It is suggested that in an uncertain prediction task heuristic processing has a larger role than it has been ascribed in previous research on informational social influence, and that under these circumstances a minority influence is not associated with systematic processing.

Keywords: Social influence, heuristic vs. systematic processing, financial judgments

In stock markets information about price trends and others’ actions are important sources of information. On the basis of this information market actors make predictions of future stock prices (Andreassen, 1988; Svedsäter, Karlsson, & Gårling, in press). Our focus in this paper is how such predictions are influenced by others’ predictions.

Much research in financial economics has examined investors’ propensity to take the same action as other investors. When making the same decisions because of direct influences or imitation, this is referred to as herding (for a review, see Hirshleifer & Teoh, 2003). However, it may also be the result of “clustering of actions” as a consequence of indirect influences (Drehmann, Oechssler, & Roider, 2005; Sias, 2004), including reacting upon common knowledge (Grinblatt, Titman, & Wermers, 1995), following the same fads (Sias, 2004), or having common investment styles (Wermers, 2000). An important challenge to empirical studies is therefore to distinguish herding from clustering of actions. Since in actual markets the bases for investors’ decision making are seldom disclosed, it becomes difficult to identify the sources of information that influence the decisions. An experimental approach that accomplishes this is therefore required.

Anderson (2001), Anderson and Holt (1997), and Celen and Kariv (2004) report experiments in which participants make choices sequentially based on
private information and information about the choices made by others preceding them. In general it is found that participants disregard their private information and instead imitate the others. This form of herding is referred to as information cascades (Bikhchandani, Hirshleifer, & Welch, 1992). The aim of experiments demonstrating information cascades is to assess the degree to which herding is rational (e.g., Drehmann et al., 2005). In the present research, a more general aim is to understand the processes that may account for herding in stock markets whether or not it is rational.

Since herding in stock markets is a form of social influence, an understanding of the phenomenon may increase by applying theories of social influence. Such theories may however also need to be further developed. In the following we first briefly review empirical findings, then theories of social influence which are potential explanations of herding in stock markets. We finally report three experiments examining how the size of a herd and the accuracy of the herd’s predictions influence predictions of a fictitious stock price.

An important distinction is made between normative and informative social influence (Deutsch & Gerard, 1955). In the former case the motive is to conform to others due to external social pressure or internalized norm systems, whereas in the latter case the motive is to acquire useful information from others. Although both types of social influence exist in stock markets (Shiller, 2000), our focus is on informative social influence. Such influences are likely to dominate when participants make individual decisions knowing that their decisions will not be disclosed to others.

Herding refers to that individuals act in a similar way because they are influenced by others, but the number of people in the herd acting similarly may vary. A main finding in research on social influence is that a group of others who are in agreement tend to be more influential, and that their influence increases with the size of the group (e.g., Bond, 2005; Bond & Smith, 1996). One reason for such an influence is that the judgments of a group are perceived to be more accurate than judgments by individuals (“wisdom of the crowd,” Surowiecki, 2004). In a stock market this occurs when the prevailing consensus forecast influences investors independently of its accuracy in predicting subsequent stock prices.

The moderating effect of group size has been conceptualized as a majority vs. minority influence. A general conclusion is that a majority has a stronger influence than a minority (e.g., Bond, 2005). Previous research has also shown that the size of a majority (e.g., Hodges & Geyer, 2006; Lascu, Bearden, & Rose, 1995), as well as of a minority (Arbuthnot & Wayner, 1982), increases its influence. It has likewise been found (e.g., Arbuthnot & Wayner, 1982; Wood, Lundgren, Ouellette, Busceme, & Blackstone, 1994) that consistency increases the influence of both minorities and majorities.
Two processes have theoretically been posited to mediate influences from others, comparisons with others and validation of these comparisons (Wood et al., 1994). According to Mugny and Perez (1991), comparisons involve identification with the others and results in influences without deliberation, whereas validation assesses the others’ arguments and results in influences after deliberation. It is assumed that comparison and validation underlie both majority and minority influences. An opposite position is maintained in Moscovici's (1985) dual-process theory of conformity and conversion, according to which people comply with the majority without thoroughly reflecting because they wish to belong to the majority (conformity). Since people are unwilling to be identified with deviant groups, minorities are in contrast incapable of eliciting a comparison process. However, a minority may trigger a validation process leading to that their arguments are considered in detail. This may result in a changed private opinion (conversion) even though the majority's opinion may still be officially proclaimed.

A related conceptualization of minority and majority influences connects the processes of comparison and validation to heuristic and systematic processing (Martin, Martin, Smith, & Hewstone, 2007; Moskowitz & Chaiken, 2001). In heuristic processing influences are triggered by some cue in the environment (signaling status or size) or are the result of the belief that “the majority is always right”, referred to as the use of a “consensus implies correctness heuristic” (Martin, Gardikiotis, & Hewstone, 2002), henceforth labeled the consensus heuristic). In systematic processing, which entails careful evaluation of arguments and interrelated information, influences occur if people are persuaded by the others. However, previous research has not settled the issue of whether majority influence is mediated by heuristic processing and minority influence by systematic processing. Systematic processing has been demonstrated in response to messages provided by a majority (Mackie, 1987) as well as a minority (Martin, Hewstone, & Martin, 2003), and under some conditions messages are processed heuristically whether provided by a majority or a minority (Martin & Hewstone, 2003). In a recent study, Bohner, Dykerna-Engblade, Tindale, and Meisenhelder (2008) crossed size (majority vs. minority) with message strength (high vs. ambiguous vs. low) and with framing of the others as either similar to the participants (socio-relational framing) or as more knowledgeable (accuracy framing). In socio-relational framing systematic processing (referred to as message processing) occurred for arguments stated by a minority but not by a majority, regardless of message quality. In accuracy framing systematic processing occurred in minority conditions when message quality was high and low, but in majority conditions when message quality was ambiguous.

A common paradigm in previous social-influence research (e.g., Erb, Bohner, Rank, & Einwiller, 2002; Martin et al., 2002; Martin et al., 2003;
Martin, Hewstone, & Martin, 2007) is to assess participants’ attitudes after they are exposed to different messages that are delayed in time. An initial message is endorsed by a group of others, and then measurements are made of how much attitude certainty is influenced by a subsequent counter-message arguing the opposite position. This task implies that participants are presented with clear information about the group’s position. Since the focus of the present research is on social influence in stock markets, the experiments examine the different task of predicting stock prices. In this task participants themselves infer which of the others constitute the majority or minority. In order to do this, perceptions of consistency are likely to be the result of observations over time of agreements between actors’ predictions. In support of this, Andersson, Hedesström, and Gärling (2008) demonstrated the role of agreement (correlation) over time among other investors’ predictions for herding to occur.

Furthermore, in a stock market actors make predictions of stock prices based on uncertain information. This uncertainty amplifies the role of informative social influences. In line with Moscovici (1985) and the findings of Bohner et al. (2008), it may then be hypothesized that the use of the consensus heuristic accounts for the influence of others’ predictions when the herd is a majority. If a minority elicits systematic processing, it may be expected that in order to have an influence, a minority needs to be accurate. In contrast, a majority would have an influence whether it makes accurate or random predictions. In support of this hypothesis, Andersson, Hedesström, and Gärling (2008) observed an influence from the majority despite that its predictions were random, suggesting that participants used the consensus heuristic. Furthermore, a financial reward for following a majority or minority herd led to an influence from the majority but no influence from the minority. These results suggest that following a majority is a strong motive. In the present experiments the attention is shifted from financial motives to accuracy motives. By varying the accuracy in the herd’s predictions, the experiments investigate whether a majority will be followed both when it makes accurate and when it makes random predictions, whereas a minority will be followed only when it makes accurate predictions.

**Overview of Experiments**

The present three experiments aim at simulating predictions of price movements in a stock market. The participants are informed that they participate in a multi-trial experiment where each trial represents a trading day. Their task on each trial (day) is to make predictions of the price of a fictitious stock on the next trial (day). The stock price varies both systematically (referred to as *price trend*) and unsystematically (referred to as *price errors*) across trials. On each trial the participants receive information about the current stock price (referred to as *price cue*) and the predictions made by five other participants who have ostensibly taking part in the experiment under identical conditions. In majority
(minority) herd conditions four (two) of the five others’ predictions are correlated across trials. Their correlated predictions are accurate (correlated with the price trend) or random (uncorrelated with the price trend). The accuracy of the participants’ predictions will be assessed by correlations with the price trend. Whether the predictions made by the majority (minority) herd has a stronger effect than the price cue will be assessed by correlations with both the herd’s average predictions (after partialing out the price trend when the herd makes accurate predictions; referred to as herd error) and the price cue.

In Experiment 1 we compare majority and minority herd influences when the herd’s predictions are accurate or random. The remaining two experiments investigate whether increasing systematic processing by inducing attention focus on accuracy will affect majority herd influences (Experiment 2) and minority herd influences (Experiment 3).

Experiment 1

Experiment 1 investigates the influence from a majority herd compared to a minority herd when level of accuracy of the herd’s predictions vary. In two conditions the herds’ predictions of a future stock price are accurate, and in two conditions the herds’ predictions are random. In two majority herd conditions four of the other five participants’ predictions are correlated, and in two minority herd conditions two of the other five participants’ predictions are correlated. If a majority herd influence is associated with heuristic processing (the use of a consensus heuristic), an accurate majority herd would not have a larger influence than a random majority herd. In contrast, if a minority herd influence is associated with systematic processing, an accurate minority herd would have a larger influence than a random minority herd.

All the participants in Experiment 1 can make accurate predictions by utilizing the price cue. Thus, their predictions will then correlate with both the price trend and the price error. They will also make accurate predictions if influenced by the predictions by the accurate herd, in which case their predictions will correlate both with the price trend and the herd error. On the other hand, if the participants are influenced by the random herd, their predictions will only correlate with the herd error. On all the measures (the correlation with price trend, the correlation with price error, and the correlation with herd error), significant interactions between herd size and herd accuracy are hypothesized. More specifically, if the majority herd influence is mediated by heuristic processing, for an accurate majority herd the participants’ predictions will correlate with the price trend and the herd error, whereas for a random majority herd their predictions will correlate with the herd error. In contrast, if the minority herd influence is mediated by systematic processing, for an accurate minority herd the participants’ predictions will correlate with the price...
trend and the herd error, whereas for a random minority herd the participants’ predictions will correlate with the price trend and the price error.

Method
Participants. The participants were 64 undergraduates (40 women and 24 men) at University of Gothenburg, Göteborg, Sweden, volunteering to participate in return for SEK 50 (approximately US$ 8.0). They were recruited through sign-up sheets and electronic mails. The women’s mean age was 26.5 years ($SD = 5.0$) and the men’s mean age 26.3 years ($SD = 4.8$).

Design. Equal numbers of participants with sex and age balanced were randomly assigned to a 2 (Herd size: Minority vs. Majority) by 2 (Herd accuracy: Accurate vs. Random) by 50 (Trial) factorial design with trial as a repeated-measures factor.

Materials. The price was varied according to an increasing (decreasing) linear trend across trials with a mean of SEK 750 and ranging from SEK 395 to SEK 1076 (from SEK 1065 to SEK 441). The price cue was obtained by adding to the price a random error sampled from a normal distribution ($M = 0, SD = SEK 92$), resulting in an $r \approx .80$ with price trend. Two different sequences of price cues were used, one in the conditions with an increasing price trend and the other in the conditions with a decreasing price trend.

The predictions by each of the five others were obtained by random sampling from a normal distribution ($M = 0, SD = SEK 92$). For the two (minority) or four (majority) others included in the herd, a common error was also added by random sampling from the same distribution. In this way correlations ($r > .95$) between the others’ predictions were created. The predictions by the others not included in the herd were uncorrelated ($r < .20$) with the predictions by the others included in the herd and among themselves.

In the accurate herd conditions the price was on each trial added to the predictions by the others included in the herd. As a consequence, their predictions were correlated with the price trend ($r \approx .80$). A constant was added to all others’ predictions in both the accurate and random herd conditions (500 when the price was increasing, 1000 when the price was decreasing), thus resulting in predictions that were uncorrelated ($r < .20$) with the price trend.

Procedure. Participants were appointed via e-mail to come to the laboratory. Upon arrival they were seated in separate cubicles facing a computer screen. The instructions and all tasks were presented on this screen. The tasks were self-paced. An experimenter was present to supervise the participants. A session lasted for approximately 25 minutes.
The participants were informed that they would be presented the current price of a fictitious stock in 50 trials, each representing a new day, and that their task was to predict the price of the stock the following day by typing a price in a box shown on the screen. They were further told that in addition to the current stock price, they would on each trial be shown the predictions made by five others (identified by letters) who previously had participated in the experiment under identical conditions. They were reminded that the price presented on the following trial made it possible to assess the accuracy of their own prediction as well as the accuracy of the others’ predictions. Finally, they were informed that their own predictions were made anonymously and would not be shown to any other participants.

After having completed their predictions, the participants were requested to answer five questions on 9-point rating scales ranging from never (1) to always (9). Two questions were related to the experimental manipulations: “Did you perceive that the others were in agreement when making their predictions?” (Awareness of agreement), and “Did you perceive that the others’ predictions were accurate?” (Accuracy beliefs). Another three questions were related to social influence: “Did you believe the same as the others when your predictions coincided with their predictions?” (Independence beliefs), “Were you influenced by the others’ predictions?” (Perceived social influence), and “Did it matter to you if your predictions deviated from the others’ predictions?” (Importance of non-compliance). Participants were finally debriefed and paid.

Pilot Study. In order to check whether the accuracy of and correlations between the others’ predictions were possible to detect, another 16 undergraduates were recruited from the same pool. Upon arrival to the laboratory they were placed in separate cubicles and received a booklet to fill out at their own pace. Each page of the booklet represented a trial on which they were shown predictions of a current stock price ostensibly made by five other fictitious participants. These other participants were each labeled by a letter (A to E, F to J, K to O, or P to T in different blocks of trials). After ten trials the participants were requested to state which of the others’ predictions correlated across trials, and how accurate each of the others’ predictions were. When reporting the correlations they underlined the letters (from none to all) representing the others who were correlated. When reporting the accuracy they rated each participants’ accuracy on 9-point rating scales ranging from completely inaccurate (1) to completely accurate (9). A within-group design was employed so that all four conditions in the experiment (random majority, accurate majority, random minority, accurate minority) were tested in a counterbalanced order. The material was constructed by selecting every fifth trial from the set of 50 trials used in the experiment.
Table 1

Mean Ratings of Accuracy of Others’ Predictions Related to Condition (Pilot Study)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Other participant</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Random minority</td>
<td>3.4ₐ (1.8)</td>
<td>3.6ₐ (1.6)</td>
<td>3.8ₐ (1.7)</td>
<td>3.9ₐ (2.0)</td>
<td>4.0ₐ (2.0)</td>
</tr>
<tr>
<td>Random majority</td>
<td>3.1ₐ (1.4)</td>
<td>3.1ₐ (1.5)</td>
<td>3.3ₐ (1.5)</td>
<td>3.1ₐ (1.2)</td>
<td>2.7ₐ (1.8)</td>
</tr>
<tr>
<td>Accurate minority</td>
<td>6.8ₐ (1.3)</td>
<td>6.9ₐ (1.1)</td>
<td>3.8ₐ (1.3)</td>
<td>3.5ₐ (1.3)</td>
<td>2.7ₐ (1.3)</td>
</tr>
<tr>
<td>Accurate majority</td>
<td>6.1ₐ (1.6)</td>
<td>6.1ₐ (1.6)</td>
<td>6.0ₐ (1.5)</td>
<td>6.2ₐ (1.2)</td>
<td>2.6ₐ (1.5)</td>
</tr>
</tbody>
</table>

Note. 1 = completely inaccurate; 9 = completely accurate; different subscripts indicate that the mean differences are significant at \( p = .05 \) in paired \( t \)-tests.

Table 1 indicates that the differences in accuracy were detected in that paired \( t \)-tests revealed statistically significant differences in the accurate conditions between the others included in the herd (A, B, C, and D or A and B) and those not included in the herd, whereas there were no significant differences in the random conditions.

Table 2

Proportion Indicated Frequency of Correlations with the Others’ Predictions Related to Condition (Pilot Study)

<table>
<thead>
<tr>
<th>Condition</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
</tr>
<tr>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>Random minority</td>
<td>.88ₐ</td>
<td>.94ₐ</td>
<td>.25ₐ</td>
<td>.19ₐ</td>
<td>.13ₐ</td>
</tr>
<tr>
<td>Random majority</td>
<td>1.00ₐ</td>
<td>.94ₐ</td>
<td>.94ₐ</td>
<td>.81ₐ</td>
<td>.06ₐ</td>
</tr>
<tr>
<td>Accurate minority</td>
<td>1.00ₐ</td>
<td>1.00ₐ</td>
<td>.06ₐ</td>
<td>.00ₐ</td>
<td>.06ₐ</td>
</tr>
<tr>
<td>Accurate majority</td>
<td>.94ₐ</td>
<td>.94ₐ</td>
<td>.94ₐ</td>
<td>1.00ₐ</td>
<td>.00ₐ</td>
</tr>
</tbody>
</table>

Note. Different subscripts indicate that the mean differences are significant at \( p = .05 \) in paired \( t \)-tests.

In Table 2 paired \( t \)-tests revealed statistically significant differences in the majority and minority conditions between those others whose predictions were
correlated (A, B, C, and D or A and B) and those others whose predictions were not correlated.

Results

Post-Experimental Questions. Mean answers to each post-experimental question are given in Table 3.

Table 3
Mean Answers to Post-Experimental Questions (Experiment 1)

<table>
<thead>
<tr>
<th>Question</th>
<th>Minority herd</th>
<th>Majority herd</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Random M (SD)</td>
<td>Accurate M (SD)</td>
</tr>
<tr>
<td>Awareness of agreement</td>
<td>4.8 (1.9)</td>
<td>2.9 (1.6)</td>
</tr>
<tr>
<td>Accuracy beliefs</td>
<td>2.8 (1.4)</td>
<td>4.1 (1.2)</td>
</tr>
<tr>
<td>Independence beliefs</td>
<td>5.3 (2.4)</td>
<td>5.9 (2.4)</td>
</tr>
<tr>
<td>Perceived social influence</td>
<td>3.4 (2.2)</td>
<td>6.0 (2.8)</td>
</tr>
<tr>
<td>Importance non-compliance</td>
<td>3.4 (2.4)</td>
<td>5.1 (2.7)</td>
</tr>
</tbody>
</table>

Note. The answers were given as frequency ratings on scales ranging from 1 to 9.

Parallel 2 (Herd size: Majority vs. Minority) by 2 (Herd accuracy: Random vs. Accurate) analyses of variance (ANOVAs) were performed on each answer. On awareness of agreement the main effects of herd size, $F(1, 60) = 12.84, p = .001$, and herd accuracy, $F(1, 60) = 7.60, p = .008$, reached significance. Participants perceived the others as being more in agreement in the majority than in the minority conditions ($M_{majority} = 5.4$ vs. $M_{minority} = 3.8$) and more in the random conditions than in the accurate conditions ($M_{random} = 5.3$ vs. $M_{accurate} = 4.0$). On perceived social influence, a significant main effect of herd size was due to that participants were more influenced by the others in the majority conditions than in the minority conditions ($M_{majority} = 6.0$ vs. $M_{minority} = 4.7$), $F(1, 60) = 4.46, p = .039$, and a significant main effect of herd accuracy was due to that participants were more influenced by the others in the accurate conditions than in the random conditions ($M_{accurate} = 6.7$ vs. $M_{random} = 3.9$), $F(1, 60) = 21.00, p < .001$. On accuracy beliefs, a main effect of accuracy was found, $F(1, 60) = 15.32, p < .001$, implying that accurate herds were perceived to be more
accurate than random herds ($M_{\text{accurate}} = 4.3$ vs. $M_{\text{random}} = 2.9$). On importance of non-compliance, a significant main effect of herd accuracy, $F(1, 60) = 7.64, p = .008$, was due to that predictions that deviated from the others’ predictions were more important in conditions with accurate herds than in conditions with random herds ($M_{\text{accurate}} = 5.1$ vs. $M_{\text{random}} = 3.4$).

**Predictions.** The three dependent variables were constructed as product-moment correlations between the participants’ predictions and the price trend, between the participants’ predictions and the price error, and between the participants’ predictions and the herd error for each participant in each of two blocks consisting of 20 trials, excluding the learning phase consisting of the initial 10 trials. All statistical analyses were performed on Fisher’s $z_r$ transformed values. Means are reported in Table 4. Since increasing vs. decreasing price trend had no effect, the means are averaged across this factor.

Table 4
**Mean Fisher z, Transformed Correlations with Price Trend, Price Error, and Herd Error Related to Herd Size, Herd Accuracy, and Block (Experiment 1)**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Block</th>
<th>Minority herd</th>
<th>Majority herd</th>
<th>Minority herd</th>
<th>Majority herd</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Random M (SD)</td>
<td>Accurate M (SD)</td>
<td>Random M (SD)</td>
<td>Accurate M (SD)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Price trend</td>
<td>1</td>
<td>0.45 (0.22)</td>
<td>0.44 (0.18)</td>
<td>0.44 (0.23)</td>
<td>0.57 (0.22)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.42 (0.28)</td>
<td>0.50 (0.28)</td>
<td>0.44 (0.25)</td>
<td>0.54 (0.17)</td>
</tr>
<tr>
<td>Price error</td>
<td>1</td>
<td>0.63 (0.29)</td>
<td>0.74 (0.34)</td>
<td>0.66 (0.42)</td>
<td>0.30 (0.27)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.63 (0.32)</td>
<td>0.51 (0.35)</td>
<td>0.48 (0.38)</td>
<td>0.24 (0.26)</td>
</tr>
<tr>
<td>Herd error</td>
<td>1</td>
<td>0.40 (0.33)</td>
<td>0.36 (0.24)</td>
<td>0.62 (0.91)</td>
<td>0.83 (0.45)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.23 (0.22)</td>
<td>0.30 (0.23)</td>
<td>0.67 (1.21)</td>
<td>0.75 (0.38)</td>
</tr>
</tbody>
</table>

Parallel 2 (Herd size: Majority vs. Minority) by 2 (Herd accuracy: Random vs. Accurate) by 2 (Block) ANOVAs with block as repeated-measures factor were performed on each dependent variable. On the correlation with the price trend, a non-significant effect of herd accuracy, $F(1, 56) = 2.99, p = .089$, $\omega^2_{\text{partial}} = .02$, was due to a tendency that the correlation was higher in the conditions with accurate herds than in the condition with random herds ($M_{\text{accurate}} = 0.51$ vs. $M_{\text{random}} = 0.29$).
No significant interaction was observed between herd size and herd accuracy, $F(1, 56) < 1$.

The ANOVA on the correlation with the price error yielded a significant main effect of herd size, $F(1, 56) = 6.49, p = .013$, $\omega^2_{\text{partial}} = .04$, due to that in the majority conditions the correlation was lower than in the minority conditions ($M_{\text{majority}} = 0.42$ vs. $M_{\text{minority}} = 0.63$). A significant main effect of block substantiated that the correlation was higher in the first block than in the second block ($M_{\text{block1}} = 0.58$ vs. $M_{\text{block2}} = 0.47$), $F(1, 56) = 10.51, p = .002$, $\omega^2_{\text{partial}} = .07$. A three-way interaction between herd accuracy, herd size, and block was also revealed, $F(1, 56) = 5.83, p = .019$, $\omega^2_{\text{partial}} = .02$. In the random majority conditions the correlation decreased over blocks ($M_{\text{random majority block 1}} = 0.66$ vs. $M_{\text{random majority block 2}} = 0.48$), but no decrease was found in the random minority conditions ($M_{\text{random minority block 1}} = 0.63$ vs. $M_{\text{random minority block 2}} = 0.63$). In the accurate majority conditions the correlation was lower and decreased less over blocks ($M_{\text{accurate majority block 1}} = 0.30$ vs. $M_{\text{accurate majority block 2}} = 0.24$) than in the accurate minority conditions ($M_{\text{accurate minority block 1}} = 0.74$ vs. $M_{\text{accurate minority block 2}} = 0.51$).

A significant main effect of herd size on the correlation with the herd error indicated that participants were more influenced by the predictions by the herd in the majority conditions than in the minority conditions ($M_{\text{majority}} = 0.72$ vs. $M_{\text{minority}} = 0.32$), $F(1, 56) = 7.67, p = .007$, $\omega^2_{\text{partial}} = .05$. The interaction between herd size and herd accuracy did not reach significance, $F(1, 56) < 1$.

**Discussion**

The results indicated that the majority herd exerted more influence on the participants’ predictions than did the minority herd. However, the hypothesized interactions between herd size and herd accuracy were not found on any of the dependent variables, implying that the participants were more influenced by a majority herd than a minority herd independently of accuracy in the herd’s predictions. If the minority herd influence is associated with systematic processing as hypothesized, the effect of herd accuracy should have been evident in the minority conditions. However, the results still suggest that the majority herd influence is mediated by heuristic processing.

An alternative interpretation that the majority herd is associated with both heuristic and systematic processing is suggested by the significant three-way interaction between herd size, herd accuracy, and block. When the herd’s predictions were random, the correlation with price error decreased over blocks for the conditions with a majority herd but not for the conditions with a minority herd. Thus, in the second block where the random herd’s predictions increasingly deviated from the price error, the participants ignored the price error and instead followed the random majority herd. This suggests the use of heuristic processing. When the the herd’s predictions instead were accurate, the
correlation with price error was lower in the majority herd conditions and decreased less over blocks than in the minority herd conditions. The low level of utilization of the price cue in both blocks in the accurate majority herd conditions further indicates that the accuracy in the herd’s predictions had an influence on the participants in the majority herd conditions but not in the minority herd conditions. This suggests the use of systematic processing.

If the participants had processed the price cue thoroughly, it would have been possible for them to infer the price trend and thus to make more accurate predictions. However, even though the results from both the pilot study and the post-experimental questions suggested that the participants were able to detect the differences in accuracy between the conditions when this was not their assigned task, detecting the price trend might have been too difficult when the assigned task was to make predictions.

In conclusion, since a random majority herd had the same influence as an accurate majority herd, majority influence appears to be associated with heuristic processing. However, the results showing that the correlation with price error increased when the majority herd was accurate may be interpreted to be the result of systematic processing. The findings of Experiment 1 thus suggest that majority influence is associated with both heuristic and systematic processing. No evidence indicated that a minority herd influence is associated with systematic processing.

**Experiment 2**

The main results of Experiment 1 indicated that a majority herd has more influence than a minority herd independently of the level of accuracy in their predictions. In Experiment 2 we ask whether the weak influence of accuracy in the majority herd’s predictions is due to the difficulty participants may have had in detecting the accuracy of the predictions. We assume that in the present experiments the participants will primarily focus their attention on the consistency in the herd’s predictions over trials. For this reason the difficulty in detecting accurate performance may be reduced if the participants are instead induced to focus their attention on the herd’s performance. We then expect an interaction between performance focus and herd accuracy on the correlation with the herd error since the influence from the random majority herd would decrease when the participants focus their attention on the herd’s performance. At the same time the correlation with the price error would increase since the price cue is utilized to a larger degree. As a consequence of being influenced by the herd’s predictions or utilizing the price cue, the participants’ predictions will correlate with the price trend when the herd is accurate and when the participants focus on a random herd’s performance.

**Method**
Participants. The participants were another 80 undergraduates (54 women and 26 men) at University of Gothenburg volunteering to participate in return for SEK 50 (approximately US$8). They were recruited through sign-up sheets and electronic mails. The women’s mean age was 28.1 years (SD = 9.7) and the men’s mean age 26.5 years (SD = 9.0).

Design. Equal numbers of participants with sex and age balanced were randomly assigned to a 2 (Accuracy: Random vs. Accurate) by 2 (Attention focus: Performance vs. Consistency) by 50 (Trial) factorial design with trial as a repeated-measures factor.

Materials and Procedure. The materials and the procedure were the same as in Experiment 1 except that after 10, 30 and 50 trials participants were requested to state either how many of the five others who made accurate predictions (performance focus conditions), or how many of the others who made consistent predictions (consistency focus conditions). Only majority conditions with four others’ correlated predictions were used.

Results

Manipulation Checks. After 10, 30 and 50 trials participants were requested in the performance-focus conditions to answer the question: “How many of the others made accurate predictions?” and in the consistency-focus conditions: “How many others were correlated with each other?” Two separate 2 (Herd accuracy: Random vs. Accurate) by 2 (Trial number: 30 vs. 50) ANOVAs with trial number as a repeated-measures factor were performed on the answers to each of the questions, respectively, excluding the answers on trial 10. Since no effect of trial number was observed, means were averaged across this factor.

Table 5
Mean Indicated Number of Others Related to Attention Focus and Herd Accuracy (Experiment 2)

<table>
<thead>
<tr>
<th></th>
<th>Performance focus</th>
<th>Consistency focus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Random</td>
<td>Accurate</td>
</tr>
<tr>
<td>M (SD)</td>
<td>0.7 (1.7)</td>
<td>1.3 (1.8)</td>
</tr>
</tbody>
</table>

Note. The responses were given as the number of others whose predictions were accurate (performance focus) or correlated with the others (consistency focus).

As may be seen in Table 5, the participants in the performance focus conditions perceived that on average only one other made accurate predictions. The
ANOVA showed no significant main effect of accuracy, $F(1, 38) = 1.61, p = .212$. In the consistency focus conditions, participants perceived that on average 3.4 participants were correlated. No significant main effect of accuracy was observed, $F(1, 38) < 1$.

**Predictions.** The correlations with price trend, price error, and herd error were computed as in Experiment 1. No significant effects of increasing or decreasing price trend were found. In Table 6 means are therefore presented averaged across this factor. All statistical analyses were performed on Fisher’s $z_r$ transformed values.

Table 6

<table>
<thead>
<tr>
<th>Measure</th>
<th>Performance focus</th>
<th>Consistency focus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Random</td>
<td>Accurate</td>
</tr>
<tr>
<td>Block M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Price trend</td>
<td>1 0.40 (0.18)</td>
<td>0.49 (0.14)</td>
</tr>
<tr>
<td></td>
<td>2 0.56 (0.27)</td>
<td>0.52 (0.21)</td>
</tr>
<tr>
<td>Price error</td>
<td>1 0.74 (0.31)</td>
<td>0.67 (0.39)</td>
</tr>
<tr>
<td></td>
<td>2 0.68 (0.36)</td>
<td>0.72 (0.35)</td>
</tr>
<tr>
<td>Herd Error</td>
<td>1 0.46 (0.31)</td>
<td>0.39 (0.39)</td>
</tr>
<tr>
<td></td>
<td>2 0.19 (0.12)</td>
<td>0.34 (0.38)</td>
</tr>
</tbody>
</table>

Parallel 2 (Herd accuracy: Random vs. Accurate) by 2 (Focus: Performance vs. Consistency) by 2 (Block) ANOVAs with block as repeated-measures factor was performed on each of the dependent variables. A significant main effect of accuracy was shown on the correlation with price trend due to that in the conditions with accurate herds the correlation was higher than in the conditions with random herds ($M_{\text{accurate}} = 0.53$ vs. $M_{\text{random}} = 0.45$), $F(1,76) = 3.96, p = .050$, $\omega^2_{\text{partial}} = .02$. No significant interaction was revealed between focus and herd accuracy, $F(1,76) = 1.75, p = .190$, $\omega^2_{\text{partial}} = .00$. The main effect of block was
significant, $F(1,76) = 6.80, p = .011, \omega^2_{\text{partial}} = .03$, due to a higher correlation in the second block than in the first block ($M_{\text{block1}} = 0.45$ vs. $M_{\text{block2}} = 0.53$). A significant interaction between herd accuracy and block was due to a difference in correlation between accurate and random herds in the first block ($M_{\text{random block 1}} = 0.35$ vs. $M_{\text{accurate block 1}} = 0.52$), but no such difference in the second block ($M_{\text{random block 2}} = 0.52$ vs. $M_{\text{accurate block 2}} = 0.53$), $F(1,76) = 5.03, p = .028, \omega^2_{\text{partial}} = .02$.

The ANOVA on the correlation with price error revealed a significant main effect of focus, $F(1,76) = 20.30, p < .001, \omega^2_{\text{partial}} = .10$. The correlation was larger in the performance focus conditions than in the consistency focus conditions ($M_{\text{performance}} = 0.70$ vs. $M_{\text{consistency}} = 0.42$). The interaction between focus and herd accuracy did not reach significance, $F(1,56) < 1$.

The ANOVA on the correlation with herd error yielded a significant main effect of focus. The correlation was lower for performance focus than consistency focus ($M_{\text{performance}} = 0.34$ vs. $M_{\text{consistency}} = 0.55$), $F(1,76) = 6.22, p = .015, \omega^2_{\text{partial}} = .03$. No significant interaction between focus and herd accuracy was found, $F(1, 56) < 1$. A significant main effect of block was due to a higher correlation in the first block than in the second block ($M_{\text{block1}} = 0.52$ vs. $M_{\text{block2}} = 0.36$), $F(1,76) = 10.42, p < .001, \omega^2_{\text{partial}} = .06$. A significant interaction between herd accuracy and block was due that the difference in correlation was larger in the random conditions ($M_{\text{random block 1}} = 0.57$ vs. $M_{\text{random block 2}} = 0.31$) than in the accurate conditions ($M_{\text{accurate block 1}} = 0.48$ vs. $M_{\text{accurate block 2}} = 0.42$), $F(1,76) = 6.46, p = .013, \omega^2_{\text{partial}} = .03$.

**Discussion**

As indicated by a lower correlation with herd error and a higher correlation with price error, inducing the participants to focus on the others’ performance reduced the majority herd influence. A possible interpretation is that the performance focus decreased heuristic and increased systematic processing, thus making the participants infer that the herd’s predictions were not accurate. However, whether the majority herd made random or accurate predictions had no effect on the correlations with price error or herd error. The participants were perhaps unable to detect the accuracy in the herd’s predictions. This was substantiated by the manipulation check showing that the participants underestimated the number of others making accurate predictions. Yet, in the first block the correlation with price trend was higher when the herd was accurate than when the herd was random.

In support of the notion that the participants primarily focus their attention on the consistency in the herd’s predictions, the manipulation check showed that they were more accurate in perceiving the correlations between the others’ predictions than they were in inferring whether the others’ predictions were accurate or random. It may also be the case that detecting accuracy in the herd’s
predictions was a more difficult task. The correlation between the herd’s average predictions and the price trend was perhaps too low. This is not refuted by the observed effect of accuracy on the correlation with price trend in the first block, since this effect was not primarily due to a difference in the performance focus conditions.

Taken together, the results of Experiments 1 and 2 suggest that a majority herd has more influence than a minority herd independently of accuracy in their predictions, and that the influence from a random majority herd may be prevented by inducing the participants to focus on the accuracy in the others’ predictions. A remaining question is whether performance focus would have an effect on minority influence.

**Experiment 3**

If the perception of a minority herd induces systematic processing, it is hypothesize that an accurate minority herd should have more influence than a random minority herd. In Experiment 3 we again investigate whether a minority herd has a larger influence when its predictions are accurate than when its predictions are random. A higher level of accuracy in the herd’s predictions than in Experiments 1 and 2 seems to be required for this to occur.

Because systematic processing is hypothesized to already be induced by a minority herd, focusing attention on performance should not moderate the influence of an accurate minority herd. Still, the degree of systematic processing may be possible to increase. We therefore investigate whether the influence of an accurate minority would increase in a performance-focus condition.

**Method**

*Participants.* The participants were another 64 undergraduates (39 women and 21 men) at University of Gothenburg volunteering to participate in return for SEK 50 (approximately US$8). They were recruited through sign-up sheets and electronic mails. The women’s mean age was 26.3 years ($SD = 8.3$) and the men’s mean age 23.6 years ($SD = 3.7$).

*Design.* Equal numbers of participants with sex and age balanced were randomly assigned to a 3 (Condition: Random vs. Accurate vs. Accurate with focus) by 50 (Trial) design with trial as a repeated-measures factor.

*Materials and Procedure.* The procedure was the same as in the preceding experiments except that the sequence of events on each trial was changed so that the price cue and the others’ predictions were first shown, then after having made their prediction, the participants were shown the correct price.

The orders of presentation of the price cue and the price were each randomly varied across trials such that they were uncorrelated ($r < .20$). The predictions made by the two others in the minority herd were uncorrelated with the price cue ($r < .20$) and either random (uncorrelated with the price, $r < .20$) or
accurate (correlated with the price, $r \approx .95$). The predictions made by the others not in the minority herd were always random, thus uncorrelated with the price cue ($r < .20$), the price ($r < .20$), and the others’ predictions ($r < .20$). Four random sequences of the others’ predictions were used for different participants in each condition.

In the performance focus condition, in which the minority herd made accurate predictions, the participants were after trials 10, 30, and 50 asked to indicate who of the five others made accurate predictions.

Results

Manipulation Checks. As may be seen in Table 7, paired t-tests confirm that after both 30 and 50 trials the participants in the minority herd (A and B) are rated as more accurate than the other participants.

Table 7
Mean Ratings of Accuracy of Others’ Predictions in Trial 30 and 50 (Experiment 3)

<table>
<thead>
<tr>
<th>Trial</th>
<th>A M (SD)</th>
<th>B M (SD)</th>
<th>C M (SD)</th>
<th>D M (SD)</th>
<th>E M (SD)</th>
<th>None M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>.55 a (.51)</td>
<td>.65 a (.49)</td>
<td>.05 b (.22)</td>
<td>.05 b (.22)</td>
<td>.10 b (.31)</td>
<td>.20 b (.41)</td>
</tr>
<tr>
<td>50</td>
<td>.65 a (.49)</td>
<td>.85 a (.37)</td>
<td>.05 b (.22)</td>
<td>0 b (0)</td>
<td>.05 b (.22)</td>
<td>.10 b (.31)</td>
</tr>
</tbody>
</table>

Note. 0 = random; 1 = accurate; different subscripts indicate that the mean differences are significant at $p=.05$ in paired t-tests.

Predictions. The product-moment correlations between the participants’ predictions and the price cue and between the participants’ predictions and the minority herd’s average predictions were computed for each participant in each of two blocks of 20 trials, excluding the initial 10 trials. All the following statistical analyses are performed on Fisher’s $z_r$ transformed values. Table 8 shows the means related to condition and block.

Table 8
Mean Fisher $z_r$ Transformed Correlations with Price Cue and the Minority Herd’s Predictions Related to Herd Accuracy and Block (Experiment 3)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Block</th>
<th>Condition</th>
<th>Random M (SD)</th>
<th>Accurate M (SD)</th>
<th>Accurate with focus M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Random</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accurate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accurate</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Price cue    1  0.58 (0.40)  0.31 (0.30)  0.19 (0.11)
            2  0.46 (0.27)  0.35 (0.29)  0.14 (0.13)
Herd’s predictions 1  0.21 (0.15)  0.54 (0.37)  1.31 (0.89)
            2  0.30 (.29)  0.57 (0.41)  1.58 (1.00)

A 3 (Herd accuracy: Random vs. Accurate vs. Accurate with performance focus) by 2 (Block) ANOVA with block as repeated-measures factor yielded a significant main effect of herd accuracy on the correlation with the price cue ($M_{random} = 0.52$ vs. $M_{accurate} = 0.33$ vs. $M_{accurate with focus} = 0.16$), $F(2,57) = 11.48$, $p < .001$, $\omega^2_{partial} = .26$. Tukey post-hoc tests only showed a significant difference between the condition with a random minority herd and the condition with an accurate minority herd with performance focus.

A parallel 3 (Herd accuracy: Random vs. Accurate vs. Accurate with focus) by 2 (Block) ANOVA with block as repeated-measures factor was performed on the correlation with the herd’s average predictions. A significant main effect was observed of accuracy ($M_{random} = 0.25$ vs. $M_{accurate} = 0.56$ vs. $M_{accurate with performance focus} = 1.44$), $F(2,57) = 25.20$, $p < .001$, $\omega^2_{partial} = .34$. Tukey post-hoc tests revealed significant differences between the accurate minority herd condition with performance focus and the other conditions, whereas the difference between the random herd and the accurate herd conditions was not significant.

**Discussion**

The results showed that the minority herd had a larger influence on the participants’ predictions when it was accurate and the participants were induced to focus their attention on the herd’s performance. This suggests that the presence of a minority herd does not by itself elicit systematic processing. Furthermore, the results indicate that in Experiment 3 the participants were able to detect that the herd’s predictions were accurate. In general detecting an accurate minority herd consisting of two in a group of five others should be more difficult than detecting an accurate majority herd consisting of four in a group of five others. Thus, for this reason an accurate majority herd should have been possible to detect in this experiment. It may then be concluded that the absence of an effect of an accurate majority herd in Experiments 1 and 2 was likely due to that, in these experiments, the task to detect accuracy was too difficult.

It should be noted that without a performance focus, the accurate herd still tended to have a larger influence on the participants’ predictions than a random
minority herd, although the difference was not significant. This suggests that in an uncertain prediction task systematic processing may vary in degree.

**General Discussion**

In past research herding in stock markets has been conceived of as rational (e.g. Drehmann et al., 2005). In the present research we compare herds making accurate predictions with herds making random predictions, thus creating the opportunity to investigate both rational and irrational herding. Furthermore, we have introduced a new perspective on herding in stock markets by relating it to research on informative social influence and different types of information processing.

A general outcome of the present experiments is that the dual-process theory proposed by Moscovici (1985) did not receive unequivocal support. Whereas this theory posits that majority influences are associated with heuristic processing and minority influences are associated with systematic processing, the present results are interpreted as showing that majority herd influences are primarily associated with heuristic processing. However, a minority herd did not elicit systematic processing. In the following we propose an alternative account, as illustrated in Figure 1.
In the uncertain prediction task like that we used, people are likely to search and evaluate the usefulness of various pieces of information or cues (Busemeyer, Byun, Delosh, & McDaniel, 1997). We assume that the participants believed that the current stock price (the price cue) is useful in predicting the future price. In fact the price cue had predictive value in Experiments 1 and 2 but not in Experiment 3. However, since the price cue only had a probabilistic relation to the future price, it is conceivable that the participants preferred to attend to the consistent predictions by others, believing that these provide more useful information. If a majority made consistent predictions, this led to a majority herd influence; if a minority made consistent predictions, this led that the price cue was utilized and no herd influence.

Figure 1. Flow diagram of information processing in the prediction task.
In Experiments 1 and 2 when the accurate herd’s predictions like the price cue had a probabilistic relation to the future price, accuracy of the majority or minority herd had no effect. Therefore, whether the herd was a majority or minority, heuristic processing was elicited. Thus, we assume that the consensus heuristic is accompanied by an “opposite” heuristic, resulting from the belief that “a minority cannot be accurate”. In further support of this, in Experiment 3 when the price cue lacked predictive value and the predictive value of the herd’s predictions were higher, the participants were still not influenced by a minority herd’s accurate predictions. In order to break heuristic processing leading to that the minority herd’s predictions had no influence, it was necessary to induce an attention focus on the herd’s performance. Only then a minority herd’s accurate predictions had an influence.

Our account is possibly incomplete in not taking into account the tendency in Experiment 3 that an accurate minority herd had larger influence than a random minority herd. If in light of this tendency the hypothesis is maintained that a minority elicits systematic processing, it may be proposed that systematic processing varies in degree. Likewise, our account ignores the possibility noted in Experiment 1 that a majority herd may sometimes be associated with systematic processing. More specific hypotheses need to be specified and investigate in additional studies.

The present results are in some respects an extension of previous findings in research on social influence (e. g. Bond, 2005). For instance, the results bear similarities to the pattern predicted by Martin, Hewstone et al. (2007), arguing that majority messages only instigate systematic processing in the presence of additional factors such as instructions increasing processing depth (Craik & Tulving, 1975). The Bohner et al. (2008) study showed different patterns of social influence and systematic processing depending on how the source of information (the majority or the minority) was framed. When framed as being similar to participants, like in our experiments, the majority influence was strong. However, the lack of support for minority influences is difficult to reconcile with either Bohner et al. (2008) or Moscovici’s (1985) theory. In conclusion, the association between majority and minority influences and the type of processing thus seem to depend on context. In our uncertain prediction task we have demonstrated that heuristic processing has a larger role than it has been ascribed in previous research on informational social influence, and that the systematic processing is not associated with minority influence under these circumstances.

In general, the use of a consensus heuristic is justified by the fact that a crowd makes predictions that are better than individuals or even experts (Surowiecki, 2004). However, a number of factors may cause failures in the crowds’ “wisdom.” One example is related to social influence; the members of the crowd may be conscious of the other members’ opinions and begin to imitate

each other rather than making predictions independently. In such cases, people who are influenced by the crowd’s consensus predictions will obtain worse outcomes. Our results suggest that people may be influenced by large crowds irrespective of both how the crowd’s wisdom was gained and how accurate its predictions are.

A bulk of previous research in financial economics concerns what information traders use when making investment decisions. An issue is the role of news media and its relationship to market actors. Based on survey data from professional traders, Oberlechner and Hocking (2004) concluded that foreign exchange traders do not consider the perceived truth and accuracy to be as important features of news as information speed, expected market impact, and anticipated market surprise. It is suggested that investors have limited time to check the accuracy in news releases, and that they anticipate other traders to be equally affected by the news regardless of its accuracy. Thus, in this respect making decisions consistent with a herd of investors may be a conscious strategy which is more important than carefully evaluating the validity of the information. A similar result regarding the relation between herd influence and accuracy of its predictions has been found when following others have not been an explicit strategy. In a survey of financial analysts (Welch, 2000) the results showed that their predictions were influenced by the established consensus forecast, but this influence was not stronger when the forecast provided accurate information. In a similar vein, the present research shows a strong herd influence, independently of the level of accuracy of the herd’s predictions. Thus, the same pattern of findings is demonstrated in our laboratory experiments, even though it is not clear whether or not a deliberate strategy was used.

Despite the noted similarity in results, a short-coming of the present experiments is that knowledge of the stock market is not investigated. Such knowledge possessed by experts may be an important determinant of actual stock investments. However, some (e.g., Shefrin, 2002; Taleb, 2004) argue that stock investments are highly influenced by random factors. Also, in real life investments, it will be difficult to judge which others constitute a herd. Furthermore, the present experiments primarily focused on the informational aspects of herding, not on reputational or normative concerns (Sias, 2004). One avenue to investigate the generalizability of the present findings is to test their invariance across different investment tasks in laboratory experiments.

References


**Author Note**

Maria Andersson, Department of Psychology, University of Gothenburg, Sweden; Ted Martin Hedestrom, Department of Psychology, University of Gothenburg, Sweden; Tommy Gaerling, Department of Psychology, University of Gothenburg, Sweden

This research was financially supported by a grant from the Swedish Foundation for Strategic Environmental Research (MISTRA) to the program "Behavioral Impediments to Sustainable Investments". Thanks are due to Isak Barbopoulos and Lisa Ohman for assistance in collecting the data.

Correspondence concerning this article should be addressed to Maria Andersson, Department of Psychology, University of Gothenburg, PO Box 500, SE-40530 Goeteborg, Sweden. E-mail: Maria.Andersson@psy.gu.se
A CASE STUDY OF A COLLAPSING HANDBALL TEAM

Erwin Apitzsch

Abstract
Collective collapse in team sports, conceived in terms of negative psychological momentum when the players on a team suddenly perform below the expected level despite having had a good start, was investigated involving nine male players from an elite handball team. Semi-structured interviews were employed. The major causes of collective collapse were found to be inappropriate behavior, failure of the role system to function properly, negative communication within the team, a change in the tactics of the opposing team, and goals being scored by that team. Factors seen as needing to be dealt with to prevent collective collapse included negative thinking, negative emotions, and negative emotional contagion. The study provides a team perspective on negative psychological momentum as well as tentative proposals for avoiding collective collapse.

Keywords: Collective collapse, communication, performance, roles, emotional contagion

Background
In team sports, sudden and unexpected shifts in performance can sometimes be observed. A soccer team may be ahead 2-0 after 70 minutes of play, only to lose by 2-3 when the final 20 minutes have been played. This can be termed collective collapse when such an outcome is due to the sudden underperformance of the players of the team originally in the lead. Apitzsch (2006) has suggested that collective collapse occurs when, in a match of considerable or decisive importance, the majority of the players on a team suddenly perform below their expected level after the team has had a good or normal start, or when they underperform from the very start. What causes collective collapse?

The following is a summary of various theoretical perspectives on collective collapse discussed earlier by Apitzsch (2006). The performance of an athlete is strongly affected in psychological terms by the person’s thoughts and emotions, and by the context at hand. It thus appears reasonable, in investigating the phenomenon of collective collapse within team sports, to examine not only the overt behavior but also the cognitions and affects associated with it. A cognitive approach that could be taken is one based on Janis’ (1982) conception of groupthink, or the tendency of a group to act primarily in accordance with normative pressures rather than on the basis of more relevant considerations. Carron, Hausenblas, and Eys (2005) maintained that groups endeavoring to solve a crisis, which the initial stages of a collective collapse indeed represent, are particularly vulnerable to groupthink. Players on a soccer team leading by 2-
0 towards the end of a match may feel, for example, that the match is already won and behave accordingly. This can mean their becoming more passive in their play and failing to put the effort needed into the game, allowing the opposing team to catch up. Taking an affective approach as a team member, regardless of whether the emotions involved are positive or negative in character, often results in emotional contagion, the person’s emotions being transferred to the other members of the team. Kelly and Barsade (2001) noted that strong emotions can markedly affect the cohesion, commitment, and performance of a team. Tickle-Degnan and Puccinelli (1999) found negative moods and emotions to be more readily transferred from one member of a team to another than positive moods and emotions are. Cacioppo and Gardner (1999), emphasizing connections between emotions and cognition and the strong effects that emotional processes have on how one thinks, noted that persons who are worried or afraid tend to devote greater attention to what are regarded as dangers or threats. Thus, a negative mood on the part of various members of a team, or of the team as a whole, can increase the probability of bad performance. Totterdell’s (2000) finding that when members of the one team become happier, members of the other team tend to become less happy appears to have direct bearing on collective collapse. For example, when a soccer team is leading by 2-0 and the opposing team makes a goal, this can put the team in the lead in a bad mood and improve the mood of the opposing team. One behavioral approach to analyzing this concerns the rationality of group members. Bion (1961) introduced the concepts of workgroup (a rational group) and basic-assumption group (an irrational group) to describe what can happen in a small-group setting. The workgroup aims at what is most rational for the group’s existence and the tasks the group is to perform. The members of a workgroup tend to be fully aware of their goals, to be organized appropriately, and to cooperate with one another in order to achieve their goals as effectively as possible. The basic assumption group, in contrast, is characterized by ineffectiveness and contradictory behavior. Although no group can be completely rational or effective, the aim should be to keep irrational behavior at a minimum. The workgroup of a soccer team consists, in a broad sense, of the 20-25 players who constitute the squad and, in a narrower sense, the 11 or so of these who tend to play regularly during a match. Collective collapse has not received much attention in the sport psychology literature, either in the latest editions of general textbooks (Cox, 2006; Morris & Summers, 2004; Tenenbaum & Eklund, 2007; Weinberg & Gould, 2007; Williams, 2001) or in books specifically devoted to team sports (Carron, Hausenblas, & Eys, 2005; Lidor & Henschen, 2003). However, a number of terms in general use are related to collective collapse, such as psychological momentum, critical moments, choking, slumps, and mistakes. Psychological momentum can be either positive or negative, in accordance with the definition
suggested by Taylor and Demick (1994, p. 54) of its being “a positive or negative change in cognition, affect, physiology, and behavior caused by an event or series of events that will result in a commensurate shift in performance and competitive outcome”. A multidimensional model proposed by Taylor and Demick (1994) is one of three different models that have been suggested as serving to explain psychological momentum, which refers to a chain of events that starts with a precipitating event which lead to changes in cognition, affect, and physiological arousal, and thus to changes in behavior, performance and outcome that can be either positive or negative. In terms of that model, both opponent factors and what the team members experience can moderate the strength and the effects of psychological momentum. The precipitating events leading up to it can be of crucial importance. The negative momentum of the one team can also precipitate positive momentum on the part of the other. Thus, a goal in soccer scored by the trailing team can result in negative momentum for the team in the lead. The other two models proposed to explain psychological momentum are the Antecedents-Consequences Model (Vallerand, Colavecchio, & Pelletier, 1988), which only deals with positive psychological momentum, and the Projected Performance Model (Cornelius, Silva, Conroy, & Peterson, 1997), which considers psychological momentum as primarily a way of describing performance shifts. Collective collapse can be incorporated into the Multidimensional Model of Taylor and Demick (1994) through its being seen as representing negative psychological momentum at the team level.

A variety of researchers (e.g., Baumeister, 1984; Carlstedt, 2004; Coward, 2006) and practitioners (e.g., Goldberg, 1998; Halden-Brown, 2003), have addressed issues related to collective collapse. Carlstedt (2004) maintained that sport psychologists and sport experts in general are unable to explain the phenomenon of a team’s collapse adequately on a scientific basis and that research has had very little to say about the relationship between psychological factors and athletic performance under various critical conditions during competition.

Choking is a term that Baumeister (1984) employs for denoting performance decrements in situations in which pressure is involved and good performance is important. Pressure refers to the presence of factors that increase the importance of performing well. Internal psychological factors characterizing the individual can be important components of pressure. Questions such as “What if I fail?” and thoughts such as “We’ve already won the match” can disrupt performance and result in an unfavorable outcome. Choking appears to be linked with the fear of being unable to live up to high expectations. In several experiments that Baumeister conducted he demonstrated that performance decrements can often be produced by directing one’s attention at one’s own performance, results which were in line with his self-awareness model of choking. The contributions that Baumeister has made to an understanding of the choking phenomenon are
substantial, especially the account he provides of the cognitive processes underlying drops in performance, yet he fails to report any results at a team level. Coward (2006), in reviewing various theories and research findings in the area, concluded that choking can well be the result of the fear of failure motive and the motive of presenting oneself favorably, where the former motive can elicit anxiety connected with thoughts of not being able to achieve the desired goal despite its being within easy reach, irrelevant thoughts tending to reduce the likelihood of success.

In Goldberg’s (1998) terminology, performance decrements are denoted as slumps. He maintains on the basis of practical experience that slumps start in the head of the athlete and are the result of a mismatch between the mind and body. Slumping athletes are considered to be negatively motivated by threats, fears and negative thinking. One of the key factors causing slumps is regarded as being competitive pressure, which makes the slumping athlete too nervous to achieve optimal performance. Goldberg noted further that in a team one player’s slump can quickly spread to the other players on the team, which can result in a devastating outcome. In conclusion, Goldberg maintains that slumps originate in cognition (negative thinking) and concomitant negative feelings at the individual level. On team level there is not much reported.

In attempting to explain the choking phenomenon, Halden-Brown (2003), an experienced coach and sport psychology instructor, argues that mistakes are more obvious in the case of bad performance than of good performance, and also that mistakes are linked to emotions, its being primarily the emotions and not the mistakes that have to be handled appropriately in order for optimal performance to be achieved.

It can be concluded, on the basis of a review of the models and the terms that relate to collective collapse, that both theorizing and research should mainly be directed at gaining an understanding of the processes at an individual level that are involved and of the role of both cognitive and emotional factors. There appears, at the same time, to be a lack of knowledge regarding the relationship here between psychological factors and performance. The fact that collective collapse is by definition a phenomenon at the group level implies there to be a social dimension that needs be taken into account if sudden decrements in team performance are to be properly understood.

In a recent review of the literature on psychological momentum, Crust and Nesti (2006) called for qualitative investigations in particular in this area, including examination of the individual experience of athletes, and of the cognitive, affective and behavioral changes associated with psychological momentum.

In the qualification series for the top division in handball in Sweden, Team A (the away team) was leading against Team B (the home team) by five goals with only 14 minutes left to play. Team A was unbeaten in the series and appeared to
clearly be on its way to victory, after having dominated the match from the start. Team B called for a time-out then, after which it went over to marking two of Team A’s top scorers man-to-man. Team B made seven goals in a row then, whereas Team A collapsed and, with 30 seconds left, missed a penalty shot that would have tied the match. Team B ended up winning by one goal. The aim of the present study was to investigate the thoughts, emotions, and behaviors of the players involved in the collective collapse of Team A.

**Method**

**Participants**

Nine handball players 18-33 years of age (mean age 23.1 years) from Team A, all of whom had played in the match just described, participated. These players, and the team to which they belonged, represented a convenience sample. Two of the players, who had left the team at the time of the interviews, were not included in the study.

**Procedure**

An interview guide consisting of 15 questions was constructed by the author for the purpose of gaining in-depth knowledge of factors connected with the occurrence of collective collapse. The questions were based on interviews with four experienced coaches and on the newspaper report of this particular match. In addition to personal background questions, there were questions concerning how the players had experienced the collective collapse that had occurred, and how they had reacted to it (their thoughts, emotions, and behavior). The coach and the key players were also asked how they reacted to the expression of emotions at the time by the coach and by their teammates. Due to time limitations and other practical considerations, it was not possible to test the interview guide prior to its use.

The interviews took place in the team’s hometown in May of 2006, four months after the collective collapse had occurred, a time chosen in part so as to not interfere with what was going on during the handball season. The participants were informed beforehand of the purpose of the study and agreed to participate. The interviews, lasting 40-60 minutes, were conducted by the author and by a master-degree student involved in research on collective collapse. The interviews began with a brief oral account of the match. The interviewers wrote down the participants’ answers to questions as exactly as possible. The three categories involved in the coding were the origin of a factor (one’s own team, the opposing team or external circumstances), the factors in question (cognitions, emotions or behaviors of various types), and their valence (positive, neutral or negative). The inter-rater agreement, measured using Cohen’s kappa, gave correlations of .79, .70, and .78, respectively. Disagreements were
discussed until the examiners reached consensus (Côté, Salmela, Baria, & Russell, 1993). Feedback to the interviewees was given upon request.

**Results**

*Thoughts and emotions when the team had a five-goal lead*

With 14 minutes left to play, Team A was leading by five goals. At this point, the thoughts and emotions of the players were obviously very positive. Eight of the players reported having only positive thoughts, and one player to have both positive and negative thoughts. The positive thoughts were expressed in such ways as the following: “Things looked good”. “It was an important match and we felt we were approaching victory, definitely not defeat”. “We were certain we were going to make it. There seemed no way we could lose”. The only player who appeared to have been doubtful said, “I felt the victory wasn’t quite safe yet, but that it was very close”. The emotions expressed were also largely positive, though less so than the thoughts that were reported. Expressions of positive emotions included the following: “Everything felt so good. We weren’t going to let this get away from us”. “It felt really good. We were the better team and were in control of the match. The opposing team seemed to have given up”. Negative feelings were expressed by two of the players: “We were tired”. “There were negative feelings we had. The pressure the fans were putting us under made us feel anxious”. Mixed feelings were expressed by one of the players: “We were excited, but also worried. We were in the lead, but it didn’t feel safe. The goals we’d scored had been easy ones, whereas the goals of the opponents were made were after good combinations and shots that were well-taken”.

*Perceived causes of the collective collapse*

The reasons subjects gave for collective collapse having occurred fell into two major categories: the behavior of one’s own team, and the actions of the opposing team. The behavior of one’s own team was seen as involving underperformance in the offensive play of nearly all of the players, fatigue, anxiety, and lack of experience. The underperformance, particularly of key players, was regarded as affecting the team as a whole. The remarks made included the following: “No one showed the ability to take the lead”. “We couldn’t change our play”. “Our play was characterized by the contagious effects of playing badly and each of us expecting someone else to take the initiative”. The sense of lacking experience was expressed, for example, as follows: “A young team is not as consistent in its performance as a team in which the players are older”. “The substitute players wanted to show off instead
of playing for the sake of the team”. Anxiety found expression in the statement, “We became anxious and began thinking about failure”. It was felt that the team was very much under stress through the opposing team being able to read off the offensive moves undertaken against them and to adjust their defensive tactics accordingly. One player remarked, “The stress we were under and our difficulties in anticipating the actions of the opponents resulted in our losing possession of the ball repeatedly and its being easy for them to score goals”.

Reactions of the players
The reactions of the players to the collective collapse that occurred involved the feeling of losing control of things. This was accompanied by largely negative thoughts, as well as negative emotions and negative behavior, i.e. behavior that failed to contribute properly to success of the team (See Figure 1).

![Figure 1. Reactions in the case of collective collapse.](image)

Although most of the thoughts that were cited were negative, two of the players mentioned positive thoughts: “We can still win the rest of our matches and end up ahead.” “Since I wanted to play, I thought mainly about myself and hoped I’d be selected to play” (a substitute player). For the most part, however, the thoughts of the players were negative: “What’s going to happen if we lose?” “We’ve been close to victory before but failed. Is this what’s about to happen again?” “Everything’s against us”. “As a junior player, there’s not much I can do”. Positive, task-oriented thoughts included the following: “I act on my own and do the best I can to challenge the opponents”. “The coach should call a time-out. We ought to be more disciplined in our play”. The emotions that were expressed were all negative. “Frustration” and “stress” were the words most frequently employed. Anger was expressed in one case: “I got angry at myself when I discovered how close the opponents’ score had gotten to ours”. Loneliness was expressed as well: “I felt completely deserted on the court and became paralyzed”. Lack of motivation was likewise expressed: “I had no
inspiration to play”. A sense of powerlessness was expressed too: “I knew I wouldn’t get any time on the court”.

The behavior subjects reported was classified as being either task-related, communicative, or irrelevant. The task-related behavior was of a sort that affected the play negatively, such as careless passes (“balls were being thrown away”), playing without making any real effort (“No one took responsibility”, “The play became stationary”), playing only as individuals (“The offensive combinations collapsed”), being regressive in one’s style of play (“We played like juniors”), and making wrong decisions (“The shots were taken too quickly and from unfamiliar positions”). Communication within the team was mainly negative, both in verbal terms (“We started to shout at each other”) and in terms of body language (giving expression to anger). There were reports of attempts to communicate constructively, but of these being unsuccessful (“I tried to encourage my teammates, but it didn’t work. They were all so down”. “I tried to give some hints on how to play”. “I tried to talk to the others and calm them down, but this didn’t succeed”). Irrelevant behavior was manifested by focusing on the referee instead of on what one was to do (“We got some calls against us and started to argue with the referee, although we knew it was pointless”).

Reactions of the coach and of the key players

According to the players, the coach was also taken by surprise by the opposing team’s change in tactics and reacted emotionally (“He was stressed, worried, and frustrated”. “He raised his voice and shouted”, “He was irritated by the calls of the referee”. “He was probably most stressed of all of us”). Answers to the question “What did he do?” gave a divided picture of things. Some players could not remember what the coach had done, whereas others remembered but were very critical (“He had no ideas. He seemed lacking in competence”). The coach called a time-out after two goals by the opposing team, which had reduced the team’s lead to three goals. The atmosphere during the time-out could be described as chaotic (“Everybody talked at the same time. It led to nothing. No encouragement was given”, “The coach didn’t know what to do”, “No changes at all were made”, “He wasn’t accustomed to a situation like this”). Some of the players questioned the decision to call a time-out (“A time-out in such a situation could result in anxiety and doubting one’s own ability.” “A time-out is only justified if it results in a goal immediately after the time-out”). Two of the key offensive players were assigned man-to-man, after which they became basically stationary in their actions. One key player had two yellow cards against him and had to play carefully to avoid being suspended. The general view of matters was that the team was taken by surprise and did not play well, and also that the offensive play failed when two of the players became marked. Shots were taken from odd positions because of stress.
The key players failed to fulfil their roles, and the man-to-man marking of two of the players clearly showed that, in the new situation that had come about, there was a lack of clarity regarding individual roles. The formal role of the team captain is that of his being the voice of the coach on the court and encouraging his teammates. He is usually an experienced player and is expected to be a good role model. The team captain in this particular match was described as trying to do his job, but failing at this (“He tried to coach us defensively, but he wasn’t alert”, “He tried to push us, but made sounds that gave kind of a whimpering effect and he was irritated”). His communications were seen as being mainly negative (“He shouted and bawled, which is negative in a match like this”), and his behavior was regarded as being directed at himself (“He focused on himself and on negative things. He became frustrated with the other players”. “He turned inwards. The team captain should be above the others, but he was like everybody else”).

Reactions to the expression of emotions by the coach and by teammates
Most of the players indicated that they were affected by the emotions of others on the team (“You can’t avoid it.” ”Yes, consciously or unconsciously. It’s contagious”. “Unconsciously, I’m affected by body language”. “The emotions of the coach affect me the most”. “As a young player I’m very much influenced by the others”). All of the players reported that negative emotions of the others had a clear influence on their own emotions and behavior. The reactions of the players varied from partial agreement with the idea that negative emotions are contagious (“It depends on how long it goes on. I try to focus on the next ball”) to complete agreement (“Of course I’m affected by any sort of negative atmosphere within the team”). There were also players who tried to do something about the negative emotions that were evident (“I tried to calm down the players who were upset, but it isn’t easy to change a whole team”). One player reacted with anger (“Why can’t they pull themselves together?”)

Discussion
It is scarcely surprising that, with a five-goal lead and only a short time left to play, most of the thoughts and emotions of the team members were positive. The players could readily interpret the match, both in cognitive and in emotional terms, as though they had already won it. A turning point in the match came when the opposing team changed its tactics. Although most of the players reported that the collective collapse that occurred then was due mainly to factors within their own team, some of their statements – such as “We didn’t succeed in changing the way in which we played“ or “No one turned out to be able to lead the team” – indicate that certain of the reactions they showed could be interpreted as responses to actions by the opponents. Fatigue, inexperience and the feeling of many of the players that it was someone else who should take the
initiative resulted in their exerting less effort than otherwise. The team’s offensive play, negatively affected as it was by what the opponents were doing, collapsed, whereas the defensive play appeared to not have changed considerably. Failure of the offensive play to function as it should resulted in an almost complete change in their thinking, emotions and behavior, which suddenly changed from being primarily positive to primarily negative in character. Their negative thoughts (“What’s going to happen if we lose?” “As a junior player, I can’t do much”), distracted them from the ongoing match and placed limitations on their behavior. Their negative emotions appeared to have a highly detrimental effect on their performance. The team was carried away emotionally instead of acting in a rational way. Noteworthy is that the few attempts to act rationally that they made, such as efforts to communicate constructively, failed. The players were not receptive. This could reflect their having too high an arousal level. The findings provide support for the views of Baumeister (1984) and of Coward (2006) regarding negative or irrelevant thoughts readily disrupting performance, and of Goldberg’s (1998) indicating performance slumps to start in the head of the athlete. They are also in line with Halden-Brown’s (2003) assertion that both frustration and anger need to be handled properly in order for top performance to be achieved. In the present case, role performance broke down – neither the coach, the captain of the team nor other key players performing in accordance with the roles they were to carry out.

The coach called a time-out but failed then to show effective leadership. Since he could not calm the players down and he gave them no apparent instructions, no change in tactics was brought about. The key players underperformed and the team captain was likewise unable to effect a change in the team’s performance. The lack of leadership was obvious. According to Carron, Hausenblas, and Eys (2005), the role of the team captain is not an easy one, the leadership demands placed on him often appearing overwhelming. In this particular match, the team captain failed to live up to the expectations placed on him, or to demonstrate a person-task compatibility. This suggests it could be a good idea, in some cases at least, to let two or more players share responsibility for the team captain role.

The few constructive initiatives taken were by players whose status appeared to be too low and were thus not listened to. Adequate role performance is highly important in team sports. Hagger and Chatzisarantis (2005) emphasize the necessity, for team success, of players fulfilling their roles. Carron, Hausenblas, and Eys (2005) underline the important role the coach has in the communication process. Team members tend readily, if the information provided them is unclear, to experience frustration and dissatisfaction and to act inappropriately. In the present case, the social function of creating a positive atmosphere failed.
It is not clear whether the team had appointed anyone as a kind of socio-emotional leader, but in any case no one was performing such a role adequately.

One can note that in various of the interviews negative emotions were reported to be contagious and to affect the entire team, both players on the court and those on the bench. This is a phenomenon that has also been reported by Totterdell (2000), who found the mood and performance of teammates on a cricket team to be contagious, the mood or performance of the one affecting that of the others. Since positive emotions readily lead to feelings of performing better and to an increase in self-confidence (Heath & Jourend, 1997), the present team could have profited from having a player who could take on a socio-emotional role for the team as a whole.

To conclude, it appears to be important to be prepared for at least those tactical changes by one’s opponents that are readily foreseeable, to withstand distracting thoughts and thus be able to focus adequately on the ongoing match, to have players who can take over important roles if the players designated to perform them fail to fulfill them properly, to always play as a team, to avoid being beset by negative emotions, and if possible to give someone the role of transferring positive emotions to the others.

**Concluding discussion**

Initially, three theoretical approaches seen as able to contribute to an understanding of collective collapse were presented. Janis’ (1982) theory of groupthink appears to be applicable to the decision making involved. In the match taken up in the study, the coach shouted at the players but failed to provide them constructive advice during the time-out. He seemed unable to adjust to the needs of the situation and, after the time-out, the passive play of the team continued. The results of the investigation as a whole support Kelly and Barsade’s (2001) findings of strong negative emotions (players’ shouting at each other) having a detrimental effect on group cohesion, as well as on commitment (loss in motivation), and performance (failing to play as a team), and of such effects being contagious, affecting the team as a whole. Support was also obtained for Totterdell’s (2000) finding of an increase in positive mood in the one team resulting in the mood of the opposing team becoming more negative. The present findings are also in line with Bion’s (1961) conceptions of assumption-group dependency (none of the other players taking over the roles of players who were performing unsatisfactorily) and of flight group (play having become passive).

The most firmly established model for the study of team sports would appear to be that of Carron, Hausenblas, and Eys (2005). This model suggests, in brief, that attributes both of the environment and of the team members form the basis for the physical and psychological structure of a team, which in turn affects those team processes that lead to the individual and team outcomes that are
achieved. Team cohesion is the variable seen as mediating between team structure and the processes involved.

Applying results of the present investigation to the conceptual framework of the model enables collective collapse be described as follows: Despite the attributes of a team, as far as the abilities of its individual members is concerned, tending to be stable during a given season and the physical condition of most of the players usually not changing much from one match to the next, the motivation of the players can change quickly, also during a single match. There can be a drop in motivation in the team which is in the lead, for example, if winning the match looks easy and this results in the team’s underestimating the abilities of the opposing team. This can possibly set the stage for collective collapse of the team which is in the lead if certain additional factors are present, in part factors specific to the situation, such as the actions of the opposing team, of the referee, and of the crowd, and the venue of the match. What the players of the opposing team do clearly has a stronger impact on whether collective collapse occurs than any of the other factors specific to the situation that were referred to. If the opposing team scores a goal, this gives that team a boost, and if the team also changes its tactics, this can confuse players on the team that had looked forward to winning easily and possibly even put them into a state of panic. The other factors just referred to that are specific to the situation appear to be of lesser effect, although some of the participants did direct attention at the need of dealing adequately with pressures produced by the spectators.

What follows then in the case of a collective collapse of the team which is or has been in the lead is a momentary disintegration of its team structure, of the communicative processes normally found, and with this of the usual team environment. The psychological structure of a team is a function of the position, the status and the roles of the various players and of norms that have been established. In a collective collapse, the structure of the team breaks down. Key players fail to do what is expected of them and the role system ceases to function properly. Team processes – in particular those of interaction, cooperation, communication, and decision making – also fail to function as they should. Constructive interactions become less frequent, and team members begin to play primarily as individuals, their also disregarding the tactics agreed upon. Communications become negative in character, players being irritated, beginning to shout at each other, and negatively affecting each others’ play. Decision making becomes poor, wrong decisions being made both by the players regarding such matters as when to shoot, and by the coach concerning questions such as when to substitute players. The achievement of the individual players declines radically due to recklessness, stress, and failure of the players to act as a team, the ultimate result being one of chaos and defeat.

The changes most frequently noted in connection with the collective collapse of a team were changes in the actions of the opponents and changes both in role
performance (team structure) and communication (team processes) within the team stricken by collective collapse. The actions of the opponents represent the only one of the three factors just mentioned that the collapsing team has no direct control over. Matters of the opposing team playing better than its ranking indicated, of its changing its tactics, and of its scoring an easy or unexpected goal were regarded as contributing to the collective collapse. Effects which collective collapse had on the team that was afflicted, and which were mentioned frequently included those of irritation, insecurity, fear of losing, shouting at each other, and chaos. The collapsing team can be seen as being faced with an increasingly difficult and discouraging situation, at the same time as the opposing team is encouraged by the events taking place, a situation described earlier by Totterdell (2000), who found that as the one team became more satisfied the other team became more and more dissatisfied.

One can suggest that a team, in order to minimize the negative influence on its play that the opposing team can have, should prepare itself tactically and mentally for those problematical actions by its opponents that are most foreseeable, and that it should train itself to react insofar as possible with positive rather than negative emotions under such circumstances. Halden-Brown’s (2003) advice is to express emotions that are genuinely appropriate to the situation, also when the going gets rough. Negative emotions such as shouting at teammates or disputing the calls of the referee distract from what needs to be done. The task at hand in many difficult situations is to refocus on the match, take position on the court and play in accordance with the tactics decided upon. Since negative emotions are usually not constructive, they should be avoided in a practical situation, both at an individual and at a group level. Remaining “cool” in the sense of not giving up, even when faced with a highly threatening situation, can be seen as the mark of a great athletic team.

The failure of players to perform the roles expected of them when collective collapse occurs could be noted in the underperformance of the key players in particular, together with the inability of the team captain to encourage his teammates, and the failure of the coach to provide clear instructions. The failure of players, under conditions of impending collapse, to do what is expected of them may be due to a lack of clarity of roles (players not understanding what behavior is expected of them, or the instructions provided them by the coach being too diffuse), to role conflicts (players lacking either the ability or the motivation needed), and/or to lack of role efficacy (players being insecure regarding their capacity to perform the role assigned them). Reasonable measures that could be taken to counteract such difficulties include the development of alternative role systems that can be employed in case of the underperformance of various players, such as players who are momentarily performing under par or who are temporarily restricted in the tasks they are to perform (such as marked players), the appointing of a team captain who is...
highly expressive emotionally and is particularly adept in influencing his teammates in a positive way (Hatfield, Cacioppo, & Rapson, 1994), and the adopting of a communication pattern that facilitates the achievement of mutual understanding.

The maintaining of adequate communication between team members on the court can be seen as particularly important and to be very much lacking in situations characterized by a high level of stress, by negative thinking, and by organizational disorder. In a situation such as that described above, in which the team is faced with a genuine crisis, constructive communication can readily decrease, communication both verbal and in terms of body language becoming negative in character. Despite its obvious importance, communication within sport teams during matches is an under-researched area within sport psychology (LaVoi, 2007; Sullivan & Feltz, 2003). It is important that someone present on the court assumes a leadership role for the team, helping the other players focus on the job at hand. It is also important that negative emotions and negative communication be avoided. The contagious effect of negative emotions readily results in negative communication, which can have a detrimental effect not only on those who are playing, but also on players on the bench. According to Barsade (2002), emotional contagion is largely an unconscious phenomenon. This suggests that an initial step to take in efforts to avoid it would be to make players aware of the phenomenon. The next step could be to endeavor to find ways of helping players, both those on the court and those on the bench, avoid negative emotions and avoid being affected by negative emotional contagion. A rational communication pattern needs to be established. LaVoi (2007) notes that certain interventions targeting communication have been successful and suggests the implementation of programs of this sort in order to improve communication patterns between the players on a team.

The non-constructive behavior that characterizes collective collapse is obviously preceded by cognitions and negative affects that stem from these. Taylor and Demick’s (1994) model for explaining positive or negative momentum, regarded as a highly useful model in this area (Mack & Stephens, 2000), postulates that momentum is the result of precipitating events, changes in both cognition and affects, and physiological changes connected with this, followed by changes in behavior, in performance, and in the outcome of efforts the person makes, without the model’s predicted what the direction of causal relationships is between cognitions, affects, and physiological changes. One can assume, on the basis of the present findings, there to be differences between how cognition, emotion and behavior are related to each other, depending upon whether the perspective taken is that at the start of a match, and thus prior to any possible collective collapse, though bearing in mind that in the course of the match collective collapse could occur, or at the close of a match, when a collective collapse, if it occurred, has already taken place. As suggested in
Figure 2, it can be assumed that, prior to a match, two possible causal chains that both are of problematical character exist, one of them of the sort negative thoughts leading to negative emotions, this resulting in a passive playing style, and the other of the sort positive thoughts leading to overconfidence, this resulting in mistakes and being followed by negative communication.

<table>
<thead>
<tr>
<th>Cognition</th>
<th>Emotion</th>
<th>Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative thoughts</td>
<td>Passivity (lack of effort)</td>
<td>Insecurity (fear)</td>
</tr>
<tr>
<td>(fear of losing)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive thoughts</td>
<td>Security (overconfidence)</td>
<td>Mistakes (and shouting)</td>
</tr>
<tr>
<td>(expectations of winning)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. The tentative relation between cognition, emotion, and behavior in collective collapse before the start of a match.

In contrast, as suggested in Figure 3, by the end of the match, various critical events can have occurred, such as mistakes by individual players, goals scored by the opponents, or whatever, that lead to negative emotional reactions followed by negative thoughts and result either in passivity and thus lack of effort, or further mistakes, followed by negative communications. Physiological changes can be the result of thoughts, of emotions or of actions that are carried out. Whether it is thoughts or emotions that appear first is more of academic than of practical interest. They appear to occur within close temporal proximity and can be assumed to influence each other.

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Cognition</th>
<th>Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frustration</td>
<td>Negative thoughts (fear of losing)</td>
<td>Passivity (lack of effort)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mistakes (and shouting)</td>
</tr>
</tbody>
</table>

Figure 3. The tentative relation between cognition, emotion, and behavior in collective collapse after critical moments during a match.

In conclusion, the major causes of a collective collapse are to be found partly in factors within the collapsing team. Such factors include those of inappropriate behavior, failure of the role system to function as it should, and negative communication. The causes are also to be found in the actions of the opposing team. These include their changing their tactics and their scoring of goals. Such external influences as false calls by the referee appear to be negligible in their effect. Collective collapse is characterized by a loss of control, accompanied by negative thoughts and negative emotions, all of which have a detrimental affect
on performance. The expression of negative emotions is also contagious, affecting the other players on the court and those on the bench.

**Limitations of the study**

The team selected constituted a convenience sample, the responses to collective collapse found to have occurred only reflecting with certainty the reactions to it of that particular handball team on the occasion in question. All the data in the study are also retrospective and may be subject to response bias. This is evidenced by the inability of some of the participants to recall what happened during the time-out that was taken. These factors represent limitations of the study carried out.

**Future studies**

In order that more adequate knowledge of collective collapse can be obtained, it is important that the occurrence and the perceptions of collective collapse be studied in a wide variety of different team sports and that intervention projects be carried out examining ways of counteracting or averting collective collapse and reducing its effects. Gender differences in this area are also very much in need of investigation. It can also be asked what men can possibly learn from women and women from men in this respect.

**References**


**Author note**

Erwin Apitzsch, Department of Psychology, Lund University.

This research was supported by grants from the Swedish Sport Research Council.

Thanks are due to Simon Granér for comments and assistance and to Susanne Persson for assistance in collecting the data.

Correspondence concerning this article should be addressed to Erwin Apitzsch, Department of Psychology, Lund University, PO Box 213, SE-221 00 Lund, Sweden. E-mail; Erwin.Apitzsch@psychology.lu.se
ORGANIZATIONAL FRAMEWORK AND PSYCHOTHERAPY SUPERVISION

Siv Boalt Boëthius & Marie-Louise Ögren

Abstract

The study aimed to highlight the functioning of the organizational framework supporting psychotherapy supervision at psychotherapy training institutions in Sweden. Data from course coordinators and supervisors was collected. Two questionnaires were constructed (Course Coordinator Questionnaire and Supervisor Questionnaire). Our results indicated that the organizational framework for different training institutions were similar with regard to the perceived primary task, i.e. supervision goals and overall framework. According to the course coordinators, there were certain differences with regard to role distribution, group composition, information, routines, regularity and accessibility. Economic and geographical conditions had a substantial influence on the design and functioning of the framework. The supervisors stressed the importance of explicit organizational frames, clear information and routines for, e.g. evaluation, how to deal with conflicts and failing, and above all the need for a competent and flexible course coordinator.

Keywords: Psychotherapy supervision, organizational framework, holding environment, course coordinator, supervisor

Introduction

This paper illuminates the design and functioning of organizational framework in psychotherapy training as a means of creating a holding environment for the learning process in psychotherapy supervision.

The concept of organizational framework can be used to indicate different organizational dimensions depending on the context. Formalization, as a dimension of organizational structure, describes to what extent explicit rules, routines, policies and procedures govern the work. It refers to who does what and how this work is performed (Abrahamsson & Andersen, 2005; Hatch & Cunliffe, 2006). Frames and routines in an organization can also be perceived as to what extent the organization is characterized by clarity and stability.

The concept of perceived organizational support (Rhoads Shanock & Eisenberger, 2006), which is used in connection with working life and work groups, could be perceived as a holding function of an organization when using a concept developed by Winnicott (1985). Based on Winnicott’s work, Amado and Ambrose (2001) have developed their theoretical model in cooperation with Harold Bridger. A key component in their theoretical framework is that changes in societal and organizational transitions involve changes both in social groups
and organizations as well as in the human beings and their concerns. Although the processes of change in these areas are quite different, they influence each other in a profound way.

Thus, a good enough holding environment in an organization consists, on a group and organizational level, of designing and arranging external conditions. On an individual level it entails being available and listening to individual needs. As described by Stapley (2006, p. 170), “We use it /the organizational holding environment/ to supply the same needs as the maternal holding environment and we apply the same emotion to it and create similar defences when it is seen as not good enough”.

In order to simplify the complex nature of the concept of an organizational holding environment Stapley suggests a division between what he calls external and internal holding environment. By the external holding environment is meant what is exposed or conscious, whereas the internal holding environment includes the subjective experiences of members of a group, organization or institution, i.e. the internal realities of the members of an organization (ibid. p.173).

In terms of an external holding environment for a training program certain main ingredients regarded as crucial to the framework are: task/s (supervision goals and principal framework), group composition (type of group formations), role distribution (leadership functions including role differentiation and delegation of responsibility), information (clarity), and regularity and accessibility (meetings and contacts with the outside world). Economic conditions as well as geographic and environmental factors are also important frame factors. In connection with a psychotherapy-training program other frame factors should be routines for evaluation and for conflict management as well as criteria for approval (Boalt Boëthius, 2003). The external holding environment, in terms of the organizational framework, could either be described by the course coordinators, as in the present study, or through documents concerning the framework of the program.

To obtain a picture of the internal holding environment the data consist of the subjective experiences and perceptions of the supervisors involved in the study in relation to the above-mentioned external organizational frame factors. However, the extent to which a certain training program could be considered as providing a good enough holding environment depends not only on the design and arrangement with regard to the external environment, or the capacity for containment. The interaction and mutual transactions between the two areas of interest are equally important.

**Organizational aspects of psychotherapy supervision**

A starting point for an analysis of the structure of the framework of a specific training program is the primary task the organization has to address. In an
organization where the main task is to stimulate and develop knowledge about psychotherapy, psychotherapy supervision as a special form of learning situation will be in focus.

Experience-based and emotional components have a prominent place in psychotherapy supervision. From a psychodynamic perspective the emotional component is a fundamental part of a process-oriented supervision. Unconscious levels are engaged and may give rise to both insecurity and resistance, as well as to various defensive attitudes (Boalt Boëthius & Ögren, 2000; Szecsödy, 1990). Personal learning and growth are therefore often associated with anxiety, which is why a secure and holding organizational base is needed (Boalt Boëthius & Ögren, 2003; Moxnes, 1991).

The present study is based on the assumption that the organizational framework affects the supervisors and their supervisees within a training program in a similar way as employees in working life. There is reason to assume that an important prerequisite for a creative learning climate in the supervision is that the supervisors have access to an organization that offers a professional and trustful structure and a holding environment (Brown & Bourne, 1996; Hughes & Pengelly, 1997; Rhoades Shanock & Eisenberger, 2006).

The importance of the organizational framework for psychotherapy supervision is discussed in the literature (Boalt Boëthius & Ögren, 2000; Ekstein & Wallerstein, 1977; Greenhalg, 2000; Proctor & Inskipp, 2001; Szecsödy, 1990). However, the impact of the organizational framework on psychotherapy supervision has until now been remarkably neglected as an area of research and systematic evaluation.

This study aims to highlight the impact of the organizational framework of psychotherapy supervision on advanced level training programs in Sweden. The study focuses both on the external holding environment, i.e. the actual design of the framework as described by the course coordinators, and the internal holding environment as it was experienced by the supervisors.

The main questions concerned how the course coordinators described the organizational framework (external holding environment) in terms of task, group composition, role distribution, information, regularity and accessibility, routines and economic and geographical aspects, and how these aspects of the framework were experienced by the supervisors (internal holding environment).

**Method**

*The training institutions*

The institutions involved in this study were all university affiliated professional psychotherapy-training units. The training units offered an advanced training program at postgraduate training level for psychotherapists. The programs were part-time stretching over a three-year period, and included theoretical seminars,
the supervisees’ own psychotherapy experiences and supervised clinical work. One to four different treatment orientations were represented at each training unit (e.g. child and adolescent, existential, family, cognitive, cognitive-behavioural, psychoanalytic and psychodynamic orientation). Both state and privately funded training programs were represented.

Research groups
Course coordinators. The course coordinators were responsible for designing and coordinating different parts of the training program. The course coordinators had a middle management leadership function. The course managements were responsible for the training program as a whole.

The sample comprised 18 course coordinators, 10 female and 8 male, from 13 (of 17 possible) training units. The average age of the course coordinators was 60 years (49 to 65 years). Most of them were licensed psychotherapists/psychoanalysts, with an average of seven years (one to 18 years) experience as a course coordinator. Half of the respondents worked at institutions that were funded by the state, i.e. universities; whereas the rest worked at institutions where the students paid for the training themselves (privately funded programs).

Supervisors. This sample comprised 27 (of 30 possible) psychotherapy supervisors, 17 women and 10 men. The average age was 53 years and 59 years, respectively. The supervisors had on average 14 years experience (three to 26 years), and had supervised at one to 10 training units. The supervisors had a psychodynamic or an existential orientation.

Instruments
Data from the course coordinators were collected via a Course Coordinator Questionnaire (CCQ). This questionnaire focused on the design and functioning of the framework in order to illuminate the external holding environment of the training program. The questionnaire consisted of open-ended questions. Two complementary interviews were performed.

Data from the psychotherapy supervisors were collected via a Supervisor Questionnaire (SQ). In order to illuminate the internal holding environment this questionnaire focused on the supervisors’ experiences of the design and functioning of the framework of different training programs they had been involved in as supervisors. The questionnaire included open-ended questions as well as ratings on a five-point scale (not very important [1] to very important [5]). The questionnaires were thus intended to generate data that highlighted the external (CCQ) as well as the internal holding environment (SQ).

Data processing
The responses from CCQ and SQ were transcribed and independently scrutinized by two of the authors. An open coding (Strauss & Corbin, 1998) was carried out by each examiner according to the research questions. Various possibilities for coding were tested separately and independently, and later discussed until consensus was reached. The analysis of data resulted in a number of themes related to the research questions.

In order to validate the course coordinators’ descriptions of the organizational framework, information and documents for the different training programs were checked by looking at the respective training program’s website. The agreement between the questionnaire responses and the existing documents was high.

**Results**

The results are presented under seven sub-headings: task, group composition, role distribution, information, regularity and accessibility, routines and economic and geographic aspects. Data from the course coordinators are presented first, followed by data from the supervisors.

**Task**

*Course coordinators.* The designing of a stable framework was considered to be the main task in order to create a climate that could facilitate the learning process in supervision.

The supervisors were prepared in various ways for their assignment. This could be by means of written descriptions about the program, the aim of the supervision, number of supervision sessions, supervisor meetings, criteria for selection of patients and for fulfilling the demands of the supervision. Information was also provided on a regular basis about schedules and changes. Some institutions also offered opportunities for their supervisors to take part in seminars within the field.

The choice of supervisors was in most programs a question for the course coordinators. In order to be employed as a supervisor one had to be an experienced psychotherapist and in most cases the supervisors were supposed to have a special two-year training program for psychotherapy supervisors. Another criterion was that potential supervisors should be well acquainted with the theoretical frame of reference of the program.

*Supervisors.* The main task for the supervisors in order to create good learning conditions for the supervision was formulated in terms of explicit and clear framework. This included information about routines regarding conflicts and failing. Regular supervisor meetings and easy access to a competent and flexible course coordinator were also stressed.
**Group composition**

*Course coordinators.* In general the course coordinators were responsible for the composition of the supervision groups or supervision pairs. However, some institutions, primarily the privately funded ones, and when the supervision took place far away from the training institute, the responsibility for the matching was left to the supervisees.

Most programs used specific criteria for matching supervisor and supervisee with regard to gender, professional background and the possibility of finding a time that would suit the parties involved. Some course coordinators mentioned several criteria, while others mainly stressed finding a common time and/or a geographically convenient place for supervision, as the most important factor/s.

Changes in group/pair compositions were rare in all programs. The changes that were reported had to do with illness, somebody leaving the program or conflicts that could not be solved otherwise.

*Supervisors.* The supervisors underlined the importance of being able to influence the matching of supervisor and supervisee/s. The course coordinator was perceived as having the overall responsibility for the matching of supervisor and supervisee/s.

As a result of the differences mentioned above in connection with the course coordinators, it was evident that there were large discrepancies between the institutions regarding the extent to which the supervisors and/or the supervisees could influence the composition of their supervision pair or supervision group. In line with this it is interesting to note that this frame factor received a low rating from the supervisors in comparison with other frame factors.

**Role distributions**

*Course coordinators.* There was a considerable variation between the different institutions regarding views about allocation of responsibility between various roles. In many programs both supervisors and course coordinators could have several roles, e.g. lecturer, seminar leader, examiner, researcher, and course paper supervisor. Having double or multiple roles was seen as quite problematic.

The responsibility for the examination of the psychotherapy supervision varied to a considerable degree between the institutions. The course coordinator was often responsible on the state funded training courses, whereas it was often the supervisor on the privately funded courses.

The responsibility for the treatment varied in a similar way between the institutions. It could either be the course coordinator, the course management, the supervisor or even the supervisee. It could also be the responsibility of the clinic where the treatment took place.

*Supervisors.* Clarity concerning the supervisor’s and the course coordinator’s areas of responsibility and mandate to take decisions was experienced as crucial,
as well as having clear definitions and delimitations for various role functions regarding formal responsibility. The supervisors emphasized the importance of having a clear mandate to assess the supervision process and client work. However, the general opinion was that the formal examination responsibility should lay with the course coordinator. It was seen as utmost important that the supervisor should be involved in as few other functions as possible in the training.

The course coordinator’s competence was stressed as a vitally important organizational prerequisite for the supervision. This necessitated a capacity for reflection, flexibility and ability to carry out both short and long-term planning. Also emphasized was the need for well considered strategies for acute situations that could arise for both supervisees and supervisors.

The course coordinator’s familiarity with both the supervision situation and the difficulties that could arise, as well as knowledge about the individual supervisors inspired trust and confidence.

The course coordinator’s capacity to contain the whole training situation was experienced as crucial. A competent course coordinator was expected to be able to take care of, protect and respect the organizational framework for the supervision. A successful course coordinator was expected to maintain the boundaries of the course (course management, supervisors and supervisees) as well as the boundaries of the program in relation to the organization as a whole, i.e. university department etc. and the outside world. The course coordinator’s ability to adhere to the agreed frames for supervision and client work was perceived as crucial in creating a profitable situation for the learning process.

Lack of clarity regarding roles and responsibility was experienced as having an extremely inhibiting effect on the supervision. This was transmitted as insecurity also to the supervisions in progress that were not primarily involved in the conflict situation.

Information

Course coordinators. Information to supervisors and supervisees was disseminated in many ways. One institution described how the supervisees and the supervisors received comprehensive written information, distributed at the start of the course, pertaining to, amongst other things, the criteria for passing the clinical component of the course. Despite this it was not unusual that both supervisors and supervisees later claimed that they had not received certain information.

Other channels of information reported were further training for the supervisors, joint literature studies, “acclimatization” via supervision assignments at basic level training and joint assessment of therapy sessions. Information to the supervisor about the selection process was also emphasized.
Supervisors: Clear information channels between course coordinator, supervisors and supervisees were emphasized. Of importance was also information about the clinical unit’s organization, routines and ethical issues regarding the treatment of clients.

It was considered important for the supervisors to have access to recurring information about current theory courses, and the supervisees’ theoretical knowledge and performance. When the supervisees were told one thing and the supervisor another, this created confusion in the supervision.

Regularity and accessibility
Course coordinators: Frequency and regularity of the supervisor meetings varied greatly. In general, supervisor meetings were held once or twice a term although at some institutions meetings were held more often. However, some training programs were not, due to economic or/and geographical reasons, able to arrange regular supervisor meetings.

The supervisor meetings focused on information to the supervisors, discussion of criteria for what the supervisees should learn, evaluations and supervision related problems. The focus could also be on factors hindering and promoting learning. At some institution supervisors were invited to group discussions with colleagues to raise individual questions concerning supervisees in order to hear the views of the other supervisors.

The regular institutional support to the supervisors varied. A few institutions did not offer any support, mainly due to economic factors. At other institutions the support to the supervisors was well developed through, e.g. supervisor meetings, supervision of supervision, further training, reading articles about training supervision, as well as external consultants.

Supervisors: Easy access to the course coordinator as well as the possibility of receiving support in the form of supervision of supervision was experienced as very valuable. Regular contact between course coordinator and supervisors (by means of supervisor meetings, information, contact opportunities with colleagues) was emphasized.

Frequent, regular and well-prepared supervisor meetings were much appreciated. Furthermore, supervisor meetings were supposed to involve the supervisors in current development work within the training and, moreover, to fulfil an important function in enabling course coordinator(s) and supervisor colleagues to discuss problematic supervision situations.

The importance of providing a scope for reflection at the supervisor meetings about current and problematic supervision scenarios contributed to recreating a holding environment for a learning climate. To receive support and help from a competent and insightful course coordinator was considered to strengthen the supervisor, the supervision group and the supervisee.
The organization’s concern for competence development for supervisors was highly appreciated. The need for recurring supervisor meetings with clear structure and content that met the supervisors’ need of discussing current issues and at the same time had a further training purpose was emphasized. Discussion of supervision cases for educational purposes was perceived as positive, as was the course coordinator’s feedback regarding current research in the area and development of the institution’s work.

Distance from and lack of contact with the course coordinator were experienced as having an inhibiting effect on the work in the supervision just as the lack of regular supervisor meetings. Uncertainty about the content in the training component of the course that was to be practiced in the supervised patient work was seen as creating uncertainty in the supervision.

Routines

Course coordinators: Routines for evaluation of the supervision’s quality varied. At certain institutions the evaluation took place chiefly through written assessments. In some cases the supervision was only assessed through the supervisor or the supervisee’s supervised work.

Routines for conflict management varied also between different institutions. Conflict management could result in a supervisee changing supervision group or receiving individual instead of group supervision. In order to prevent the development of serious conflicts close co-operation between the course coordinator, the course management, the supervisors and the supervisees was stressed.

Routines for failing were seen as a major factor. Supervisor meetings were emphasized as important for identifying the risk of a student failing the course. Some institutions were able to offer a new supervisor and/or a prolonged supervision when failing a supervisee. The importance of admission interviews was emphasized.

Supervisors: Clear routines within the organization were considered to be fundamental for the supervision process. This concerned all areas, i.e. established frames and routines for the supervision regarding frequency, length of supervision, clinical work, documentation, session reports, diagnostic and therapy summary reports and routines for evaluation and passing/failing.

Recurring evaluations imposed by the training organization were valuable for checking on current work in the supervision and for nurturing the learning space.

A readiness on the organization’s part, in the form of clear routines about how to manage conflict situations and antagonisms arising in connection with the supervision, was perceived as increasing the security for both the supervisor and the supervisee.
The lack of fixed routines clearly communicated between the course coordinator, supervisor, and supervisee was seen as contributing to splitting and acting out, which in turn reduced the learning space in the supervision. A lack of clear routines, that all parties are informed about, makes both the assessment of a student’s client work and supervision as well as failing more difficult in those cases where it was deemed necessary.

Economic and Geographical aspects

Course coordinators: There appeared to be a certain variation in the design and function of the organizational framework due to the training institutions’ economic situation and/or geographical location. Regularity in supervisor meetings turned out to be problematic when the economic situation was tight, and when the supervision as well as the clinical work was performed in another geographical area than that of the training institute.

If a student failed, the training institution’s economic situation was crucial if the supervisee had to be offered extra resources, e.g. continued supervision after the completion of the program.

As indicated above, the financial basis for the training program had a distinct influence on the design and the functioning of the program. One example of this was that the supervisors on the privately funded courses were often supposed to shoulder the whole responsibility for both the supervision and the final examination. On the state funded courses these functions were separated.

Supervisors: One advantage pointed out was that a privately funded training program could mean a less rigid bureaucracy. However, this could also easily lead to an increased insecurity and in turn to a decreased stability of the general framework. Clarity in role distribution and mandate, responsibility and continuity could be affected in a negative way.

When the training program was privately funded, and as a consequence the course coordinator was often involved in many different roles, the structure and clarity could be affected in a negative way. On the other hand, it was also argued that course coordinators as well as students on privately funded training programs could exhibit an extra strong commitment.

The supervisors reported that the economic interdependence between the training institute and the student/supervisor could be quite complicated. Supervisors experienced seriously questioning or failing a student on a privately funded program as more complicated regarding the consequences for the student as well as the training institute. In general, however, most supervisors were of the opinion that funding was not the most critical issue, provided the training institute had the competence to establish and maintain a clear and integrated framework where responsibilities, roles, mandates, routines and regular supervisor meetings were taken seriously.
When a training program had its resources cut back so that the number of supervision sessions were suddenly reduced, confusion and uncertainty arose in the supervision work and the creativity of the persons involved was affected.

**Discussion**

Our point of departure was that the external and internal holding environments are intimately woven together into a complex pattern. This complexity contributes to difficulties when striving for highlighting the external and internal environment separately. However, in this preliminary state of research the model of external and internal holding environment was considered to be of value in order to illuminate part of the complexity.

Our results indicated that, according to the course coordinators, the organizational framework (external holding environment) for different training institutions was fairly similar with regard to the perceived main task. All course coordinators showed an awareness of the importance of the framework for the psychotherapy supervision included in the training program. However, distinct differences between the institutions were reported regarding specific aspects of the design and function of the organizational frames. There were clear differences with regard to the management of role distribution, group composition, information, regularity and accessibility and routines, and the design of the framework was sometimes limited due to economic and/or geographical conditions.

According to the supervisors there was a consensus about the importance of the organizational framework in terms of external holding environment. With regard to the internal holding environment the supervisors stressed the course coordinator’s competence, flexibility and accessibility (e.g. regular supervisor meetings) as being crucial for a functioning supervision situation. Explicit information to all concerned parties and clear routines, e.g. for evaluation, passing/failing and conflict management were also considered to be of major importance. There was an overall consensus regarding the value of clarity in questions of frames, routines, roles and areas of responsibility. One factor emphasized by the supervisors was the importance of the total ability of the course coordinator to plan and keep the program on track. The mixture of roles was an obstacle to clarity concerning mandate and areas of responsibility that were regarded as prerequisites for creating a secure basis for supervision work. Economic or geographical obstacles could, in certain cases, contribute to the lack of maintenance of a satisfactory framework with consequences for the holding capacity of the training program. However, the course coordinator’s competence and capacity to contain the training program as a whole was of crucial importance.

One interesting aspect of the results was that the routines for group- and pair formation by the supervisor and supervisees varied between different
institutions. Certain training programs had carefully thought out routines for how a supervision constellation was created, whereas others could transfer responsibility for group formation and finding a supervisor to the supervisees (from the instructions on stipulated supervision competence).

Supervisors meant that certain participation in the composition of the supervision constellation were important but that ideally the ultimate responsibility for the composition of the supervision constellation should lie with the course coordinator. One may ask whether or not a training organisation primarily tends to be governed by trying to make the logistics of the supervision function practically at the cost of current knowledge regarding, e.g. principles for the composition of supervision groups (Boalt Boëthius, 1993; Boalt Boëthius, Sundin & Ögren, 2006; Ögren, Apelman & Klavitter, 2000).

The economic frame had evidently a certain importance for the framework as a whole concerning psychotherapy supervision, e.g. distribution of responsibility, information, supervisor meetings, and access to course coordinators. A well formalized framework seemed to be perceived as a quality control guaranteeing that similar situations were treated in the same way, in order to secure the situation for all parties: course coordinators, supervisors and supervisees (Abrahamsson & Andersen, 2005; Hatch & Cunliffe, 2006; Mintzberg, 1983). Moreover, a stable organizational framework may be a necessary basis for guaranteeing a satisfactory quality of the supervised clinical work and the patients being treated (Rönnestad & Reichelt, 1999; Watkins, 1997).

However, formalized training situations may have advantages as well as disadvantages. Apart from increased knowledge and work satisfaction there is a risk for a normative culture that promotes too much conformity. Clarity and transparency with regard to the organizational framework could limit this risk as, e.g. clear roles and routines tend to allow the individual supervisor and supervisee more freedom. It is likely that the design and function of the organizational framework could be even more important when supervision occurs in a group format, as so many persons and interactions within and between the supervision group(s) are involved.

The study has primarily focused upon the design of the organizational frames on advanced level psychotherapy training programs in Sweden. As these programs led to state licensing, they were subjected to meticulous state quality control at regular intervals. This could have contributed to the fact that the organizational frames were fairly well developed and planned.

The results of this study are limited to Swedish conditions but are in accord with previous international studies. The importance of a holding and supportive organization for the employees (in this case the course coordinators and the supervisors) so that they in their turn will be able to create favourable conditions
for their employees (in this case the supervisees) is emphasized by, for example, Rhoades Shanock and Eisenberger (2006) and Stapley (2006).

There is an evident need for further studies within this area in order to find out how to best use the organizational framework to achieve the goals of supervision. It would be of interest in a future study to examine to what extent and in what way treatment goals and acquired learning are related to specific aspects of an organizational structure and framework.

References


**Author Note**

Siv Boalt Boëthius, Ph.D., Professor em., Department of Education, Stockholm University, Sweden. Email: siv@ped.su.se and Marie-Louise Ögren, Ph.D., Associate professor, Department of Psychology, Stockholm University, Sweden. Email: mlon@psychology.su.se

Support for this research was provided by The Swedish Research Council, 721-2002-2586 (to Marie-Louise Ögren).

Correspondence concerning this article should be addressed to Marie-Louise Ögren, Department of Psychology, Stockholm University, Sweden. Email: mlon@psychology.su.se
DOES ANSWERING A QUESTIONNAIRE PROMOTE TRAFFIC SAFETY?

Birgitta Falk

Abstract

This study explores if answering a questionnaire on behaviour in traffic may decrease risky driving behaviour among young males. Two studies indicate that answering a questionnaire regarding personal risky driving behaviour may result in a decrease in self-reported risky driving behaviour some six weeks later. In Study 1 participants (193 men, 18-20 years old) also reported more concern about hurting others, increased subjective probability of accidents, but less thinking about injuries at follow-up. Results are discussed in terms of the question-behaviour effect, that is, questioning a person about a behaviour can influence the future performance of that behaviour. Assuming that most young male drivers essentially disapprove of traffic violations, it is argued that answering the questionnaire served as an intervention that made attitudes more accessible and led to a polarization towards stronger disapproval of traffic violations, which in turn influenced driving behaviour.

Keywords: Persuasion, self-report, young drivers

Background

Young drivers, and particularly young male drivers, are over-represented in traffic crashes. They also perform risky behaviours like, for example, speeding and reckless driving more often than other drivers do (e.g., Clarke, Ward & Truman, 2005; Boyce & Geller, 2002). It is, however, also acknowledged that such risky behaviours are difficult to change, and that this especially applies to high-risk groups (e.g., Glendon, 2007; Lewis, Watson, & Tay, 2007; OECD, 1994, Ulleberg, 2001). Factors connected to high crash risks of young drivers were illuminated in a recent comprehensive publication by the OECD (2006) where the urgent need for developing new countermeasures was also stated.

Studies from areas other than traffic psychology have shown that personal mental elaboration (i.e., reflective thinking on an issue) can be a powerful agent behind changes in attitudes as well as in behaviour (Petty, Haugtvedt & Smith, 1995). Research on the relationship between mental elaboration and attitude and/or behaviour change has used different methods for inducing elaboration. Some of the research concerns changes obtained without interference of any specific persuasive message or request for a change.

In an attempt to explore the potential of techniques based on reflective thinking, Falk and Montgomery (2009) tested the effects of making young men elaborate on the issue of personal negative consequences of risky driving behaviour. Effects on attitudes were measured by means of seven attitude scales
and effects on driving behaviour by means of a scale regarding self-reported risky driving behaviour.

There were some differences in attitudes between the experimental and control groups immediately after the intervention, but at follow-up some six weeks later the differences did not remain. Regarding self-reported risky driving behaviour, the study, however, revealed an unexpected finding. No effects on behaviour of the elaborative intervention could be established, but all three groups, the control group as well as the two experimental groups, reported less risk-taking behaviour at follow-up than they reported before the intervention, as can be seen in Figure 1 below. The change in self-reported driving behaviour over time was significant \((p<0.001)\) with partial \(\eta^2 = .18\).

![Mean scores on the Risky Driving Behaviour scale](image)

**Figure 1.** Difference in mean scores on Risky Driving Behaviour scale at pre-intervention and at follow-up for the experimental and control conditions in the study by Falk & Montgomery (2009). High scores indicate risky driving behaviour.

This was an intriguing result and it was decided to further explore its origins. It has been shown that unprompted reflection can both influence and polarize previously held attitudes (Tesser, Martin & Mendolia, 1995). To exemplify, in a study by Richard, van der Pligt and de Vries, (1996), changes in both attitudes and behaviour towards condom use were accomplished after participants had been asked seemingly unobtrusive questions about how they anticipated they would feel after having had unprotected sex with a stranger. More specifically, it has also been established that simply questioning people about a specific behaviour may influence their future attitudes as well as future behaviour – a phenomenon recently named the question-behaviour effect (Sprott et al., 2006). Could it be that simply responding to items in the questionnaire used by Falk and Montgomery (2009) had affected the participants’ subsequent self-reported driving behaviour? This would be in line with one theoretical underpinning of
the question-behaviour effect: that answering a question raises the accessibility of attitudes related to the question, and this in turn may result in behavioural change (Morwitz & Fitzsimons, 2004).

In order to investigate whether answering a questionnaire regarding personal driving habits and attitudes can lead to a self-reported decrease in risky driving behaviour two studies were conducted. Study 1 was done in order to replicate and substantiate the findings from the study by Falk and Montgomery (2009). Study 2 was conducted to explore what specific components of the questionnaire could have stimulated the decrease in self-reported risky driving.

**Study 1**

*Background and aim*

The main aim of Study I was, as mentioned above, to replicate the findings regarding the decrease in self-reported risky driving behaviour found in the study by Falk and Montgomery (2009). I especially wanted to control for possible effects of having met the researchers for an interview, of different contexts when answering the initial and follow-up questionnaires (at the enrolment centre/regiment versus at home) and of unpredictable external factors in the period between answering the initial and follow-up questionnaires.

*Method*

A total of 193 males with a driver’s licence participated in the study, of which 142 completed it satisfactorily. Of the participants, 99.3% were 18 or 19 years old and all had a driver’s licence. The questionnaire used was identical to the one used in the study by Falk and Montgomery (2009). The scale measuring driving behaviour, the “Risky Driving Behaviour scale”, was composed partly of items from the “violations” factor in the Swedish version of the Swedish Driving Behaviour Questionnaire (DBQ-SWE, Åberg & Rimmö, 1998), and partly of items developed and used in a study by Ulleberg and Rundmo (2002). The participant was asked to indicate how often (on a Likert-type response scale with end points 1= never to 5= very often) he committed 21 different violations of traffic rules.

Attitudes towards risk-taking behaviour were measured by three scales, developed and validated by Ulleberg and Rundmo (2002). The three scales were: “Speeding” (5 items), “Funriding” (3 items) and “Traffic Flow vs. Rule Obedience” (9 items). In the study by Ulleberg and Rundmo (2002) these scales had the highest correspondence with self-reported behaviour, predicting between 12% (“Traffic Flow vs. Rule Obedience”) and 18% (“Speeding”) of the
respondents’ variability in risk-taking behaviour, and had reliabilities exceeding Cronbach’s $\alpha = .75$.

Attitudes towards accidents and injuries in traffic were measured by four scales. “Risk of Accidents” (3 items) was constructed and validated by Ulleberg and Rundmo (2002). “Injury Reflection” (3 items) was constructed specifically for the study by Falk and Montgomery (2009) and contained items regarding reflection on possibilities of hurting oneself or others in traffic. The third scale, “Concern about Hurting Others” (5 items) was based on a scale developed by Ulleberg and Rundmo (2002) to which two other items were added. Items on these attitude scales were rated on a Likert-type five-point scale with the terminal points 1 = ‘Completely disagree’ and 5 = ‘Totally agree’. The fourth scale “Subjective probability of accidents or mishaps in traffic” (5 items) was also constructed specifically for the study by Falk and Montgomery (in press). It contained questions regarding the subjective probability of accidents or mishaps in traffic during the next three years: a collision, oneself being injured, injuring someone else, one’s friend/family being injured, and being caught speeding. These items were rated on a five-point scale with the terminal points 1 = ‘Totally unlikely’ and 5 = ‘Very likely’.

Participants were asked to volunteer for the study by staff at three centres of the Swedish National Service Administration when enrolling for compulsory military service. They were assigned to one of three conditions. In condition (1) “Enrolment – Home” (EH) the initial questionnaire was completed at the enrolment centre and the follow-up questionnaire at home, in conformity with questionnaire administration in the Falk and Montgomery (2009) study. In condition (2) “Home-Home” (HH) the initial as well as the follow-up questionnaires were completed at home. The follow-up questionnaire was administered approximately four weeks after the initial questionnaire. In the third condition (3), “Sole questionnaire” (S), only the initial questionnaire was completed, but at the same point in time that the follow-up questionnaire was sent to participants in the other two conditions. The purpose of condition S was to control for the influence of external factors on conditions EH and HH from completion of the initial questionnaire to the completion of the follow-up questionnaire. Participants in condition S thus did not receive a follow-up questionnaire. The questionnaires used were identical to those in Study II. An overview of the study design and principal measures used can be found in Figure 2:
Results

Self-reported risky driving behaviour at initial and follow-up measurement. In order to test whether there was a main effect of time a factorial 2 (time: initial measurement, follow-up) x 2 (condition: EH, HH) repeated measures ANOVA was conducted. There was a main effect of time, $F_{(1, 91)} = 16.16; p < .001$, partial $\eta^2 = .15$. Thus, the decrease in self-reported risky driving behaviour found in Study II was replicated without any intervention or personal contact with researchers, and the results indicated no effects due to different contexts when filling out the initial questionnaire (at the enrolment centre or at home). There were no significant differences between conditions, either regarding the initial questionnaire or at follow-up. Nor was there any significant interaction between condition and time.

The possibility that the influence of external factors could explain the decrease in self-reported risky driving behaviour was also ruled out. There was a significant difference between the scores on the Risky Driving Behaviour scale for condition S (whereby the questionnaire was completed only once, but at a point in time corresponding to when conditions HH and EH answered their follow-up questionnaire) versus the scores for conditions EH and HH at follow-up, as displayed in Figure 3:
Figure 3. Means of scores on the Risky Driving Behaviour scale for the three conditions. High scores indicate risky driving behaviour. Note that the questionnaire was administered only once to group S, at a point in time corresponding to the administration of the follow-up questionnaire to groups EH and HH.

Attitudes towards risk-taking at initial and follow-up measurement. A doubly-multivariate analysis of variance was conducted to investigate whether attitudes had changed from initial measurement to follow-up. (See Tabachnik & Fidell (2001) for a full description of the MANOVAs used in the present investigations.) Independent variables were time (initial measurement and follow-up) and condition (HH and EH). Six attitude scales – “Speeding”, “Traffic flow vs. rule obedience”, “Funriding”, “Concern over hurting others”, “Injury reflection” and “Subjective accident probability” – were included as dependent variables. (The scale Risk of accidents was excluded from analysis due to low reliability.) There was a significant main effect of time F(6.85) = 8.37 p < .001; Wilks’ Lambda = .628, but no main effects of condition. Tests of within-subject contrasts showed a significant main effect of time for three of the scales: “Concern for hurting others”; F(1) = 11.57 p < .001, “Subjective Accident probability” F(1) = 8.64 p = .004 and “Injury reflection” F(1) = 8.42 p = .005, meaning that over time participants claimed to be more concerned over hurting others and experienced a higher subjective probability of accidents, but reported thinking less about the possibility of injuries at follow-up than at the initial measurement.

Discussion and conclusion
The decrease in self-reported risky driving behaviour found in the study by Falk and Montgomery (2009) was substantiated. I therefore decided to conduct Study
2 in order to explore what specific components of the questionnaire could have stimulated the decrease.

Study 2

Background and aim
The results from Study 1 were interpreted as a question-behaviour effect (Sprott et al., 2006) resulting from completing a questionnaire regarding behaviour and attitudes related to risky driving. The main aim of Study 2 was to investigate whether the change in self-reported risky driving behaviour found in Study 1 would be elicited by simply asking questions about personal driving behaviour, or whether additional questions about attitudes relating to risky driving were needed to obtain it. In order to accomplish this aim, I used three different versions of the initial questionnaire and studied their effects on the Risky Driving Behaviour scale at follow-up.

Method
A total of 149 male participants with a driver’s licence signed up for participation in the study and 133 completed it successfully. Of the participants, 95% were 18 years old. They were recruited and assigned to conditions (i.e., version of the initial questionnaire) by the staff at one of the offices of Swedish National Service Administration. All participants completed the initial questionnaire at the enrolment centre.

The first version of the initial questionnaire, condition “Driving Behaviour Only” (DBO) consisted of only the “Risky Driving Behaviour scale”, i.e., questions about personal driving behaviour.

The second version of the questionnaire, condition “Attitudes to Risk-Taking” (RiskAtt) consisted of the Risky Driving Behaviour scale plus the three scales regarding attitudes towards various risk-taking behaviours in traffic. The rationale behind studying the impact, if any, of this version was as follows: As a consequence of being exposed to questions about actual personal risky driving behaviour as well as questions about his attitudes towards similar behaviour, the participant may become aware of discrepancies between his attitudes and behaviour. This in turn might lead to a change in self-reported risk-taking behaviour. Such discrepancies may evoke cognitive dissonance and motivate changes in behaviour or attitudes, as shown in a number of previous studies (for an overview see, e.g., Olson & Stone, 2005).

The third version of the initial questionnaire, condition “Attitudes to Accidents” (AccAtt) consisted of the Risky Driving Behaviour scale plus four scales regarding attitudes towards accidents and injury in traffic. The rationale behind using this version was that questions about consequences of risky driving behaviour might raise the participants’ awareness of negative affective consequences of risky driving, and thus motivate decreased risk-taking. This
would be in line with the results from the study by Richard et al., (1996) who found effects on attitudes as well as behaviour by raising awareness of the negative affective consequences of unsafe sex.

The follow-up questionnaire was administered some four to five weeks after completion of the initial questionnaire. The follow-up questionnaire only contained the Risky Driving Behaviour scale, no attitude scales.

**Results**

To investigate whether the change in self-reported risky driving behaviour found in Study 1 would be affected by the content of the initial questionnaire, a 2 (time: initial questionnaire, follow-up questionnaire) x 3 (condition: DBO, RiskAtt, AccAtt) repeated measures ANOVA was conducted. The Risky Driving Behaviour scale was the dependent variable. There was a main effect of time, $F_{(1, 128)} = 35.76; p < .001$, partial $\eta^2 = .22$. Separate one-way repeated measures ANOVAs showed that the change in scores from initial questionnaire to follow-up questionnaire was significant for all three conditions at $p = .001$. In other words, all conditions reported a decrease in risky driving behaviour.

However, there was a significant overall difference, $F_{(2, 128)} = 5.25; p = .006$, between conditions on the Risky Driving Behaviour scale, as can be seen in Figure 3. The effect size was not large (partial $\eta^2 = .08$) and pairwise comparisons (Bonferroni correction) revealed that the difference was significant only for condition DBO versus condition RiskAtt. For these two conditions the effect was significant regarding the initial questionnaire as well as the main effect of time. Concerning the follow-up questionnaire, there were no significant differences on the Risky Driving Behaviour scale between any of the conditions.

![Figure 4. Means of scores on the Risky Driving Behaviour scale at the initial and the follow-up questionnaire for the three conditions. High scores indicate risky driving behaviour.](image-url)

**Discussion and conclusion**
Study 2 was conducted to explore if specific components of the questionnaire could have caused the decrease in self-reported risky driving. The decrease in self-reported risky driving behaviour was replicated again and was significant for all three conditions, i.e., regardless of the content of the initial questionnaire. Thus the decrease was caused by simply asking questions about personal driving behaviour. The additional questions about attitudes relating to risky driving did not add significantly to the effect. The comparatively high score on the Risky Driving Behaviour scale on the initial questionnaire for the DBO condition is difficult to explain. The only difference between conditions regarding background factors between was place of domicile, but analyses revealed no general differences between urban and provincial participants regarding driving behaviour. One possible explanation could be that, in conditions RiskAtt and AccAtt, the attitude questions influenced the response to the behavioural questions on the initial questionnaire. I find this explanation less likely as the attitude questions were posed after the behavioural questions, but Schwarz and Hippler (1995) have shown that regarding self-administered questionnaires subsequent questions can influence preceding questions as respondents may go back and forth between questions. Nevertheless, whatever the reason for the high scores of condition DBO on the Risky Driving Behaviour scale, it does not obscure the fact that in all three conditions the scores had changed in the expected direction at follow-up.

**General discussion and conclusion**

The two studies presented in this article show that answering questions regarding personal risky driving behaviour leads to a self-reported decrease in risky driving behaviour about one month later. This is a robust finding, further supported by the results reported by Falk and Montgomery (2009).

A number of questions remain to be explored though. The first and perhaps most important one is whether the decrease in self-reported risky driving mirrors safer actual driving behaviour. However, several studies have found self-reports of driving behaviour being predictive of accidents (e.g., Hatakka, 1998; Iversen & Rundmo, 2004; Parker, Reason, Manstead, & Stradling, 1995; Parker, West, Stradling, & Manstead, 1995; Ulleberg, 2001) as well as being in fair agreement with observed behaviour (Lajunen & Summala, 2003). Nevertheless, it is not unreasonable to infer that the self-report measure of risky driving behaviour used in the present studies actually served as an intervention in itself. As Campbell (1966) has pointed out, this represents a threat to the validity of self-report measurement. It also underlines the importance of using a control group when evaluating the effects of experimental interventions by means of self-reports.
Another important question that remains to be answered is why the decrease in self-reported risky driving behaviour occurred. It has long been acknowledged that test-retest designs may lead to changes in scores from time 1 to time 2, without any intervening experimental variable (Campbell, 1996; Windle, 1954) or any other kind of external variable that can explain the change. Changes that seem to be the result of repeated self-report measurement, so-called test-retest effects, have been reported in many different contexts. The term refers to a tendency of respondents in the control group as well as the experimental group to report more well-being or fewer symptoms when retested (e.g., Arrindell, 2001; Durham et al., 2002; Fendrich & Yun Soo Kim, 2001; Longwell & Truax, 2005; Stice, Chase, Stormer, & Appel, 2001). Although a number of explanations regarding the origin of the phenomenon have been proposed, the picture of its nature and causes is still far from clear (see, e.g., Arrindell, 2001; Knowles, Coker, Scott, & Cook, 1996 for overviews). The question-behaviour effect (Sprott et al., 2006) and the impact of self-generated attitude change, as discussed below, constitute other likely interpretations of the test-retest effect and deserve to be explored further.

Increased accessibility of attitudes and beliefs as an effect of answering questions regarding personal driving behaviour constitutes the most likely explanation of the decrease in self-reported risky driving behaviour in my view. Answering questions in a questionnaire requires accessing and scrutinizing issue-relevant internal knowledge, that is, personal reflection. Previous research has shown that merely thinking about an issue tends to make evaluations connected to it more extreme, accessible and stable (Tesser, Martin, & Mendolia, 1995). One reason may be that thinking changes the cognitive representation related to the issue (Valenti & Tesser, 1981) and affects the attitude structure (Millar & Tesser, 1986) which in turn may change beliefs regarding a specific behaviour. Fishbein, Ajzen, and McArdle (1980) argued that changing underlying beliefs, or making certain beliefs more salient, should lead to changes in attitudes, thereby influencing intention and possibly behaviour. Thinking about an issue may also continue after a task has been completed, producing entirely new thoughts during the process (Sadler & Tesser, 1973). Taken together, assuming that most young male drivers essentially disapprove of traffic violations (as portrayed in items of the Risky Driving Behaviour scale), answering the questionnaire could have led to more accessible and polarized attitudes about the issue. This in turn could have influenced driving behaviour. Using the analogy of an inner dialogue when answering the first questionnaire, the process could be verbalized as: “Do I often overtake the car in front when it is driving at the speed limit? Well, it does happen. 5 means very often, 1 means never – so I’ll set my mark at 3. Hmmm. Overtaking a car driving at the speed limit is really not a very good thing to do. I really shouldn’t do it.”
One way of approaching differences in scores between time 1 and time 2 in longitudinal studies was suggested by Golembiewski, Billingsley, and Yeager (1975), who defined three different types of change that could be reflected in the change scores. In view of the present study, the concepts of Alpha and Beta change suggested by Golembiewski et al. (1975) are interesting to discuss. Alpha change, according to Golembiewski et al:

> involves a variation in the level of some existential state, given a constantly calibrated measuring instrument related to a constant conceptual domain. (Golembiewski et al., 1975, p. 134)

In other words, Alpha change represents a “real” effect of an intervention. Beta change, on the other hand, according to Golembiewski et al:

> involves a variation in the level of some existential state, complicated by the fact that some intervals of the measurement continuum associated with a constant conceptual domain have been recalibrated. (Golembiewski et al., 1975, p. 135)

In other words, for some reason the individual’s “subjective yardstick” has changed from first measurement to follow-up. This phenomenon has also been called response-shift bias (Howard & Dailey, 1979) and instrumentation bias (Campbell, Stanley & Gage, 1963). In the present context it would mean that while the behaviour itself has not changed, it is judged in a different way. Again using the analogy of an inner dialogue when answering the questionnaires, the process could be verbalized as follows. At time 1: ”Do I often overtake the car in front when it is driving at the speed limit? Well, a couple of times a week. Quite often really. I’ll mark 4.” At time 2: ”Do I often overtake the car in front when it is driving at the speed limit? Well, a couple of times a week. Not often really. I’ll mark 2.” Reasons behind the change in the “subjective yardstick” can only be hypothesized. Thus, while I consider it likely that the reported decrease in risky driving behaviour is a result of (in the terms of Golembiewski et al., 1975) Alpha change, the possibility that it actually represents a Beta change deserves further research.

It should be pointed out that sometimes the effects of repeated measurement may mistakenly be referred to as regression towards the mean. Regression towards the mean occurs when extreme groups (e.g., much better or much worse than average) are selected on the basis of a time 1 measurement (or because they for some other reason have extreme scores). When these groups are measured again at time 2, the better ones will on average score worse and the worse ones will on average score better. Both will therefore be closer to the (population) “grand mean” at time 2. However, this regression does not affect the grand mean if the extreme groups are randomly assigned to conditions (Bland & Altman, 1994a, 1994b).
A general conclusion from the present studies is that being unobtrusively stimulated to think about an issue might affect both the belief structure and the attitudes of an individual and make both more accessible. The phenomenon has been touched on by several authors and from different theoretical angles (e.g., Feldman & Lynch, 1988; Knowles & Condon, 2000; Loken, 2006; Robinson & Clore, 2002; Wyer & Albarracín, 2005). Research on interventions building on unprompted reflection is still in its infancy. Further research in order to elucidate the potential of such interventions to increase traffic safety is much needed.

References


Author Note

Birgitta Falk, Department of Psychology, Stockholm University, Sweden.

This research was financed by the Swedish Road Administration and Länsförsäkringars Forskningsstiftelse.

I am grateful to Professors Henry Montgomery and Ola Svenson for valuable comments on earlier versions this manuscript. Thanks also to the staff at the Swedish National Service Administration who enabled data collection.
PREREQUISITES FOR MEANINGFUL GROUP WORK - STUDENTS’ EXPERIENCES OF CO-OPERATION

Eva Hammar Chiriac & Kjell Granström

Abstract

Which features do students emphasize when they describe constructive and destructive group work? This was the research question of the present study. Data was collected through focus group interviews with students from three different schools. The results reveal that the students consider six preconditions of importance for their conception of group work: organisation, mode of working, task, report, assessment and the role of the teacher. To sum up, what the students ask for is participation, quality and relevance (PQR).

Keywords: Group work, co-operation, learning conditions

Group work is used as a common method at all levels of the educational system in Sweden, from pre-school classes to university courses. Group work is used as an intended means for learning, but may also serve as an opportunity for creating social relations with schoolmates. However, teachers often use group work without considering its prerequisites. Such a mode of working may sometimes lead to positive experiences and learning, while in other cases a reversed outcome. Even though there is a lot of research carried out on group work there is still an important aspect that need to be elucidated further, and that is the participants’ conceptions of group work.

The aim of this study is to address the students’ experiences and conceptions of group work. What prerequisites do students ask for when they describe good group work, and what conditions do they think result in a bad outcome?

Definition of group work in school

The students seem to have a clear conception on what constitutes as a group work and what does not. According to the students, group work is characterized by co-operation among students on an assignment from the teacher. Group work is described as working together with classmates on a common task. The following excerpt may serve as an example of the students’ depiction of group work:

But it is group work when you get a special task (5).

The students distinguish between group work (as described above) and collaboration, where the latter is more of helping a classmate sitting next to you.
Students are working on their own task and the choice to collaborate is their own. The reply below is an example:

When you, sort of, assist one another or collaborate you might … [help] somebody with a number in maths, you might not do that every lesson, but just this time (5).

The quotations also reveal that group work is something you do in a more co-operative way during an allotted period of time while assisting a classmate is a more spontaneous and temporary activity. Table 1 summarises the different modes of working based on the students’ descriptions.

Table 1

<table>
<thead>
<tr>
<th>Aspects</th>
<th>Co-operation</th>
<th>Collaboration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>Work together</td>
<td>Help each other</td>
</tr>
<tr>
<td>Content</td>
<td>A common task</td>
<td>Individual task</td>
</tr>
<tr>
<td>Time</td>
<td>Planned</td>
<td>Spontaneous</td>
</tr>
<tr>
<td></td>
<td>Allotted time</td>
<td>Occasionally</td>
</tr>
</tbody>
</table>

The conclusion is that group work, as described by the students, is characterized by co-operation on a common and planned task during an allotted period of time. The purpose of this study is to investigate conditions for such co-operation in a school setting, and to summarize the prerequisites for fruitful work in groups.

**Method**

Data was collected by use of six focus groups (Bryman, 2001). Since the aim of this study was to explore a specific theme, focus groups were considered as an effective and well-tried method for collecting data (Millward, 2007).

The focus group interviews were carried out with students from three different schools in the south of Sweden (Table 2). Schools from different social and geographical areas were included and two focus groups of students from each school participated in the study. The class teacher accomplished the selection of students to participate in the focus groups but in some cases there was a substitute teacher in the class and the choice was based on voluntaries. The research group consisted of, in all, 41 informants, 21 male and 20 female. Each
focus group lasted a mean of 40 minutes, which means a total data source of four hours.

Table 2

Description of participants

<table>
<thead>
<tr>
<th>Focus group</th>
<th>School</th>
<th>Form</th>
<th>Participants</th>
<th>Fe/M</th>
<th>Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>U-school</td>
<td>7</td>
<td>7</td>
<td>4 / 3</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>U-school</td>
<td>9</td>
<td>8</td>
<td>2 / 6</td>
<td>36</td>
</tr>
<tr>
<td>3</td>
<td>P-school</td>
<td>8</td>
<td>6</td>
<td>3 / 3</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>P-school</td>
<td>8</td>
<td>6</td>
<td>3 / 3</td>
<td>38</td>
</tr>
<tr>
<td>5</td>
<td>K-school</td>
<td>8</td>
<td>8</td>
<td>6 / 2</td>
<td>42</td>
</tr>
<tr>
<td>6</td>
<td>K-school</td>
<td>8</td>
<td>6</td>
<td>3 / 3</td>
<td>41</td>
</tr>
</tbody>
</table>

Table 2 displays more information about the participants and shows the allotment of female and male students taking part in each focus group. In all the numbers of female and male students were equal.

When all six focus groups were accomplished and transcribed, a qualitative content analysis was carried out (Bryman, 2001, Millward, 2007). The qualitative content analysis, which emphasised the construction of meaning, was carried out in several steps by two independent researchers. After a closer examination, six major factors turned out to be important for the understanding of the students’ opinions of group work in school. A comparison with a control question revealed congruency.

This study has followed BPS’ four ethical principles based on (a) respect, (b) competence, (c) responsibility and (d) integrity (BPS, 2006). The ethical principals emphasize the concern for participants’ interest, i.e. informed consent, confidentiality, withdrawals and use of material (APA, 2002).

Results

In accordance with the purpose to catch the students’ opinions of prerequisites for meaningful group work and experience of co-operation, the major part of the result presentation will deal with the students’ descriptions of high-quality and low-quality group work in school. In order to find out if the results can be verified by findings from previous and recent scientific studies concerning group work in schools and in general, each aspect mentioned by the students will be commented on in the concluding discussion of the results, in order to validate the results and to ascertain the relevance for daily work in schools.

What students perceive as high and low quality group work

The analysis revealed six key factors as being significant for group work. Each factor may, according to the students, either facilitate or inhibit group work. The
six factors as they derived from the data, were (a) organisation, (b) mode of working, (c) task, (d) report, (e) assessment and (f) role of the teacher. The informants considered that these factors were important to the process and the outcome of group work. Certainly, they did not present their opinions consciously in these six fields. The categorization is a result of the content analysis carried out by the researchers.

In Table 3 the six factors are presented in an abbreviated form. Each factor describes prerequisites that may promote high quality as well as prohibit successful working.

Table 3
Factors and aspects promoting and inhibiting high quality group work
## Factor/aspect

<table>
<thead>
<tr>
<th>Factor/aspect</th>
<th>Promoting co-operation</th>
<th>Inhibiting co-operation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ORGANISATION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Group size</td>
<td>2-5 students</td>
<td>More than 6 students</td>
</tr>
<tr>
<td>• Group composition</td>
<td>Not too heterogeneous</td>
<td>Too heterogeneous</td>
</tr>
<tr>
<td>• Time</td>
<td>Suited to task</td>
<td>Too short time</td>
</tr>
<tr>
<td>• Location</td>
<td>Calm</td>
<td>Noisy</td>
</tr>
<tr>
<td><strong>MODE OF WORKING</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Focus</td>
<td>On the task</td>
<td>On other things</td>
</tr>
<tr>
<td>• Participation</td>
<td>All members</td>
<td>Some members</td>
</tr>
<tr>
<td><strong>TASK</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Intelligible</td>
<td>Comprehensible</td>
<td>Incomprehensible</td>
</tr>
<tr>
<td>• Stimulating</td>
<td>Interesting</td>
<td>Boring</td>
</tr>
<tr>
<td><strong>REPORT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Content</td>
<td>Understandable</td>
<td>Messy</td>
</tr>
<tr>
<td>• Performance</td>
<td>Lively</td>
<td>Hard to follow</td>
</tr>
<tr>
<td>• Responsibility</td>
<td>All group members</td>
<td>Just a few members</td>
</tr>
<tr>
<td><strong>ASSESSMENT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Transparency</td>
<td>Known criteria</td>
<td>Hidden or no criteria</td>
</tr>
<tr>
<td><strong>ROLE OF THE TEACHER</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Arranger</td>
<td>Promote co-operation</td>
<td>Fail to structure</td>
</tr>
<tr>
<td>• Supporter</td>
<td>Being attendant</td>
<td>Abandoning the students</td>
</tr>
</tbody>
</table>

The purpose of Table 3 is to present the result in a clear way, easy to grasp. Certainly, this is a simplification, but the table elucidates that the crucial factors put forward by the students, applies to both good and bad group work. Below, each factor will be presented and illustrated by excerpts taken from the entire data record.

**Factor I: Organisation**

As shown in Table 3, there are four main aspects, *size, composition, time and location*, constituting the factor of organisation. Organisation concerns the issue of group composition, which is one of the factors that seem to have a great impact on the groups’ work, process and outcome. Each of these aspects will be commented on with respect to its promoting or inhibiting functions.
**Group size.** It is evident that a group with about three members is considered to be the optimal size. The students consider larger groups, i.e. with more than six members a hindrance to good group work. Some quotations from the students illustrate their opinion on group size:

2-5, you mustn’t be too many (4).

According to the reports a too small group may also lead to negative consequences. It is important for the students that the group size is just large enough to carry out the task and that each group member have work to do on the assigned task. If some members feel unwanted or superfluous, the group size is too large.

[The group is too large] when one or several persons succeed in sitting quiet and doing nothing … (6).

**Mix of members.** The students seem to prefer a group composition that is not too heterogeneous as to interest and competence. It seems to be important to relate to at least one other member in the group. The students have no decided opinion as to gender aspects. It depends on the content and the class in general whether boys and girls can profit on a mixed composition. It is obvious that they believe that groups composed by students that do not like each other are a disadvantage to a good group work.

Yes, or it depends on what you are supposed to do, if it is a rather large so… If you kind of need to work … really hard then it works best when the teacher decide so everybody works together (4).

Implicit in the quotation above is the issue about who shall decide the group composition; the students or the teacher. The replies reveal an uncertainty among the participants. The determining factors seem to be the aim of the group work, to learn a subject or to learn to work together.

The subject and whom you work with (2).

Thus, group composition seems to be an important prerequisite for students’ positive or negative experiences of group work.

**Time.** Organisation of group work also includes frames and context for the group work. The students emphasise the time aspect which can be seen as a self-evident condition. It is important that the time allotted to the group is adapted to the task. The students seem to have an extensive experience of too short time allocated to working in groups, nevertheless they usually try to complete the work at school. They may use breaks to finish the task, which the quotation below illuminates.
On natural science … then we often have to use breaks between lessons to write
the report because you don’t have time for that during the lesson (6).

_location_. An important condition for successful group work is access to a
calm place to work in. A disorderly and noisy milieu will spoil the possibilities
to produce good work. Access to group rooms might be one solution to diminish
a stressful surrounding and promote high quality group work.

Yes it is silence around you (3).

The organisation of group work, together with the next factor on the list, mode
of working, were the aspects discussed most extended and energetic during all
the focus groups.

_factor II: Mode of working_

There are primarily two aspects that are mentioned as important to the mode of
working, i.e. whether the members *focus* on the core task or not and also
whether all members or just a few of them *participate* in the common work.

Focus on the task. A successful group work is signified by the fact that all
the members are focused on the task, which means that they spend time to
understand the task and to plan the work.

Talked through what we were supposed to do and so on, kind of planned (1).

In failing groups the group members are focused on other things, for instance
social or private issues. The reason for that is not explicitly pronounced but a
qualified guess is that the task may be hard to capture.

When everybody contributes, you know and participate in the discussions and so
on. When nobody is just sitting there and … pretend to participate … [Laugh]
(3).

As long as there is an effective use of time a variation of working modes is
possible. Sometimes it is in the groups’ best interest to divide the task among the
group members and after a while reassemble and put the pieces together. On
other occasions the group may benefit from working closely assembled during
the whole group work. The excerpt below illustrates a negative experience
where a lack of co-operation is obvious.

Everybody is doing his or her own little piece of work. You might even write
about the same subject all of us … if you don’t talk with each other (2).

Participation. A central point seems to be that all members take part in the
work, which means that all group participants are at the group’s disposal with
their special competences. Of course the contribution among members may vary
but it is essential to spend time and effort on the work and not try to take advantage of the other participants’ contributions without doing something in return. When some members, for some reason, do not participate in the common work such an attitude is to the detriment of the common work. An example on such an avoiding action is given below.

When you play around (...) and don’t talk about it but about a lot of other stuff and forget the task, kind of (4).

Active participation and willingness to contribute to the common goal is closely related to a realistic ambition. The students reveal that it is important to keep the ambition on a reasonable level otherwise it might be destructive for the groups’ work, process and outcome.

Beside group composition, participation was the most discussed subject in the focus groups concerning successful group work in school. An interpretation is that these two aspects are highly important for successful group work.

**Factor III: Tasks**

The third aspect brings to the fore a discussion about the joint task the students are supposed to handle. Two aspects seem to be in focus, i.e. the task must be *intelligible* as well as *stimulating*.

*Intelligible task.* The students emphasise the importance of intelligible tasks. A comprehensive task promote co-operation and the willingness to work, while an incomprehensible task on the contrary, might inhibit any kind of group work. If the students do not understand the assignment, they are unable to start working, which in turn might divert them into other activities. In the excerpt below one of the groups explains that one characteristic of bad group work is when you do not grasp the task.

And then when you don’t understand what you are supposed to do? (3).

*Stimulating task.* In addition to being intelligible, the task also needs to be stimulating and arise the students’ curiosity. An interesting task seems to promote high quality group work and have a positive influence on the students’ willingness to work. A boring task, on the other hand, can have a devastating effect on the students’ motivation and lead to a low participation and bad quality. The students give some examples of group work when one or two persons have done all the work.

Or they don’t even try at all and you end up doing everything yourself instead (1).

The excerpt discloses that only the most dutiful students try to carry out the task, maybe in order to secure their mark.
Factor IV: Reports

Even though the report of the group work usually is accounted for at the very end of the work it is still an important prerequisite for meaningful work and for the students’ conception of well-being. Some students’ conceptions of low quality group work have to do with bad experiences of report procedures and performance. Three aspects are crucial; content, performance and responsibility.

Content of a report. The core content of the report is self-evidently of importance to the result reported by the group. However, it is also important to present the content in an understandable way. The audience does not value a messy presentation.

Performance style. A good performance is, according to the informants, lively and attracts the audience’s attention. The presentation must not be hard to follow, otherwise you will lose the audience. An understandable language, not too much boring facts and just the right pace are elements of great importance for a high quality presentation. The example below constitutes an illustration of how the informants highlight both content and performance.

A different report
When it is interesting
Mysterious
Yes … when, that is to say, when you notice that …
[Laugh]
…when you notice that people are listening and so on (3).

Responsibility for the report. In accordance with the previous aspects responsibility, once again, seems to be an essential element. All group members must be involved and take responsibility for the groups’ report presentation. Bad experiences reported by the students concerning group reports are often connected to reports where just a few of the members have contributed in the groups’ common presentation.

And everybody must participate in the oral report is my opinion or else it might mm … If only a few report the group one might not have understood, he or she just stand there … And if everybody report the audience understand that everybody have participated (1).

Factor V: Assessments

Even though an assessment can be seen as a rather important factor concerning group work it is obvious that the students mostly do not know if they are assessed at all, neither during the process nor in the presentation. One aspect concerning assessments stands out as the most crucial, and that is transparency.
What the students ask for is understandable and transparent rules for assessment and marking.

**Transparency.** There seems to be an apprehension among the students that they will be assessed but they do not know how or when the assessment takes place. Are they being assessed as individuals or as a group? Is the teacher’s assessment based on the work process or based on the outcome, the report? The uncertainty is expressed in the quotation below.

> It is different from teacher to teacher
> Some teachers give you, type of ….
> What they think
> … both on your own achievement and the groups’ (2).

Furthermore the students are unaware of if there are any criteria for the assessments. The students lack written evaluations in connection with the report, as well as feedback for possible improvement of future group work.

> It is also good if you get sort of … you know if somebody tell you something after, you know some comments and so on. … Then you can still get something written and write something like this and this you might improve … and think of that for next time … (6).

The quotation points to a wish to be assessed in a way that can help the students to improve, i.e. they are interested in being informed of both their advantages and disadvantages. The informants also ask for an oral supportive evaluation close to the oral presentation, while written accounts also may contain more negative aspects and be given individually.

**Factor VI: Role of the teacher**

The teacher is the very last, but not the least, of the six important preconditions emerging from the focus groups. The result shows that the teacher has an important role concerning the realisation and outcome of the group work. Two crucial “teacher aspects” have been identified, namely the teacher as an *arranger* and *supporter*.

*The teacher as an arranger.* One of the teacher’s roles is to organise the frames for group work. This includes arranging a context and giving means that promote co-operation. Failing to structure appropriate arrangement may have a negative impact on the groups’ possibilities to co-operate or to work in groups at all.

> It must be an attentive teacher who sees that everybody is working (2).

The students also emphasise that the teachers need to have the situation under control, which includes both the context and the students involved in the group activities. Nearly connected to control is support, which comes next.
The teacher as a supporter. The teacher’s whole-hearted support is the be-all and end-all for the process and for the outcome of the group work. The teacher needs to be attentive and must not abandon the students during their work. If the students perceive a lack of interest from the teacher or if the teacher is leaving the students alone to sort out problems, they lose interest in the task. The students hold the opinion that a present and supportive teacher, prepared to answer questions, promotes high quality group work.

They check on how you are doing … if you need any help then … you get, then they assist, and if you sit and chat (5).

It is obvious that the students ask for an engaged and attentive teacher, with a keen ear for their needs.

Conclusions
Each of the factors presented in Table 3 have in this presentation been shown to contain important prerequisites for meaningful group work and co-operation, or in other words, aspects promoting as well as prohibiting group work. The two factors at the top of the table (group work organisation and mode of working) seem to take up a lot of the informants’ discussion in the focus groups. The most essential aspects to the students proved to be whom you are working together with, i.e. group composition and the participants’ responsibility and contribution.

One remarkable conclusion is that the students included in this study, have an impressive knowledge and understanding of the realm of group work. According to the students the ideal group work is composed of a mix of 3-4 members willing to participate and focus on the work. Additionally the arrangement and support from the teacher, assuring the group enough time and resources to accomplish an intelligible and distinct task and the report, will probably lead to a high quality group work. Finally, a transparent proceeding for assessment must be present.

Discussion
This study reveals that the students consider six preconditions as important for a successful group work (organisation, mode of working, task, report, assessment and role of the teacher). Certainly these six factors could be seen as a lay theory and the students’ excuses for not taking responsibility for the common work. However, all six factors recur somehow or other in scientific small group research. We will just briefly mention some of them below.

Students’ opinions related to current research
Factor 1. The impact of organisation, for instance, has been demonstrated by a number of studies. The optimal size of groups with adults has been estimated
to five persons or less (Shaw, 1976; Brown, 2001). Larger groups have problems with processing information. Another aspect of importance to the students was the group composition. The group must not be too heterogeneous. Nothing in the present study gives evidence for gender being an important aspect of heterogeneity. Other aspects, such as interest and ability seem to be more important. Wheelan (1996) argues that differences in status might be even more important than gender differences. Forslund Frykedal (2008) found similar patterns in a study concerning group work in compulsory school classes. The condition concerning time and location suited for group work coincide with Schlecty’s (1976) theory of a clear and understandable organisation for learning. Thus, the students’ opinions are supported by research.

Factor II. Research has clearly shown that mode of working in groups are dependent on two prerequisites; the students’ skill for the task (Peterson and Miller, 2004) and participation (Cohen, 1994). Working in groups is not an inborn ability, but has to be trained (Ashman & Gilles, 1997; Cohen, 1994). Students will experience positive involvement in group work when they have appropriate skills for the task (Peterson & Miller, 2004). Cohen (1994) also stresses the importance of a teacher convincing the students that no single group member has all the abilities to complete the given task, in order to create expectations about co-operation. Once again the students’ opinions are supported by research.

Factor III. The importance of the task for the groups’ performance and processes has been verified in a number of studies (Steiner, 1966, 1972, 1976) and in studies concerning education (Forslund Frykedal, 2008; Hammar Chiriac, 2008). “Real group work” is characterised by a common effort, assembled utilization of the groups’ competence and some sort of joint problem solving (Hammar Chiriac, 2008), a kind of group work rarely occurring in Swedish schools (Granström, 2006). Group work often consists of one or a few of the group’s members’ contributions, while the others are doing something else. The importance of teachers’ engaging and appealing introduction of the group task is highlighted (Gilles, Ashman, & Terwel, 2008). Thus, current research confirms the students’ demand for adequate and understandable tasks.

Factor IV. According to earlier research about report of group work there is an ongoing discussion of the pros and cons of group versus individual accountability (Johnson & Johnson, 2008). Some studies suggest that both forms of accountability may increase responsibility as well as interdependence among group members. The understanding of the task is important and may reflect the presentations of the groups’ performance during the report (Granström, 2006; Huber & Huber, 2008). Educational research, as well as the students in this study, point to the importance of clear frames for reporting group work.

Factor V. Concerning assessment earlier studies have shown that feedback and rewards are highly relevant factors when organising group work (Hammar Chiriac, 2008b; Steiner, 1972, 1976; Underwood. 2003). It is important that the
reward system match the groups’ assignment. Group reward can direct the group members to work together “as a group”, while individual reward can encourage “working in a group”. The fundamental issue may not be whether the actual reward system promotes individual performances or a group product, but what the students believe they are rewarded for. If a teacher wants a group to collaborate this must be one of the assessment criteria which will be rewarded. Equally important is to enlighten the students about the criteria for assessments (Jaques, 2000). The students’ need for transparency and clear criteria for assessment coincide with contemporary research findings.

Factor VI. Recent research gives evidence of the informants’ need for a teacher role that is attentive, supportive, interactive and clarifying (Gillies, 2008; Huber, Sorrentino, Davidson, Eppler & Roth, 1992; Johnson & Johnson, 2008; Peterson & Miller, 2004). The role of the teacher in group work must not be underrated or neglected (Huber, Sorrentino, Davidson, Eppler & Roth, 1992). Group work is most effective when the teacher presents clearly structured situations including guidelines. Furthermore, teachers should focus their attention on designing reasonably challenging tasks that help students to reach their goals, and ensuring that students have the necessary skills for succeeding in their tasks (Johnson & Johnson, 2008; Peterson and Miller, 2004). Thus, the students have, in accordance to the research findings, encircled the teacher as a key factor for successful group work.

Participation, Quality and relevance

In a previous study concerning boredom at meetings (Granström, 2001) three factors signified a sense of boredom. These were lack of participation (P), lack of quality (Q) and lack of relevance (R), i.e. lack of PQR. Quite the reverse signified a meaningful session, which means the presence of PQR. In fruitful cooperation all members are participating (P), the form of working is well organised and the task is understandable, i.e. the work conditions have a good quality (Q). Finally, the issues at hand need to be relevant and meaningful to the members (R). The model has been transferred to and verified in a study concerning the teacher training programme (Hammar Chiriac, 2007). The teacher training students confirmed that participation, quality and relevance were crucial incentives for learning. Group work and practical teacher-training situations were, according to the informants, important learning opportunities. Presence of PQR may promote high qualitative learning while the opposite, low level of, or the absence of PQR may lead to low quality learning.

Applying these three, evidence-based criteria, to the factors and their different aspects, which emerged in the present study, confirms that the students ask for PQR. In Table 4 the students’ definitions of a good group work is related to the PQR theory. As can be seen a most important precondition of successful group work is participation (P). The arrangements (group size, group composition,
time and location allotted) must be such that they support all members’ possibility to take part in the work. However, the conditions must also guarantee a high quality (Q), which means understandable tasks and interesting reports as well as a chance to consult the teacher. Finally, the work conditions, the task and the assessment need to be relevant and adjusted to the students’ abilities and interests (R).

Table 4  
*Factors and aspects promoting co-operation related to PQR*

<table>
<thead>
<tr>
<th>Factor/aspect</th>
<th>Promoting co-operation</th>
<th>Aspects of PQR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ORGANISATION</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Group size</td>
<td>2-5 students</td>
<td>P</td>
</tr>
<tr>
<td>• Group composition</td>
<td>Not too heterogeneous</td>
<td>P</td>
</tr>
<tr>
<td>• Time</td>
<td>Suited to task</td>
<td>P</td>
</tr>
<tr>
<td>• Location</td>
<td>Calm</td>
<td>P</td>
</tr>
<tr>
<td><strong>MODE OF WORKING</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Focus</td>
<td>On the task</td>
<td>R</td>
</tr>
<tr>
<td>• Participation</td>
<td>All members take part</td>
<td>P</td>
</tr>
<tr>
<td><strong>TASK</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Intelligible</td>
<td>Comprehensible</td>
<td>Q</td>
</tr>
<tr>
<td>• Stimulating</td>
<td>Interesting</td>
<td>R</td>
</tr>
<tr>
<td><strong>REPORT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Content</td>
<td>Understandable</td>
<td>Q</td>
</tr>
<tr>
<td>• Performance</td>
<td>Lively</td>
<td>Q</td>
</tr>
<tr>
<td>• Responsibility</td>
<td>All group members</td>
<td>P</td>
</tr>
<tr>
<td><strong>ASSESSMENT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Transparency</td>
<td>Known criteria</td>
<td>R</td>
</tr>
<tr>
<td><strong>ROLE OF THE TEACHER</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Arranger</td>
<td>Promote co-operation</td>
<td>P</td>
</tr>
<tr>
<td>• Supporter</td>
<td>Being attendant</td>
<td>Q</td>
</tr>
</tbody>
</table>

The consequences of this study is important and a challenge to teachers. It is obvious that all aspects with relevance to good group work have to be dealt with by the teacher in a proper way. The PQR model could fairly well be a sort of checklist, when planning and realizing group work in schools and other contexts.
Concerning organisation, the teacher has to decide group size, compose groups, allot time and guarantee a calm location. The teacher needs to train the students in the mode of group work, which means regular training in group work; how to plan a group work and how to communicate in a group situation. The teacher has to give understandable tasks and train the students to report a group task. The assessment needs to be formulated beforehand concerning form and content. The role of the teacher has to be as an arranger and in course of the group work the teacher needs to be a supporter. The teacher may not leave the students alone to form the group work. This is not a pedagogical or democratic approach, it is just to abandon the students and to give up a responsible teacher role. When considering these six factors, the PQR model could be helpful for teachers to ascertain group work of high quality (Q) with high relevance (R) to the students, and to offer them opportunity to participate (P).

If a group work fails, the teacher has to be blamed – not the students.

References


Author Note

Eva Hammar Chiriac, PhD, senior lecturer, Department of Behavioural Sciences and Learning, Linköping University,
Kjell Granström, professor in educational psychology, reader in communication and authorized psychologist. Linköping University
Correspondence concerning this article should be addressed to Eva Hammar Chiriac, Department of Behavioural Sciences and Learning, Linköping University, SE-581 83 Linköping, E-mail: eva.hammar.chiriac@liu.se
LÄRA I SAMARBETSSITUATIONER

Karin Forslund Frykedal

Abstract
Group-work is a method that facilitates students’ learning of team skills and academic knowledge. However, research shows that the method is unusual in the Swedish schools. The present article accounts for a theoretical model, which describes the pattern of interaction, and the students’ experiences that emerge as they try to handle a group-work situation. The model elucidates four different conditions, which may have influence on the students’ work. The four conditions are (a) a difference of ambition, (b) group assignment, (c) group composition and (d) creation of trust. These conditions are discussed and problemized, with respect to the role of teachers and their influence when using group-work and how they can realize a well-functioning group-work.

Keywords: Group, group-work, grounded theory, theoretical model

Inledning
avhandling som kom 2008 med titeln Elevers tillvägagångssätt vid grupparbete: Om ambitionsnivå och interaktionsmönster i samarbets-situationer.

Syftet med texten har varit att visa några förutsättningar som påverkar eleverna när det går in i ett grupparbete och även visa några aspekter på hur de försöker påverka grupparbetet så att det ska passa in i deras ambition. Utifrån syftet växer två frågeställningar fram som jag ska försöka besvara i diskussionen: Med utgångspunkt i den empiriskt skapade ETG-modellen, vad är möjligt för lärare att påverka när de använder grupparbete som arbetsform i sin undervisning? Hur kan de påverka och skapa förutsättningar för att utveckla ett väl genomfört grupparbete?


Metod


I grounded theory sker insamlandet av data genom teoretiskt urval (theoretical sampling) vilket innebär att insamling samt analys av data försiggår sida vid sida och är svåra att särskilja från varandra. Det teoretiska urvalet sker genom ett medvetet val av data från olika källor (Glaser & Strauss, 1967).

I inledningen bestod datainsamlingen av observationer med låg grad av struktur med mig som deltagande observatör (Hammersley & Atkinson, 1995; Einarsson & Chiriac, 2002). I början var observationerna öppna där jag försökte, med Glaser och Strauss (1967) term ”main concern”, se vad som hände i klassrummet och vad som upptog elevernas intresse. Ganska snart kom ett behov av fokus, som blev att observera eleverna i deras grupparbeten.

Observationerna ägde rum under två till tre dagar i veckan under en termin, nästkommande termin genomfördes 13 semistrukturerade intervjuer med eleverna (Kvale, 1997). Efter att den empiriska insamlingen avslutats fortskred
analysprocessen och i senare delen använde jag mig även av litteratur inom områden som beskriver liknande fenomen som framkom i studien. Ur analysen av empirisk data konstruerades kategorier och begrepp samt relationer mellan dessa utvecklades. De ”fylldes på” och blev tätare med hjälp av komparativ metod (Glaser & Strauss, 1967).

Läsningen av utvalda specifika teorier gav vidare information om och tydliggjorde ytterligare dimensioner i kategorierna och begreppen samt dess relationer. När kategorierna var konsistenta och som grounded theory förespråkar ”mätta” avspeglade de och återgav respondenternas erfarenheter av den fokuserade företeelserna och gav en teoretisk begreppslig förståelse, som visserligen inte helt kan återspeglas fenomenet som fanns i klassrummet, men som kan ge en teoretisk bild (Charmaz, 2006). I studien skapades en teoretisk modell, ETG-modellen, som hjälper till att förklara fenomen och skeenden i grupparbetet.

**Resultat**


*Ömsesidigt beroende och ambitionsdifferens skapar ambitionsbetroende*


**Påverka gruppsammansättning**

*Elevernas olikheter* vad gäller ambition och föreställningar om beroende, visas när de möts och ska presta en uppgift i ett grupparbete. Det blir större eller upplevs av individerna som större om de placeras i grupper av någon annan. Grupperna komponeras ofta av lärarnas vars intentioner är att de ska vara sammansatta av elever med olika *ambitionsnivåer* och förmågor. Eleverna försöker dock ständigt påverka gruppsammansättningarna genom att utöva inflytande på vilka de ska arbeta tillsammans med i syfte att utjämna ambitionsdifferensen i de blandade grupperna. Detta sker med skiftande framgång. Man söker efter möjligheten att arbeta med utvalda kamrater av antingen sociala eller ämneskunskapsmässiga skäl. Kamraterna ska antingen ha liknande ambition eller befina sig relationsmässigt nära. Därför anstränger sig eleverna för att både i och utanför arbetet försöka sätta sig bredvid de utvalda istället för att sitta kvar i de av lärarna skapade grupperna.

För en del elever är det viktigt vem de arbetar med och de strävar efter att skapa gruppsammansättningar där kamrater ingår. För andra elever är inte personen lika viktig utan miljön i gruppen är viktigare. Därför är dessa elevers strävan mer att skapa en positiv miljö än att söka sig till speciella elever.

**Påverka uppgiften**

*Uppgiften* som ges till ett grupparbete försöker eleverna konstruera utifrån sin ambition vilket skapar olika ansvarstaganden. År den ämneskunskapsmässiga ambitionen hög blir ansvaret över uppgiften genomförande också hög vilket kan ske genom styrning och/eller ett stödjande förhållningssätt. År däremot den ämneskunskapsmässiga ambitionen lägre eller om den relationsmässiga är hög blir ansvarstagandet i uppgiften lägre vilket kan ske genom flykt från uppgiften eller genom att lifta och åka snålskjuts på att andra tar mer ansvar.

I uppgiften har fyra olika aspekter synliggjorts som är mer eller mindre påverkbara för eleverna. Vilken *karaktär* en uppgift har visar hur delbar eller


Tillitsskapande
Eleverna har oftast inte möjlighet att påverka gruppssammansättning och uppgiften fullt ut i den beroendesituation som skapas i grupparbetet utan försöker istället hantera situationen med utgångspunkt i den egna ambitionen. Detta är i de flesta situationer inte alltid reflekterande handlingar som eleverna verballiserar, på liknande sätt som de pratar om att vilja arbeta med bästisen eller att dela uppgiften i individuella delar. Det eleverna inte kan påverka försöker de hantera genom ett tillitsskapande i vilket de söker tillit till egen och andras prestation samt i de sociala relationerna i gruppen. Är den ämneskunskapsmässiga ambitionen stor blir även förväntningarna på både egna och andras arbeten stor och är ambitionen låg blir förväntningarna på prestationen lägre. Är den relationella ambitionen stor är det viktigt att skapa
och underhålla relationer vilket gör att arbetet kan komma i andra hand, är den inte det kommer för det mesta arbetet i första hand. Förutom ambitionen är även föreställningar om beroendet viktigt och om beroendet upplevs som en börda eller som en möjlighet. Detta utvecklar olika typer av tillitsskapanden hos eleverna.

**Hantera genom olika tillvägagångssätt**

Eleverna sätt att skapa tillit i grupparbetet sker genom olika *tillvägagångssätt*. I tillvägagångssätten gestaltas ambitionen, föreställningar om beroende, kamratrelationer och ansvarstagande för uppgiften. Fem olika tillvägagångssätt har synliggjorts, *dirigerande*, *sporrande*, *assisterande*, *gömmande* och *liftande*.


Tillvägagångssätten skapas i situationen och blir till olika sätt att hantera ambitionsberoendet hos eleverna. De är ständigt föränderliga och varje enskild elev använder flera under arbetets gång även om någon eller några är mer återkommande. I ett grupparbete kan alla förekomma men behöver inte göra det, samtidigt kan ett eller flera tillvägagångssätt visas av flera elever.

Följande figur vill visa sambanden mellan de olika delarna i den teoretiska modellen.
Individer med olika ambitionsnivå

Får

En gruppuppgift

Vilket skapar

Ambitions-
differens

Beroende mellan
eleverna

Ambitions
beroende

I det ambitionsberoende som uppstår försöker eleverna minska beroendet och ambitionsdifferensen genom att påverka

Grupp-
sammansättning

Uppgiftens
konstruktion

Kvarstående skillnader i ambitionen och beroende hanteras genom ett

TILLITSSKAPANDE

Vilket konstrueras genom olika tillvägagångssätt

Dirigerande Assisterande Gömmande Sporrande Liftande

Figur 5. Den teoretiska modellen Elevers tillitsskapande i grupparbete, ETG-
modellen.

 Diskussion

Syftet med texten har varit att visa några förutsättningar som påverkar eleverna när det går in i ett grupparbete och även visa några aspekter på hur de försöker påverka grupparbetet så att det ska passa in i deras ambition. Syftet besvaras
med hjälp av en empirisk grundad teoretisk modell. Utifrån syftet växer två frågeställningar fram som jag ska försöka besvara i diskussionen: Med utgångspunkt i ETG-modellen, vad är möjligt för lärare att påverka när de använder grupparbete som arbetsform i sin undervisning? Hur kan de påverka och skapa förutsättningar för att utveckla ett väl genomfört grupparbete?

I ETG-modellen visas att det finns minst fyra delar som lärare har möjlighet att påverka för att kunna komponera väl fungerande grupparbeten. Dessa är (a) uppgiften (b) gruppsammansättning (c) ambitionsdifferensen (d) skapa tillit.

**Utnyttja grupuppgiftens möjligheter**


**Medvetenhet om gruppsammansättningens betydelse**

Individer skiljer sig åt vad gäller personligheten samt sociala och ämneskunskapsmässiga förmågor som de tar med sig in i ett grupparbete. Ju större olikheter bland individerna, desto större förväntas den kreativa korsbefruktningen av idéer bli. I mer homogena grupper saknar individerna möjlighet att ge divergerande infallsvinklar i en diskussion, de tenderar att


Vikten man lägger kring vilka man helst vill arbeta med skiljer sig åt bland eleverna och är också kontextberoende. En del elever ser inte en specifik individ som viktig i samarbetet utan istället är gruppklimatet viktigt. Varje elev har därför sin uppfattning om vilka de anser vara de bästa gruppmedlemmarna. Det är därför viktigt för lärare att inför varje gruppsammansättning fundera över hur grupperna ska formas och under arbetets gång också ha klart för sig hur mycket man som lärare anser att grupperna ska få förändras när eleverna försöker förändra gruppsammansättningen. Vill man som lärare att grupperna ska vara heterogena eller homogena och når ska valet av gruppsammansättningen ske av läraren och när ska det eleverna få välja vilka de ska arbeta tillsammans med? Det finns inget självlklart svar på detta men viktigt är en medvetenhet hos läraren om att han eller hon kan påverka lärandet genom att styra över hur grupper ska formas.

**Att skapa förutsättningar för att olika ambition får ta plats**

Lärarna i denna studie har intentionen att grupperna ska bestå av elever med olika ambition och utvecklad förmåga. Den blandningen skapar en repertoar av olika perspektiv vilka eleverna skulle kunna dra nytta av och både få och ge stöd åt varandra i uppgiften (Lotan, 1997b). Istället blir dock flertalet elever frustrerade över arbetssituationen när de arbetar i grupp och därför försöker de ständigt påverka gruppsammansättningarna för att minska ambitionsdifferensen.
Den här studien har visat att det finns samband mellan elevernas föreställningar om att arbeta i grupp och deras ambition. Elever med högre ambition vad gäller ämneskunskapen anser att grupparbete blir en börda därför att de får driva både sitt eget och gruppens projekt. Elever med lägre ämneskunskapsambition kan däremot tycka att det är bekvämt men det finns också en grupp bland dessa elever som anser att grupparbete skapar osäkerhet, därför att de inte kan dölja den egna förmågan.


Proaktivt arbete med tillit till varandra och varandras prestation


Sammanfattningsvis
ETG-modellen har visat att det finns minst fyra aspekter som är möjligt att påverka när man arrangerar undervisningen för ett eleinteraktivt lärande. Jag har här diskuterat hur man kan skapa förutsättningar för att påverka dessa både med stöd av vad min forskning men även vad tidigare forskning visat. Diskussionen väcker också ett antal frågor. Frågor som jag är nyfiken att titta närmare på i fortsatta studier. Vad skulle implementering av metakognitiva samtal kring grupprocesser och interaktivt lärande få för konsekvenser vad gäller att kunna skapa väl fungerande grupparbeten? Vad skulle transparabla, väl

genomtänka bedömningar, kamratvärderingar och självvärderingar av kunskap och kompetens som utvecklas hos individerna i grupperbetet innebära för utarbetandet av väl fungerande grupperbeten?

Referenser


**Author note**

Karin Forslund Frykedal, PhD, senior lecturer, Department of Behavioural Sciences and Learning, Linköping University.

The text builds on choosen parts of the thesis Elevers tillvägagångssätt vid grupparbete: Om ambitionsnivå och interaktionsmönster i samarbetssituationer.

Correspondence concerning this article should be addressed to Karin Forslund Frykedal, Department of Behavioural Sciences and Learning, Linköping University, SE-581 83 Linköping, Sweden, E-mail: karin.forslund.frykedal@liu.se
INTER-GROUP PLAY AND SYMBOLS OF A MASS EVENT
AT THE WORLD CUP IN FOOTBALL 2006

Stephan Hau & Johan Näslund

Abstract
In this article we will describe group behaviour out from three different theoretical perspectives: Social psychology concepts and symbol theory are used to explain intergroup behaviour and behaviours within a group. In addition we apply a psychological theory in order to describe processes within the individual group member, which precede and are related to specific group behaviours. This includes the abilities of mental space and play. These capacities are corner stones in identity development processes. Symbolic play, an intermediate space, neither tallying the outside world nor the inside world is another condition for peaceful interactions with others. We try to apply these theories, developed with respect to the individual, in order to expand their ambit to group processes, creating a multi dimensional picture of group processes.

Keywords: Mass events, symbol, symbolic activity, play, transitional space

Introduction
In later years violence in connection with sports activities get increased publicity, especially when it comes to football. In these contexts people often makes distinction between ordinary supporters (the fans) and the hooligans who are making trouble. A German Police Officer (Hau, 2008) makes distinction between three groups of supporters; (a) the ordinary supporter who never makes troubles, (b) ordinary supporters that can be triggered to participate in violence, and (c) hooligans that always makes troubles. Most of the interest in media and society are about the hooligans and their significant contribution for violence in connection with football matches. A sociological viewpoint of football supporters is that society have changed and so the supporters. From traditionally to market oriented spectators and from those who are identifying with the football club to spectators that not are interested in the football club (Giulianotti, 2002).

Symbols
Symbols are objects that represent abstract ideas or concepts. Symbols can appear in form of objects, character, figures or actions. The connection between the symbol and the concept that it represents is not naturally given but a result of a negotiation process that leads to a convention. From that perspective all symbols have a communicative function. They refer to something that is absent in an allegorical way, i.e. they represent but do not directly codify the related
concepts. In contrast to symbols signs can be seen as having a discrete unit of meaning – in a denotative or connotative way (cp. Pierce 1960, de Saussure 1922, Jakobson 1971).

Victor Turner (1967, 22) describes that ”symbols … produce action … Groups mobilize around them, worship before them, perform other symbolic activities near them, and add other symbolic objects to them, often to make composite shrines” (ref. see Womack 1995). Womack (2005, 5) sees the meaning of symbols as arbitrary. Even though it is not a completely arbitrary process, the actual meaning of a symbol is the result of a mutual negotiation process of members of a society/group. An interactive negotiation process does not automatically succeed. Were no agreement about the mutual meaning of symbols can be achieved the risk for escalation and violent occurrences is rising.

Symbols and football
Even though, a football match is not a real war it can be interpreted as a real fight between football teams. The symbolic fight, in analogy with the real fight, is between the different football fans. During the match the real and the symbolic fight is interconnected, but before the football match a symbolic fight probably will takes place. This way of seeing the situation reminds us about the way Granström, Guvå, Hylander, & Rosander (2006) interpreted the demonstrations in connection with the EU-assemble in Gothenburg in 2000. Based on intergroup theory they talked about peace or war producing actions. Thoughts, feelings, and actions of the group is heavily influenced by the actions of the group, how they interpret their own and other groups actions and intentions.

In order to study ways of symbolic processes in group play situations we chose for this investigation to observe supporter fans of two football teams which played a group match during the football world championship in Germany, summer 2006. In contrast to political demonstration the city of Dortmund could be liken to a festival and play. A play that brought our thoughts to the child plays were adults could interfere and finish it of (food is ready) or one playgroup has won the game (today you won but tomorrow we will win). They must to come to an agreement about who won this time. This agreement facilitates by adults that interfere and keeping track of time and place. Furthermore, several conditions have to be given – on a group level as well as on an individual level, which are described and discussed in the next chapters.

Self and play
In order to develop a feeling of identity in a group and to regulate fantasies and affects, some preconditions have to be given in the individual that are conceptualised by theories about transitional objects and transitional space (Winnicott 1971). They serve as models that help to create a link between
individual experiences and group-processes in mass events. Winnicott describes the onset of symbolisation activity and fantasy life. Small children use to cope with stressful situations by creating an inner representation of an object that has the ability to calm and comfort (e.g. a Teddy-bear, a blanket, a soft object). These “transitional objects” are regularly equipped with the same features of important caregivers and children can find some symbolic replacement of the caregiver’s qualities. Thus they are able to stand and to manage the stressful situations on their own. Transitional objects are neither 100% part of the outside reality nor 100% part of the inside psychological world but exist on the interface between the two and can be used in order to find compromise solutions between the outside reality (e.g. were one is being left alone) and inner psychological wishes (e.g. togetherness). As a result the child creates an intermediate space in which both sides exist and can be manipulated – the inside world as well as outer reality – which helps to cope with both worlds in a more and more competent way. Later in life the functions of the transitional objects are replaced by a transitional space, i.e. by the child’s capacity to play. However the capacity requires some defined space where the play action takes place as well as a limitation by time. Defined starting points and end points are needed, which help to ensure the play activity.

While playing, this transitional space and the events that take place, the stories that develop, can be actively manipulated by the individual player, without having to take regulations and constraints of reality into consideration. The transitional space is not identical with the inner world of thoughts, fantasies, wishes, anxieties, etc. Furthermore, it functions like a stage on which different fantasies, ideas, problem solving, wish fulfilment, etc. can be enacted. These enactments have conscious and unconscious components.

The Dutch philosopher and anthropologist Johan Huizinga writes about the irrational play as a well-defined action different from ‘ordinary life’, “as a contest for something or a representation of something” (Huizinga, 1950, 13). Play is performed in a playground that has its well-defined boundaries and time, and could be likened with a feast. A temporary suspension of the normal life where the group fits the play as “the hat to the head”. The players often disguise themselves and follow the rules of the play so they not violate the play.

Indeed, as soon as the rules are transgressed the whole play-world collapses. The game is over. The umpire’s whistle breaks the spell and sets “real” life going again.
(Huizinga, 1950, 11)

When dealing with play and interaction, transient intensification of emotions occurs. The level of affectivity rises in a safe situation, as the possibility of negative and unwanted consequences is lowered. At the same time transient partial exclusion of critical reality testing can be observed. However, when
looking at a play situation a part of reality testing of the participants must still be intact. Otherwise they would not be able to distinguish between play and reality and probably would not find their way back, once the play is over or the playground is left. Children who are emotionally deeply involved in play activity can develop difficulties to find their way back to reality. We assume the same processes within groups and during group interaction. Once deeply involved in processes involving high affectivity, it might be difficult for group members to calm down again once the game is over. Especially rivalling groups can easily develop states of high affectivity.

Meanings and symbols that are ‘valid’ and that can be understood by both playing partners have to be negotiated and defined. The players interact permanently by exchanging and communicating about symbols, their meaning and understanding on the playground of their mutual transitional space.

In groups constituted by adult persons these negotiation processes can be assumed at work as well. However they can be facilitated by pre-formed symbols, especially by oversimplified patterns, rituals, dress codes, etc. These symbols as well as symbolic activities (expressed via four areas of human anatomy: face, voice, hands, and legs, cp. McPhail 2006) allow to easily getting access to the mutual transitional space. They are subject to change (new rituals, songs, symbols) but also include a certain part of inconsistencies, which are important for identity processes.

**Group and intergroup**

When children or adults meet in a group play, they know who they and others are, as group members. The membership with the own group and discrimination of other group members (us and them) demand rules, norms, and borders. The group-playing children are heavily influenced by other children outside the group. The leader (motherhood) has been replaced by other persons and groups, in a fantasy about “they”, which affect the group (we) play. Gibbard, Hartman & Mann (1974, p 271) indicate that the activity of “the group is the potential space between member and external objects”. Winnicott (1971, p 41) talks about same line of thoughts when he writes about play that leads to health and development and group relationship. Nitsun (1996) writes about the anti-group that struggle against development of the group. He also has a word of caution. The tension between reality and fantasy may get out of control and destroy the play.

Historically man has argued that under certain conditions, for instance during riots, they has lost their individual responsibility and was lost in the actions of the mob (group) (LeBon 1995/1896). Allport (1924) is writing in favour of a more individualistic explanation. Personal identity refers to “me versus not me” and social identity refers to “us versus them” (Onorato & Turner, 2004). From a sociological viewpoint, Stryker & Burke (2000) talk about “multiple identities” linked to roles and behaviour through meaning and Roccas
& Brewer (2002), from an inter-group viewpoint, talks about multiple identities linked to a multicultural society. In later years researchers from Europe has elaborated more advanced theories about social identity (for example ESIM; Reicher, 1996; Drury & Reicher, 2000; Drury & Reicher, 2005, and multiple social identities; Brewer, 2008) as a way of interpreting group actions from an intergroup perspective.

Ashbach & Schermer (1987) argue that there are different levels in transference phenomenon. From the individual, through the group, to culture and society. Beside these levels there is a fourth, the intergroup level (Wheelan, 2005). All these levels are under discussion in this article, as a gestalt approach. Agazarian & Peters (1981) writes about concept of isomorphism from a systemic angle. This concept is not under discussion in this paper but is an underlying concept in explaining how different levels have an influence on each other.

Another conceptual approach we use a concept that describes typical phases and features of group processes. Different theories have been developed in order to better understand group behaviour (Wheelan, 2005). McPhail (2006) described in his systematic investigation of a religious rally several distinguishable features of group behaviour that are of value for our approach. He refers to a “taxonomy of elementary forms of collective action” (Schweingruber & McPhail 2000) in which three different phases of temporary group gatherings are defined. As characteristic features of group processes their variations, alterations and flexibilities are named. Thus, the crowd concept was critically questioned by creating general categories describing the life course of temporary gathering, but still guaranteeing enough flexibility to give room for descriptions of rare and short-lived group events. The first phase is seen as “assembling process” in which the crowd/group appear on the place where the gathering shall take place. Various forms of assembling can be described. Being at the place where an event will take place, the second phase begins: “the gathering” which is followed by the third phase, “the dispersing process” when group or crowd members leave. Once again, the dispersing process can take place in different ways.

An additional categorical tool described by McPhail (2006) is the introduction of four distinctive groups of symbolic activities, connected to human anatomy. McPhail arrived at these categories when systematically observing and analysing the rally events. When two or more persons interact one can systematise their interaction with respect to “face, voice, hands, and legs” (441). To visualise symbols of the own national team (flags, t-shirts, scarf’s, etc.), shout or sing supporter songs, clapping hands or making specific gestures, as well as moving activity (dancing, jumping, running, etc.) refer to different modes of symbolic interaction which have a specific importance in the life of football supporter groups.
Method
In order to perform a naturalistic study of mass events (group and intergroup) at the World Cup in football in Germany (June-July 2006) a group of senior psychology researchers from Linköping University in Sweden went to Dortmund for observing a potentially “high-risk match” (according to the police in Dortmund) between Germany and Poland. This match was chosen for several reasons. In the run up of the match several groups of German and Polish hooligans had announced (e.g. on the internet) to meet for clashes and fights on match day. Even in the media (e.g. in Germany and in Great Britain) reports were published referring to the explosive situation that might develop in Dortmund. Thus, Dortmund police prepared by concentrating several hundred special police forces - trained to handle large group situations - in and around the city.

The researchers intended to study the life in Dortmund, before, during, and after the football match between Germany and Poland, from a gestalt perspective. A case study strategy that “are an ideal method for studying groups in their real context over time /.../ allow for more ecologically valid and complex treatments of groups as systems” (Berdahl, & Henry, 2005, 32).

Observations
In Dortmund, about 60,000 fans attended the match in the Stadium while 12,000 supporters watched the game on a big screen on ”Friedensplatz”, a large square in downtown Dortmund. Another 10,000 people were watching the game in Westfalenhalle, a big concert and event hall close to the stadium. In addition, uncounted numbers of fans followed the game in bars and restaurant, in outdoor locations and cafes throughout the town.

The situation was embedded in the events of the World Championship in which Germany tried to present itself as a friendly host (a time to make friends) towards the international guests. As far as Poland is concerned, the overall (cultural) background played a certain role.

The group of six researchers - divided in pairs of two – observed the events on Friedensplatz (“public view” area, 2 researchers), in Westfalenhalle (“public view” area, 2 researchers), and in central Dortmund (2 researchers). Another place that we temporarily observed was Dortmund central train station when the supporter train from Warsaw arrived. The data collected existed of; (1) recorded interviews and “songs”; (2) photography of artefacts and events; (3) field notes during the entire data collection. Data has been collected between 14.00 and 02.00 the day of the match.

Observations, phenomenological descriptions and some interpretations
The focus of the observations is the symbolic activities in connection with the football match. We distinguish between signs (symbols that have a distinct and visible purpose) and symbols (symbols that can be interpreted in different ways) (Womack, 2005). The theoretical model is from Frey & Sunwolf (2004) who describe the symbolic practice and the context of the group.

Based on the literature we created four main categories for the analytical work. We categorised them as; structure (time, place and overall culture); play (transitional space and mentalisation); symbols (body-movements, body features, voice, and hands); group activities (within and between groups). As far as the groups were concerned we categorized them in; Polish fans; German fans; police/referees; and spectators. The police and the referees of the match are both groups that operate at the borders of the system (gatekeepers). The categories of time (the assembling, the gathering and the dispersing) are the same as before, during, and after the match. The places can be categorized in different playgrounds; (a) the place for the match (arena and screens), (b) the places for play to occur, and (c) the city centre (the festival). If we place the playgrounds in a figure we can describe them as below (inspired by Kurt Lewin, 1948).

![Figure 1: Different playgrounds in Dortmund](image)

In the assembling process we realised that much play took place. The fans from both Germany and Poland prepared themselves before the match in playground (b). The police was overseeing the borders between playgrounds (b) where the fans were and playground (c) where the festival took place. In the gathering, during the match, every ones focus was on the playground (a) (football stadium and the widescreen-places), the players (who are representatives of the fans), and the referees (gatekeepers). The police (gatekeepers) was overseeing the match-arenas. In the dispersing, after the match, everything switched over to a festival in (c) the city centre. This description shows that if we wanted to observe “playing” we had to observe the “middle-ground” of the playgrounds, before the match. Neither the match itself nor the time after the match seemed suited to reveal fantasies and expectations of the supporters, most predominant in form of symbolic activities.

**The “fight” about the fountain**
In the frame below an example of intergroup-play is described in full length and detail. The playing process is subdivided into different episodes. In the detailed analysis of the play-event we refer to each of the 8 episodes.

**Episode 1**
In the inner part of the city police in riot uniforms (Helmets, cudgels, plastic handcuffs, etc.) were visible. German and Polish fans stroll in small groups through the streets. Small booths offer food and drinks. The city looks a bit like a market place. In the middle of the old town – on the place called “Alter Markt” (old market) – we see for the first time two large groups of Polish and German fans in front of each other, singing their songs. Some German fans start to jump into the water of the fountain and begin to sing and play in the water. As a reaction, the Polish fans come closer to the fountain. Suddenly, four policemen place themselves between the two groups. German and Polish fans continue to sing. A fan has painted a German and a Polish flag on his cheeks. A Polish fan is painted in German colours. The German fans start to sing “steht auf, wenn ihr Deutsche seid” (get up if you are Germans). The Polish fans answer with the same song in Polish.

**Episode 2**
In the next minute a Polish fan climbs on top of the fountain. Reactions: 1. The German fans start to hoot. 2. They start to sing the song “Ihr könnt’ nach Hause fahr’n” (You are allowed to drive home). 3. A German fan tries to climb the fountain as well.

**Episode 3**
But the Polish fan climbs the last stretch to the top and sits on top of the fountain. Both fans hug each other but the top place stays occupied by the Polish fan. The German fan decides to climb down, while the Polish fan starts to laugh and incites the Polish supporters to continue to sing.

**Episode 4**
The German supporters react on the Polish singing and start a new song: „Deutschland ist das schönste Land der Welt – Deutschland ist das stärkste Land der Welt“ (Germany is the most beautiful country of the world – Germany is the strongest country of the world). Afterwards the German supporters start to somehow ignore the situation and sing more to themselves “Oh wie ist das schön” (oh- this is so wonderful/beautiful).

**Episode 5**
Then the Germans turn back to the Polish group and sing: “Berlin, Berlin, wir fahren nach Berlin” (Berlin, Berlin, we will travel to Berlin – which is an allusion to the city where the final will take place). The Germans demonstrate indirectly that they will win at least all the matches on the way to the final and then the final as well.

**Episode 6**
The Polish fan is still sitting on top of the fountain, now with a flag in his hand. He is not very successful in motivating the Polish supporters to sing. However, the German fans react on the flag. Several Germans come and try to cover the entire fountain with a giant German flag so that the polish fan will “disappear”, won’t be seen any more (magic thinking: making invisible?!). The Polish fan, sitting on top, tries to defend himself against the covering. The Germans however fail, their flag is not big enough. The top of the fountain and the Polish fan remain visible.

**Episode 7**
Meanwhile the Polish fan wears sunglasses, sits relaxed and smirks, enjoys the situation being on top. The police have disappeared. Many Germans sit or stand around the fountain while the Polish fan starts to leave the fountain not without trying to put up his flag on top of the fountain.

**Episode 8**

At the same time the German fans sit down on the ground and listen to a man with a megaphone, shouting letters to the group: H – U – M – B – A. The group repeats each letter. All of a sudden, after the last letter, all jump up and dance while singing an old German carnival song from the 1950s: “Humba-Humba-Täterä” (which has no meaning but copies a marching band by onomatopoeia). The dancing and jumping merges into an occupation of the fountain and the Polish flag is immediately stripped off. The Polish fans somehow don’t seem to care. The whole scene lasts about 20 minutes.

**Analysis**

The observations described above will form the basis of our analysis. The course of events is divided into 8 episodes in order to clearly mark the interactive processes. Minor groups of fans strolled through the city and formed two larger groups when arriving at the place “Alter Markt”. The place with its fountain was not arbitrarily chosen. It served as circumscribed arena, as a frame, both the German and the Polish fans agreed upon. Here, over a certain period of time, interaction between the two groups was about to take place. The basic conditions for playing were installed by defining this transient space. As both groups had to communicate, they must have mentalised the understanding of the own signals and symbols by the other group including the reactions to be expected. Each group supposed that the other group had the same ideas about playing and understood the signals that were sent and reacted to them in a predictable way. The main process of interaction happened between the groups but interactions within the groups must have also taken place because each of the fan-groups had to decide that “Alter Markt” was a suitable location and frame for playing. However, as the fan groups had already assembled when we arrived at Alter Markt, we could not observe clearly these within-group processes.

**Episode 1**

Once the two supporter groups had assembled on the place “Alter Markt” they clearly formed two separated units in front of each other. They did not intermingle while forming and demonstrating their group identities. By singing their national fan songs (often with the same melody but with different German or Polish texts) they tried to demonstrate their superiority. But at the same time it became clear that there also was a mutual play, e.g. when shared musical symbols were used. Both groups prepared the playground: They had to assume
that the other group would understand and react in a certain way towards the singing and the symbols the own group presented. The agreement included the place (to gather around the fountain at a central place in the town) and to choose the highest point – the top of the statue in the middle of the fountain as the symbolic representation of victory/winning of the game. The groups also agreed that they will only play with or against each other, which meant a limitation that none of the groups would escalate the situation. Thus, We can assume that even while focused on their own rituals, songs, and symbols, the two groups interacted intensely.

Once the frame was created and maintained (by factors like the course of the football tournament, the conditions in the city of Dortmund but also by the police) the play could start. By placing several police-cars at different locations in the inner city (signs), police also demonstrate that they were present, thus representing an outer reality of society and law. At the same time, these conditions – if successfully implemented – supported a peaceful interaction and had a de-escalating effect. Within this frame, the place for playing and expressing fantasies and wishes (intermediate space) was created. The play-episodes of actions and interactions we observed had different lengths (from several minutes to just a few seconds).

That both groups sang the same melodies but with texts from different languages were a clear indicator for interaction between the two groups. In using the same melodies they demonstrated that they “knew” about the songs the other group was using. Besides all the differences, this involved a signal of shared symbols. The same was true for other symbols the fans in each group carried and displayed. Both groups were equipped with flags, they had painted the national colours on their cheeks, presented shirts, hats, scarf’s, and other fan-equipment. The two groups could identify each other as being involved in the same field, having the same goal/fantasy and using the same symbols and unification methods in order to express their emotions, fantasies, and wishes.

However, there are not only interactions between the groups, one also must assume interactions within each of the two groups. The German fans started the play by occupying the area of play: the fountain. They jumped into the water and started to sing and play. At that moment the German fans played within their own group, keeping the Polish fans away from the refreshing water on this warm day in Dortmund. Implicitly they probably assumed that the members of the Polish group, which had no access to the fountain, had to experience themselves as excluded. As a reaction, the polish fans came closer. This was a critical moment. It seemed not clear to which results this move was about to lead. Thus, policemen appeared and placed themselves between the two groups. One could assume they served as a referee by representing restrictions from the outside reality; they symbolised the reality of society, the law and that there is a force that defines what is allowed and what is illegal. At the same time they
served as a guarantor for the continuation of the play, that the potential space continued to exist and that the play remained framed.

Within the group of Polish supporters a delegation took place. The group created a representative who functioned as an actor for the entire group. This person was about to climb the fountain and thereby “win” the game for the Polish group. He was the “player” in the transitional space who was equipped by the group with the power to lead. The group reacted on his signals and requests to sing certain songs, to applaud or to start special whoops. Even the German supporter group chose a delegate who represented the group in the playful argument with the Polish representative (i.e. the Polish group). Here we could observe both processes at the same time. Within each group the members consulted about the next steps. Once put into action they had effects on the other group creating an inter-group process.

One reason why this playful argument did not escalate could be seen in the overall atmosphere and will to celebrate a festival but also in the cautious but clear presence of the police. However, another reason could be found in the capacities of both groups to mentalise and represent the other group and the possible inner motives of the individual members, their understanding and their possible reactions. Despite the conflictive aim of the play (to defeat), the situation calmed down when the German and the Polish fans sang together (i.e. they used the same melodies but different texts). This fact and the presentation of symbols of the other group within the own group (German and Polish flags on the cheeks of fans in both groups) signalled that mentalization processes were intact: Each group recognised the other group as having the same wishes (to win and triumph) and each group was aware of the fact that the other group tried to counteract against the actions of the own group. The song the German supporters began to sing (“steht auf, wenn Ihr Deutsche seid) marked the end of the introduction phase. To stand up (if you are Germans) meant not only to raise from the ground (where actually many Germans were sitting) but also to show and demonstrate one’s powers and strengths. With this gesture the German fans tried to create feelings of inferiority in the enemy/play partner. The others should feel insecure and hesitate. As a matter of fact the Polish fans answered with the same gesture.

At this point the first episode comes to an end. Both groups had agreed upon place and time and about the forms of the play. Both groups had exchanged unifying messages (underlining sameness) as well as diversifying messages of being Polish or German (underscoring the own social identity and the competitive character of the play and the idea that one party will win [= continue to play] while the other party will have to stop to play and fantasise).

Episode 2
The second episode was started by the Polish fans: As the fountain was the agreed upon playground, the Polish fans acted on it. One fan started to climb on top of the fountain. On the one hand the Polish fans had to be aware of the symbolic meaning of being on top, at the highest point, on the other hand they had to be aware of the fact that also the German fans were about to know about the symbolic meaning of this action and that they had to react on it. Thus, the competition was on its way. Here one could clearly distinguish between the framing factors brought and formed by the members of each group (the capacity to form a transitional space on which playing can take place) and factors that steered the interactions while playing (mentalisation) that had very strong connection to (group-) identity processes.

The reaction of the German fans came immediately: They started to hoot, showed their disapproval. They certainly had understood the “provocation” that was implied in the trial to “get on top”, be the best, be highest up and thereby the winner in the end. However, the Germans had no spontaneous answer they could show but hooting. Then, as a second reaction, they started to sing. The symbolic climb seemed to have a strong impact on the Germans. The content of the song contained an “allowance” for the Polish fans to return to their home country. Thus, the German fans fantasised themselves back into a superior position (on top) and to take actions first. However, the attempt failed to remove the Polish fans symbolically and sent them back home. The Polish fan was still sitting on top of the fountain. As a consequence, the Germans started to act in a similar way. One German fan tried to climb the fountain.

The Play-like character of the situation could be seen in this action-response. It was not a large number of Germans that tried to remove the Polish fan from the fountain. It would have been easy to remove the Polish fan from the fountain if the Germans had acted mutually. Instead they continued to “play” and sent their representative in order to solve the problem but to continue to play at the same time.

Here we might state that the exchange of shared symbols, the inclusion of symbols of the other group in the repertoire of the own group may have had a calming effect and enabled to continue to play. The two groups “understood” each other. To put it in different words: The mentalisation processes were continuously successfully working, allowing to take in each others perspective and being able to predict the behaviour of the other group with a high level of likeliness and thus helping the play to continue.

**Episode 3**

As a reaction towards the German fan, which was following the Polish fan up the fountain, the Polish man was climbing the final steps to the top of the fountain and started to sit and dominate the top position. The play was designated to continue when both fans hugged each other as a symbol of peace.
However, the Polish fan did not finish the play but continued to occupy the top position, knowing exactly the Germans experienced this as a loss (defeat) and that they had to do something against this situation. The situation was quite comfortable for the Polish fans as it was not possible for the Germans to take in the top position on the fountain. When the German fan gave way, laughter’s and the incitation to sing expressed the triumph of the Polish fans.

Singing can be seen as a triumph over the other but also as a reassurance to be alive and as a demonstration of strength. This works because every party can expect (mentalise) that the other group understands the actions and symbols in use.

**Episode 4**
The next episode is more an intermezzo, initiated by the Polish when acting and singing.
It become clear that the German group was really about to loose the play. They reacted by disrupting inter-group processes, by taking away the focus from the Polish side and reassured on their own qualities and strengths as a German group (identity): Subsequently, as a result of the within-group process of interaction they started to sing: Germany is he most beautiful country of the world – Germany is the strongest country of the world. It sounded like reassuring themselves their own identity, which was seen as still intact. Somehow, this reconstitution worked and the mood of the group was shifting again towards joy while singing “oh- this is so wonderful/beautiful”.

**Episode 5**
The fifth episode was initiated by the German group, which took up the initial play again. It seemed that the German supporter group had to reconstitute the formal aspects and the capacity for playing, before continuing to interact with the Polish fans. When continuing they displaced the focus of attention to another place for playing during the World Football Championship: Berlin, the city where the final match was about to take place. By declaring that they would travel to Berlin, the German fans implied that they were about to win against Poland. The tendency was clearly directed to leave the actual “playground”. The play at the fountain was lost somehow and the fans orientated themselves towards new play opportunities: to the “real” match that would end with a German victory, thus enabling the German fans to proceed to the final. The functions of these fantasies were to replace the momentary loss by referring to future successes and to regain the mental capacity to play. The moment the German fans became aware that they had lost the play about the top position on the fountain, the “reality” of the loss ended all fantasies and hopes about a future victory. As the intermediate space could not be maintained, the actual stage for playing was left symbolically.
Episode 6

The final phase of the play was reached when the Polish fan showed a symbol that announced the Polish victory. He presented a flag in his hand and put it on top of the statue. As if demonstrating the successful climb on top of the peak of a mountain, the flag was signalling victory, like a flag on a battlefield. The German group reacted with a last, almost desperate attempt to annul this event by trying to cover the whole statue, including the Polish fan and his flag, with a giant German flag. This attempt showed signs of magical thinking, as if one could reverse events by making the results disappear.

As soon as it became clear that this attempt would fail, there was a clearly visible tendency within the German group to be engaged with in-group interactions. As a result, cathexis was withdrawn from the play with the other group, which had a restitution effect on the German fan-group.

Episode 7

The groups still celebrated some time but the overall tendency showed clearly that the affectively laden situation of the “fight” about the fountain was over. This could also be seen in the fact that the police forces had vanished. In what followed one could see a continuation of events of episode one (on the way to the place where the match was about to be observed), however, the groups were engaged with themselves.

Episode 8

In the final episode, which was defined as being part of the play because of the location (the episode still took place at the fountain) the German group was completely focused towards their own members. A leader directed the group and stimulated to sing a carnival song. Being asked about the meaning of the ritual connected to the song (first the Germans shouted every letter of the word “Humba” while sitting on the ground, then they jumped up and danced wildly while singing the song that imitates by onomatopoeia a marching band) the answers were similar: “it is just for fun”. The song has “no deeper meaning”.

However, on a symbolic level one could argue that this episode enacts a restitution process. McPhearson describes different forms of symbolic group activity like body movements, making use of voice, hands, etc. In episode 8 the German fans made use of all of these possible forms of symbolic actions: They jumped, danced, clapped their hands, shouted and sang. It seemed like a mighty recovery, like a phoenix rising from the ashes after a loss. The assumption of the latent meaning of this episode was supported by the time it took place: After the defeat, after the loss of the symbolic play/fight with the Polish group. In order to continue to play, to continue at a different time and place, the group had to recover. Precisely this recovery process is included in the last episode. The
Polish fans had left the scene, the German fans were focused on themselves and reinstalled their identity before continuing to play at another place in the city.

In selecting a carnival song from the times of the German “Wirtschaftswunder” after the Second World War, in which economic recovery took place and that was experienced by many Germans as a miracle, the situation of being defeated and destroyed is turned into its contrary. A similar recovery was experienced psychologically when Germany won the Football World Championship tournament in Switzerland in 1954. In singing this song at the end of the play episode, which ended with a loss for the German supporters, the self-identity of a strong and mighty group is re-installed. By enacting the rise from the ground (defeat) to the dancing party situation (in contrast to destruction, mourning and loss) the parallels are quite striking: On several symbolic levels the restitution processes can be seen not only connected to the loss of the actual play but at the same time – on a symbolic level – as connected to a process that affected a nation (as a mass-group) after the immense destructions during World War II.

To play and to symbolize in a transitional space

The example of the symbolic fight about the fountain help us to describe interactions taking place within each supporter group as well as between the two groups. In a second step we will now discuss our results and relate them to our theoretical frame: symbolic activities within and between groups, mentalisation in group members, play theory, as well as the concepts of transitional space.

Sport events as possibilities for fantasies and emotional catharsis

Mass group events can be described by defining various states within a group event. This can be done e.g. by differentiating gathering, assembling and dispersing processes of by defining the frame and the conditions that have to be given in order to create the possibility for the assembly. Factors of time, duration and space are of interest here. Once masses have gathered and a sport event takes place, certain psychological processes take place that are characteristic for mass situations.

The football field like any other playground, functions as a projection screen for fantasies, hopes, wishes, etc. of the fans in the stadium who attend the match. The same function is true for the television screens of the public viewing areas on which the matches were broadcasted and collectively watched. Football field as well as TV-screens served as catalysts in order to focus attention and projections of thousands of supporters watching the game.

As the masses gathered in a stadium or in a public viewing area at the same time not only the location for the play was defined but also the time when the play was about to end – after 90 minutes, when the referee ended the match.
In the hours preceding the match, the processes of symbolic play had developed spontaneously, whenever two supporter groups had met. The fight about the fountain was one example for these types of plays and symbolic interaction. The two groups agreed upon time (duration) and space (location) for the play. This was in the gathering phase. During the assembling process the factors time and space for the play were defined by the world football organisation in cooperation with the local authorities. They all watched the game (in the stadium, in the public viewing areas, on hundreds of other places in the city. They all stared on a circumscribed area where 22 players fought about a ball. They all got involved emotionally, shouted, hoped, felt fear, chock, relief, in the end ecstasy, joy and happiness or desperation and mourning. One could clearly see how affectively involved all spectators got (c.p. Granström & Hylander 2007) that they projected parts of their inner feelings, emotions and psychological conditions onto the playground, thus participating in the play. This engagement of the spectators is not arbitrary. The playground and the actions taking place supply a flexibility and openness for various kinds of identification and projection processes, which vary and change during the course of a match in their intensities and their vitalities.

One of the major advantages of a projection screen is that what is observable allows multiple ways of interpretation. A football match can be seen as a war, as a play, as a hunt. The ball can be seen as a weapon for attack, as an object that has to be defended, as a shared valuable tool, etc. Many different interpretations could be added. Listening to the vocabulary of sport reporters is a good way of learning to know some more of the projected contents. What makes the sport so attractive is the quick change in projected contents, the free-floating possibilities to get rid of emotional pressures, the intensity of vitality affects (c.p. Stern 1995).

It requires certain preconditions that a football or sport event in which masses take part as spectators, remain peaceful and joyful and are experienced like a fest, a party rather than a fight or war (c.f. Granström & Hylander 2007, Guvå & Rosander 2007), in other words, that the playground remains a playground, that it is interpreted as such by all participants during the entire time of playing.

Football can create a collective experience of power and ecstasy as well as deep falls into abysses of depression, desperation and mourning. All takes place in the collective situation of a large group, a mass. However, what is open for discussion, to what extent psychological capacities within the individual group member remain intact, which features of the self and the identity of the individuals remain the same and to which extent regressive phenomena can be observed, especially in case of outbursts of violence.

Mental processes within the group members
Focussing on the actions that take place between groups and relate them with an interactive perspective seems not enough when trying to explain group behaviour. The problem with the phenomenological analyses is that they start at a point in the process of events where several steps of the process have already taken place, above all internal processes in the members of the groups. They have to be assumed because no human behaviour takes place independently from inner processes. Therefore, we go a step further by focussing on processes within the group members that have to be assumed when group behaviour within a group and between groups is analysed. The factors constituting an external frame are accompanied by factors constituting an internal frame of reference. These factors can be described by the capacity to symbolise. Let us go back to the scene at the fountain. The members of both groups were able to create an interaction with each other in order to understand, to negotiate and to agree upon e.g. conditions of play or meanings of symbols. In order to act and react in an adequate way, the participants of the play must be able to develop hypotheses about the inner mental states of the other participants (from the own group as well as from the other group). The mentalisation capacity is connected to fantasies and assumptions about the inner world of another person, his or her inner motives, wishes, psychological reactions, etc. Only this capacity allows to anticipate possible reactions of the other, which can be seen as initiated by own actions and reciprocally prepare reactions to possible assume actions from others.

From this perspective the behaviour to sing the same song with different texts as an initial event creates from an internal point of view an intermediate space within the group members. They identify each other as football fans (sharing the same melody) but at the same time they also point out the difference (Polish and German texts). Even they might not understand the content of the text the other group sings one can be sure that both groups assumed that the other group was singing the same, some sort of text that strengthens the self and the identity as fans.

When in episode 2 the Polish fan is starting to climb the fountain, again this can only be the result of complex within group processes as well as inner processes of the group members. The Polish group decided to send a representative into the play. Thus, they must have come to the conclusion that the German group is ready to play, they must have developed (by within group communication) the idea of letting one of the members play, representing the entire group (i.e. creating a transitional playground [fountain] onto which the members can relate, identify, project fantasies) and they must have expected that the German supporters will understand the symbolic play of climbing the fountain. Playing, creating a transitional space for play, that is neither completely identical with the external reality nor with the internal world, implies a number of mental activities within the individual.
In states of high arousal, anger, rage, fear, and anxiety the capacity for these mental capacities are reduced. In a relaxed atmosphere, in the context of a fest, these mentalisation steps are possible for all participants. The initial negotiation process between the groups with aim to play has a calming effect by itself. The Polish group does not respond aggressively, instead they send one person to “play” which signals clearly that the group is ready for play. At the same time, the Polish supporters must have assumed that the Germans will understand their signals and they must have calculated that the Germans will start to playfully respond.

**Symbolic play activity**

Once the groups have gathered in a stadium or in a public viewing area they start to interact. If they do not fight directly with each other the interactions takes place mostly on a symbolic level. In Dortmund it was the strategy of the organisers that the different supporter groups should have possibilities to meet, mingle and communicate before the match (“a time to make friends”). The idea was to create possibilities for interactions, communications – and mutual play. The fans made use of these possibilities in many different ways. We observed different modes of symbolic interactions: dancing, singing, presenting different dress-codes, waving flags, showing symbols of the own national team, acting by gestures, clapping, running, etc. The fans used their entire body in order to express in a playful way the fantasies and wishes of each supporter group. These interactions can only take place once when time and space, the overall frame is defined. The organisers of an event can create this frame, but it can also be “organised” by the different groups.

The groups have to negotiate and to agree upon the place where the play should take place. This is not arbitrary and in most cases leads to interactions at places that have a symbolic relevance. The next factor that is important is time. The play has to be framed by a starting signal as well as by a stop signal. While the starting signal is created out of a flow of preparative interactions, the stop signal is quite evident: Once a group has won/lost, the play is over. This is the case when no more fantasies about a future victory are possible, when reality succeeds (e.g. when the referee ends the game).

In an intergroup perspective the Polish fans were sharing the same fantasies as the German fans and social identity was created (the situation process). This we-ness came out of the other-ness, and vice versa. The situation with an evident other-group and some more unconscious clues that lead to perform the play, was after a while not the same. The Polish fans (the we-group) did not back up the climber with the flag, the police (gatekeepers) left the market and the German fans (the other-group) ignored the climber and started to sing their own songs. These circumstances ended the play and the climber got down from the fountain. We can imagine the situation if the police had interfered more
aggressively against one or both groups, or a situation where none of the (anti) groups had done peaceful activities. Both of these situations had probably lead to confrontations. The overall culture in Dortmund that day (the festival) where also in favour of play similar to the play of the child. These, and other structural influences, as time and place, are here facilitating play.

**Conclusion**

All these processes described above must be located in the mental lives of individuals forming a group. It would be interesting to further investigate how the “decisions” are made, e.g. sending this one specific person to climb the fountain. However, it becomes evident that we must take these internal processes into consideration when we try to understand mass behaviour and play, especially with respect to violent outbursts.

We suggest that in the case of a peaceful situation both groups create a transitional space where the play action can take place. As long as this space is maintained, a space that is neither identical with inner realities nor is it identical with outside reality, play activities can continue. Reversely, as long as two groups play, they can be located on the transitional space. Thus, the danger of real aggressive showdowns can be minimised.

**References**


de Saussure, F. (1922). *Cours de linguistique générale*. In B. Séchehaye (Ed.), *Notebooks of Saussurre's students 1907-1911*.


psychic reality in borderline patients. *International Journal of Psycho-
Analysis, 81*, 853-873.

psychic reality in borderline patients. *International Journal of Psycho-
Analysis, 81*, 853-873.

Press.

Freud, A. (1965). *Normality and pathology in childhood*. Harmondsworth,
Middx: Penguin.


scottish fans in Sweden. *International Review for Sociology of Sport, 30*,
191-217.

& social Issues, 26*, 25-46.

Hau, S. (2008). *Communication as the most important Police Strategy at the
Football World Cup Final 2006* (FOG-Report no 61). Linköping: Department of
Behavioural Sciences and Learning (IBL), Linköping University.


Klein, M. (1975/1946). Notes on some scizoid mechanism. In M. Klein (Ed.),

M. Klein (Ed.), *Love, guilt and reparation, and other works 1921-1945* (pp.
344-369). New York: Macmillan


257-278.

and Social Psychology Review, 6*, 88-106.


and Violent Behavior., 9, 353-378.

Author note
Stephan Hau, Associate Professor, PhD, Psychoanalyst (IPA, SPF, DPV), Psychologist, licensed psychotherapist, University of Linköping (Department for Behavioural Sciences and Learning – IBL). stephan.hau@liu.se

Johan Näslund, PhD, Linköping University (Department for Behavioural Sciences and Learning – IBL). johan.naslund@liu.se
This research was funded by the Swedish Emergency Management Agency (KBM).
ATTRIBUTIONAL STYLE, ACADEMIC SELF-EFFICACY, AND ATTEMPS TO INFLUENCE THE STUDY ENVIRONMENT

Tomas Jungert & Stefan Gustafson

Abstract

Relationships between students’ attributional styles, self-efficacy and strategies to influence and take control over their study situation are explored and a comparison between female and male students investigated. Participants were 271 students enrolled in two Masters Programmes in Engineering. The participants completed a questionnaire measuring academic self-efficacy, a questionnaire measuring strategies to influence the study situation and a questionnaire measuring satisfaction with study results and attributional style. Internal and global attributional styles were related to higher self-efficacy and higher beliefs in opportunities to influence the study situation. External attributional style was associated with formal strategies to influence the studies while attributions on the globality dimension were associated with social strategies to take control over their studies. Female students who perceived their study results as successful were more likely than male students to attribute their results as unstable.

Keywords: Attribution, self-efficacy, student influence, higher education

Attributional style refers to how people explain the causes of their own performances. It may be defined as the causal explanation that individuals give to various intrapersonal and interpersonal actions in their lives (Bell-Dolan & Anderson, 1999). In this paper, the focus is on causal attributions on three attributional dimensions, which are believed to be significant in an academic context (Peterson, Buchanan & Seligman, 1995); the locus dimension (e.g. Heider, 1958; Rotter, 1966; Tollefson, 2000), the stability dimension (e.g. Reyna, 2000; Weiner, 1986; Weiner et al., 1972), and the globality dimension (e.g. Abramson, Seligman & Teasdale, 1978; Foster, 2000). Earlier research has found relationships between students’ attributional styles and their academic self-efficacy (Schunk & Cox, 1986) and to study strategies (Ferla, Valcke & Schyten, 2007). A number of studies have also found differences in attributional style between female and male students (Howell, 1993). However, no earlier studies have investigated the relationship between attributional style and strategies to influence and take control over the study environment. In this study, we will explore the relationships between students’ attributional styles, their self-efficacy, and their strategies to influence and take control over their study environment, and investigate whether there are differences between female and male students.
Theoretical background

The locus dimension refers to whether causes are internal (e.g. by personality factors such as ability and effort) or external (e.g. by situational circumstances such as luck). The stability dimension refers to causes that are fluctuating or stable over time (Bell-Dolan & Andersson, 1999). Both attribution on the locus dimension and stability dimension have been found to be related to success expectancy. Individuals with internal attributions for success tend to show higher success expectancies for the future than individuals with external attributions for success (Anderson & Weiner, 1992). Stable attributions for success have been related to increased success expectancies and achievement motivation, while stable attributions for failure tend to result in decreased success expectancy and achievement motivation (Weiner, 1986). Individuals who attribute their causes of achievement as global believe e.g. that it will influence their results in other university courses and in other contexts, whereas individuals who explain causes of their achievement as specific believe that it is going to influence that specific situation only. Peterson, Buchanan and Seligman (1995) suggested that individuals who give external, unstable and specific explanations to particular events have a negative attributional style while a positive attributional style refers to internal, stable and global explanations because such attributions involve a greater sense of control.

Previous research has demonstrated that a distinction needs to be made between attributions to failure and to success (Peterson & Barrett, 1987). Research has found that students who attribute failure internally tend to focus on failure and do not consider ways to remediate it. Furthermore, Diener and Dweck (1978) found that mastery-oriented children attributed their failure to lack of effort while helpless children attributed failure to incompetence. Also, children who attribute their failure to lack of effort tend to avoid future learning situations in which failure is expected to recur.

Student levels of attributions and self-efficacy for academic success have been found to partially determine their study strategy. For example, Ferla, Valcke and Schyten (2007) found that students with a reproductive conception of their learning, i.e. a reproductive study strategy, attributed academic success to external causes. This raises the question whether students who use different strategies to influence their study situation attribute study results to different causes.

Attributions are further related to self-efficacy beliefs. Self-efficacy is the students’ judgments of their own capability to organize and perform courses of action necessary to reach selected types of performance, e.g. pass an exam or learn a course-content (Bandura, 1986; 1997). Self-efficacy concerns the students’ perceived capabilities to reach goals and achieve results and affects their choices of activities, efforts, or preservations of behaviour. Efficacy beliefs furthermore influence how students think, feel, motivate themselves and act.
(Bandura, 1995). Self-efficacy differs from other similar constructs, such as self-esteem and self-confidence, as it is more predisposed to the contextual factors and concerns a specific goal and influences effort and persistence (Bandura, 1997).

It is assumed that people with a high level of self-efficacy try challenging tasks more frequently and persist for longer with them (Bandura, 1986). Furthermore, they recognize that they can overcome obstacles and focus on opportunities, which is why they perceive stressful situations as more challenging than people with low self-efficacy (Jerusalem & Schwarzer, 1992). Students’ self-efficacy concerned with a certain task has been found to be related to effort-attributional feedback (Schunk & Cox, 1986). When students received effort-attributional feedback for problem-solving successes their self-efficacy increased.

Research has found that girls tend to attribute their success in Mathematics and the Natural Sciences to external variables and their failures to internal, uncontrollable and stable causes more often than boys (e.g. Lloyd, Walsh & Shehni Yailagh, 2005). Numerous studies have additionally suggested that females are more likely than boys to attribute their successes in Computer Science to luck and their failures to lack of ability (e.g Bernstein, 1991; Howell, 1993). A more recent study found no significant gender differences in attributional style (Wilson, 2002). Other studies have found differences in girls’ and boys’ self-efficacy as regards in particular mathematics (Pajares, 1996).

**The present study**

Self-efficacy was a focus area in a previous longitudinal project, with the overall purpose to compare the expectations and experiences of four cohorts of engineering students (author, 2008 (blind review)). In conducted interviews in that study, students describe their beliefs in their capabilities and how they relate to various forms of influencing their study situation, e.g. (a) To fill in the course evaluations, (b) To seek contact with teachers out of class to influence courses, (c) To seek contact with teachers in class to influence courses, (d) To seek support from class-mates, and (e) To talk about the study situation with peers. In that study (author, 2008) students’ achievement motivation and beliefs in their capabilities to achieve were related to their strategies to influence and take control over their studies. Based on results from that study, a questionnaire was developed in order to explore relationships with student strategies to influence their study situation and other relevant variables. In a recent study (author, in press) an informal strategy to influence the studies, e.g. to seek contacts with teachers to influence courses, was related to a strategic approach to studying, achievement orientation and monitoring effectiveness and an informal strategy to influence the studies, e.g. to fill in course evaluations, was associated with high levels of achieving, indicating that they were very motivated to achieve
well. Earlier studies on strategies to influence and take control over the study situation thus has found relationships between use of strategy to influence the studies and variables such as motivation, strategic approach to studying and monitoring effectiveness.

**Aims and hypotheses**

In the present study, the purpose is to explore if students’ causal attributions of their study results are related to their strategies to influence and take control over their studies and their self-efficacy beliefs. Therefore, three general factors were studied: (a) how students attributed the causes of positive and negative study outcomes, (b) the strategies students used to influence and take control over their study situation, (c) their levels of academic self-efficacy. In addition, this study explored if these factors demonstrate gender differences and the relationship between attributional styles, strategies to influence the study situation, and academic self-efficacy. Based on findings of the longitudinal study (author, 2008; in press) and on research on attribution theory and academic self-efficacy, it was expected that there would be a positive correlation between the internal, stable and global attributional style and their level of academic self-efficacy. It was furthermore expected that internal, stable and global attributions to positive study outcomes would be positively correlated to beliefs in good opportunities to influence the study environment and the use of more strategies to influence and take control the study environment. Based on earlier research on gender differences in how students attribute their study outcomes, a final hypothesis was that female students to a greater extent than male students would attribute negative study outcomes to internal, stable and global causes.

**Method**

**Participants**

A total of 271 students from two programmes participated in the study. The students were enrolled in the first, the third and the fifth semester in a 4.5 year Master’s programme in Applied Physics and Electrical Engineering and in the first and fifth semester in and a 4.5 year Master’s programme in Computer Engineering. In week one of the semester, the students filled in a Self-efficacy questionnaire because self-efficacy beliefs are future oriented expectations for taking personal control over achievement tasks. When the students had had their first exam in the semester, they filed in the second questionnaire measuring opportunities to influence and take control over the study situation and the third questionnaire measuring attributional styles. Table 1 records the response rates. Both programmes had few women enrolled (9 and 12 % respectively). It was reflected in the responses (around 12 % of the answers were from women for the questionnaires). The average age of the participants was 21.3 years ($SD = 2.0$).
Table 1
Response rates for male and female students on each of the three questionnaires and on all of them

<table>
<thead>
<tr>
<th>Participants</th>
<th>Responses</th>
<th>1st Q</th>
<th>2nd Q</th>
<th>3rd Q</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>34</td>
<td>30</td>
<td>23</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>Male</td>
<td>237</td>
<td>199</td>
<td>193</td>
<td>156</td>
<td>132</td>
</tr>
<tr>
<td>Total</td>
<td>271</td>
<td>229</td>
<td>216</td>
<td>175</td>
<td>150</td>
</tr>
</tbody>
</table>

The students filled in the questionnaires with informed consent and the data were coded by the researchers to make it confidential.

Measures

Academic self-efficacy A self-efficacy scale was utilized to assess the participants’ general academic self-efficacy and their semester specific academic self-efficacy. The scale had twenty items. The first ten items were adapted from the Swedish version of the General Self-Efficacy (GSE) scale, developed by Schwarzer and Jerusalem (1995). The original scale measures a general self-efficacy, in which participants are asked to rate their confidence in successfully solving general, “everyday” problems and reach general, “everyday” goals. The adaptation involved changing the referent to “everyday” problems and goals to study related problems and goals. For the original scale, high reliability, stability, and construct validity have been established in numerous studies (e.g. Luszczynska et al., 2005; Schwarzer et al. 1999). An example item is “When I am confronted with a study related problem, I can usually find several solutions”. The following ten items concerned a more specific, semester related academic self-efficacy. Students were asked to rate their confidence in being able to successfully learn and perform well in various elements during their semester. An example item in this questionnaire is “I am confident that I have the capability to learn the content of the courses this semester”. Ratings were recorded on a 7-point Likert continuum (1 = no confidence, 7 = complete confidence). These items were developed by the authors in order to measure perceived self-efficacy conceptualized in the current semester (i.e. a situation-specific manner, which is recommended by Bandura (1997; 2005)).

An explorative factor analysis of the items covering self-efficacy resulted in the two factors; general academic self-efficacy and semester specific self-efficacy when the result was rotated using varimax. The factors were extracted using a principal component analysis. The reliability of the two factors was
calculated by the use of Cronbach alpha. Both factors were suitable as scales as the value was .88 for the first factor and .90 for the second factor.

Influence To measure students’ strategies to influence and take control over their study situation, this measure has been developed (author, in press). This questionnaire was distributed to the participants during a certain week in the semester when the students had had at least one exam. In general, this week occurred during the sixth week of the semesters. The questionnaire comprised seven items. In the first question, students were asked to rate how great their opportunities to influence their study situation were on a 7-point Likert scale (1 = no confidence, 7 = complete confidence). Question two to seven assessed how students acted during the specific week (an item for influencing during the week, for example, was “I sought contact with teachers in class in order to influence courses”). For each item, students were asked to answer “yes” or “no”.

Academic attributional style In order to measure the participants’ academic attributional style, a concluding measure comprised of four items was developed. Three items assessed internal/external, stable/unstable, and global/specific attributional styles. The item indicating external/internal attribution was, for example, “My study results this semester depend on my capability and my efforts rather than on external factors”. As research has found that it is important to distinguish attributions of successful academic results and weak academic results (Weiner, 1986), a fourth item was added to the questionnaire. The fourth item concerned the participants’ satisfaction with their academic achievements during the specific semester. This item was necessary in order to determine differences between causal attributions to study results that the students were satisfied with and study results that they were not satisfied with. Responses were scored on a 7-point scale ranging from “do not agree at all” to “fully agree”.

Participants who scored one to three on the satisfaction item were categorised as students who attribute negative outcomes and participants who scored five to seven on the satisfaction item were categorised as students who attribute positive outcomes (i.e. successful study results). Separate analyses have been performed for students who attribute positive outcomes and negative outcomes.

Analysis The analysis included descriptive statistics for the total sample. T-tests were computed to determine if there were any differences in attributional style and strategies to influence and take control over the study situation between female and male students. In these analyses, we generalize to Swedish students in Master’s programmes in Engineering.
Pearson’s correlation coefficients were computed to assess the relationships between attributional style, strategies to influence and take control over the study situation, and academic self-efficacy. T-tests were computed to assess if students who had used certain strategies to influence and take control over their study situation differed in attributional style from students who had not used such strategies. A series of two-way ANOVAs were conducted to evaluate the relationships between the independent variables gender, strategy to influence the study situation and satisfaction with study results and the dependent variables attributional style and self-efficacy.

Results

Academic Attributional style and strategies to influence the study situation

Table 2 shows the mean values for the attributional styles. These results indicate that the students’ in general attribute their study results more to internal, stable and global causes than to external, unstable and specific causes. Students with positive study outcomes attributed their outcomes to more internal, stable and global causes than students with negative outcomes. The mean score for perceived opportunities to influence the study situation was rather high, which means that the students in general believe that they have rather good influencing opportunities.

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Locus, positive outcomes</td>
<td>98</td>
<td>5.63</td>
<td>1.25</td>
</tr>
<tr>
<td>Locus, negative outcomes</td>
<td>51</td>
<td>4.78</td>
<td>1.67</td>
</tr>
<tr>
<td>Stability, positive outcomes</td>
<td>92</td>
<td>5.64</td>
<td>1.57</td>
</tr>
<tr>
<td>Stability, negative outcomes</td>
<td>48</td>
<td>4.50</td>
<td>1.88</td>
</tr>
<tr>
<td>Globality, positive outcomes</td>
<td>91</td>
<td>5.43</td>
<td>1.55</td>
</tr>
<tr>
<td>Globality, negative outcomes</td>
<td>49</td>
<td>4.45</td>
<td>1.77</td>
</tr>
<tr>
<td>Opportunities to influence</td>
<td>216</td>
<td>4.56</td>
<td>1.41</td>
</tr>
<tr>
<td>General academic self-efficacy</td>
<td>229</td>
<td>5.21</td>
<td>0.96</td>
</tr>
<tr>
<td>Semester specific self-efficacy</td>
<td>229</td>
<td>4.59</td>
<td>0.89</td>
</tr>
</tbody>
</table>

In order to examine the relationship between academic attributional styles and opportunities to influence the study situation, Pearson’s correlation coefficients were computed. For students who attributed positive outcomes (i.e. who were satisfied with their study results), significant correlations ($p < .05$) were obtained between perceived opportunities to influence the study situation and stable attribution (.21) and global attribution (.27). This indicates that students who
attributed positive outcomes to stable and global causes believed more in their opportunities to influence their study situation than students who attributed positive outcomes to unstable and specific causes. For students who attributed negative outcomes (i.e. who were dissatisfied with their study results), no significant correlations were obtained between perceived opportunities to influence the study situation and attribution on any dimension.

The computed t-tests showed some differences in attributional style between students with different strategies to influence and take control over their study situation. Students who attributed positive outcomes and tried to influence their study situation by talking with peer students about their study situation and cooperating with them had a more global attributional style than students who did not try to influence their study situation with that strategy, \( t(87) = 2.47, \ p < .05 \). The global attribution mean value for students who had used that strategy to influence and take control over their studies was 5.68 compared to 4.81 for students who had not used that particular strategy. On the other hand, students who had looked for social support from peer students during the specific week had a more specific attributional style, i.e. less global, than students who had not looked for social support from peer students, \( t(87) = 2.47, \ p < .05 \). The global attribution mean value for students who had used that strategy to influence and take control over their studies was 5.20 compared to 6.15 for students who had not used that particular strategy.

Students who attributed negative outcomes and tried to improve their courses by filling in course evaluations had a more external attributional style (the mean value was 5.68) than students who had not tried to improve their courses by that strategy (the mean value was 4.81). This difference was significant, \( t(48) = 2.16, \ p < .05 \). Students who attributed negative outcomes and tried to improve their study situation by seeking contact with teachers in class had a less stable attributional style \( t(45) = 2.04, \ p < .05 \). The mean value for students who had used that strategy to influence and take control over their studies was 6.00 on the stability dimension, compared to 4.26 for students who had not used that particular strategy.

Academic attributional style and academic self-efficacy

As shown in Table 2, the mean values for self-efficacy beliefs were high. In general, the students believe in their capabilities to achieve well in the programme and in the current semester.

In order to examine the relationship between academic attributional style and academic self-efficacy beliefs, Pearson’s correlation coefficients were computed. A series of two-way ANOVAs were conducted for attributional style, satisfaction with study results and self-efficacy. None of these analyses that explored the relationships among these variables were significant.
Differences between female and male students

T-tests were calculated in order to assess if there were any differences between female and male students as regards their attributional style, self-efficacy and perceptions of opportunities to influence the study situation. No significant differences were obtained for any of these variables, see Table 3.

Table 3
Means, and Standard Deviations of Self-Efficacy and opportunities to influence the study situation for female and male students

<table>
<thead>
<tr>
<th></th>
<th>Sex</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semester specific academic self-efficacy</td>
<td>Female</td>
<td>30</td>
<td>5.08</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>199</td>
<td>5.22</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>4.44</td>
<td>0.80</td>
</tr>
<tr>
<td>General self-efficacy</td>
<td>Female</td>
<td>199</td>
<td>4.61</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opportunities to influence</td>
<td>Female</td>
<td>23</td>
<td>4.83</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>192</td>
<td>4.53</td>
<td>1.41</td>
</tr>
</tbody>
</table>

When investigating how people attribute their study results, it is, however, important to distinguish between attributions of success and failure. A series of two-way ANOVA designs were thus used to estimate if there were any interactions between the two independent variables, gender and satisfaction with study results, as regards attributional style on the three dimensions and self-efficacy. An interaction effect was obtained for gender and satisfaction with study results concerning attributional style on the stability dimension, $F(1, 140) = 6.51, p < .05$ (see Table 4).

Table 4
Means, and Standard Deviations of attribution for female and male students

<table>
<thead>
<tr>
<th>Attribution</th>
<th>Study Outcomes</th>
<th>Sex</th>
<th>n</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stability*</td>
<td></td>
<td>Female</td>
<td>7</td>
<td>2.86</td>
<td>1.68</td>
</tr>
<tr>
<td>Negative</td>
<td>Male</td>
<td>41</td>
<td></td>
<td>4.72</td>
<td>1.79</td>
</tr>
<tr>
<td>Positive</td>
<td>Female</td>
<td>10</td>
<td></td>
<td>5.90</td>
<td>1.29</td>
</tr>
<tr>
<td>Positive</td>
<td>Male</td>
<td>82</td>
<td></td>
<td>5.61</td>
<td>1.60</td>
</tr>
<tr>
<td>Negative</td>
<td>Female</td>
<td>7</td>
<td></td>
<td>4.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Negative</td>
<td>Male</td>
<td>44</td>
<td></td>
<td>4.91</td>
<td>1.60</td>
</tr>
<tr>
<td>Positive</td>
<td>Female</td>
<td>11</td>
<td></td>
<td>5.36</td>
<td>1.43</td>
</tr>
<tr>
<td>Positive</td>
<td>Male</td>
<td>87</td>
<td></td>
<td>5.67</td>
<td>1.23</td>
</tr>
<tr>
<td>Negative</td>
<td>Female</td>
<td>7</td>
<td></td>
<td>3.43</td>
<td>1.90</td>
</tr>
<tr>
<td>Negative</td>
<td>Male</td>
<td>42</td>
<td></td>
<td>4.62</td>
<td>1.71</td>
</tr>
<tr>
<td>Positive</td>
<td>Female</td>
<td>10</td>
<td></td>
<td>5.2</td>
<td>1.32</td>
</tr>
</tbody>
</table>
Both male and female students who were satisfied with their study results attributed their perceived success to highly stable causes. Female students who were not satisfied with their study results attributed their perceived failure in their course to significantly less stable causes than male students who were dissatisfied with their study outcomes. No interaction effect was obtained for gender and satisfaction with study results concerning attributional style on the locus or globality dimensions. For complete information about the ANOVAs, see Table 5.

**Discussion**

Earlier studies have found indirect relationships between attributional style and study strategy (Ferla, Valcke & Schyten, 2007). In the current study, students who tried to take control over their studies by seeking cooperation with peer students attributed the causes of positive outcomes as global. On the other hand, students who did not cooperate with peer students in order to influence or take control over their study situation attributed the causes of positive outcomes as specific, i.e. they did not believe that their good study results would have an impact on other parts of their lives. Students with positive study outcomes who sought social support from peers in order to influence and take control over their study situation, however, had a more specific attributional style, whereas students who did not look for social support had a more global attributional style.
Students who attributed successful study results to global causes believed that their study results were important for their lives in general. It is likely that students with such an attributional style have good relations with peer students and interact with them because they are highly integrated in the student community. They believed in their opportunities to influence their studies, but at the same time, they did not use any of the other strategies to influence their study situation that were assessed in this study. They did, however, collaborate with peer students. The students in this study who had a global attributional style show similarities with students who had a cooperative approach to influencing their studies in another study with a longitudinal and qualitative design (author, 2008). Both students with a global attributional style and students with a cooperative approach to their studies (author, 2008) emphasised cooperation with peer students as important in order to enhance their feelings of control because it increased their sense of sharing a highly demanding study situation with peer students and that they could manage problems together.

Students who did not cooperate with peer students but were looking for social support attributed the causes of their study results as specific to their study situation. This signifies that their study results were not highly important for their lives in general. These students may be students who are not as integrated with the student community. They study alone either because they prefer to or because it is more convenient for them. To seek social support from peers may be a strategy to keep some relations with the student community.

The reason why students who we believe are integrated in the student community did not look for social support may be because social support seeking was not considered a strategy to influence or take control over the study situation. Students who are highly integrated in the student community and who cooperate well with their peers may not consider being social with peers as seeking social support, but something that they naturally do more or less continuously in everyday life. However, actively cooperating and discussing courses with peer students nevertheless is a strategy to take control over the studies (author, 2008; in press).

For students who attributed negative outcomes, differences in attributional style were found between students who filled in and did not fill in course evaluations as well as between students who sought contact with teachers in class and students who did not. Students who attribute negative outcomes to external causes may have filled in course evaluations more often because they did not believe that their own efforts or capabilities to achieve well were sufficient. By filling in course evaluations, they may have anticipated that this could be significant for the future design and quality of courses in the programme. It was a strategy to increase their opportunities to succeed better in the future. Typically, different courses have different teachers in a Swedish higher educational context. The teacher may thus be considered as an unstable
factor rather than as a stable factor. The students who had tried to influence their teacher but later failed in their course, attributed their failure to an unstable cause. Perhaps these students exaggerated the importance of the teacher in their studies.

Attributional style did not correlate significantly with academic self-efficacy, which was unexpected since earlier research has found that individuals with internal attributions for success show higher success expectancies (Anderson & Weiner, 1992).

The percentage of women in this study is only about twelve percent and reflects the percentage of women enrolled in the two programmes in this study. A consequence of the shortage of female students in this study is that the comparisons between sexes should be interpreted with some caution. The results suggest only one significant difference between men and women as regards the variables that were investigated in this study. The one significant difference that was found concerned attribution on the stability dimension. Female students who were not satisfied with their study results tended to have a more unstable attributional style than male students who were not satisfied with their study results. This result is not line with research on gender differences and attributional style. Findings of this study do thus not lend support to earlier studies that have found that females more often than males attribute their successes in computer science and natural sciences to luck and their failures to lack of ability (Bernstein, 1991; Howell, 1993; Lloyd, Walsh & Shehni Yailagh, 2005). Other studies on gender and attribution have been carried out with other questionnaires and in other contexts than the Swedish higher educational context. It could, however, be claimed that the small percentage of women who do decide to continue studying in male dominated fields such as Masters’ programmes in Engineering, have more self-confidence and higher self-efficacy beliefs than academically capable women who chose to enrol in other domains. This could explain why only one difference between female and males was obtained, and that this finding is contrary to earlier research on gender and attributional style. Another possible explanation for these results is that teachers attempt to match coursework and exam questions and give rich feedback so that their students will not perceive that their success was a coincidence or an isolated case.

Overall, the results show that students who attribute negative study outcomes demonstrate different attributional styles than students who attribute positive study outcomes. This finding is in line with earlier research (Peterson & Barrett, 1987) and underlines the importance to distinguish between attributions of causes to negative and positive outcomes.

Limitations
It should be noted that some of the reported statistically significant results, such as the correlations below .2, were not very strong even though they were significant.

This study is limited to Swedish higher education, and although the findings are relevant and of interest to an international audience, differences do exist between Sweden and other countries as regards the designs and context of programmes in Masters of Engineering.

Conclusions

The present study shed light on the patterns of attributions and may increase the understanding of links between attribution and strategies that students use to influence their study situation. These findings can be important for understanding how students perceive their opportunities to influence their studies and when designing courses. As cooperating with students seems to be connected with a global attributional style, more courses could be designed as projects in order to increase student cooperation. Furthermore, when feedback is given to students about academic achievements, it should be specific, concern what the students may improve, and be given as soon as possible.

Students having a global attributional style cooperated more with their peer students than students with a specific attributional style. It may be claimed that students who attribute study results globally are integrated in the student community, which has many positive connotations. To integrate students in the student community can be facilitated by various means, e.g. by designing courses so that more students cooperate, designing the campus so that more can live in the student community, and by making more work spaces available to increase cooperation opportunities.

It is somewhat surprising that there was only one difference between females and males given the previous research that has found several differences between female and male students (Bernstein, 1991; Howell, 1993; Lloyd, Walsh & Shehni Yailagh, 2005). Furthermore, unlike previous research, female students in the present study had a more positive attributional style than male students. It could be argued that the small percentage of women who persist their studies in male dominated programs have higher self-efficacy beliefs and more positive attributional styles than women who choose other careers. Future research needs to explore differences in self-efficacy and attributional style between female and males before they enrol in study programs in higher education.

References


I’m so smart, why don’t I know it? Canadian Journal of Education, 28(3), 384-408.


Tomas Jungert, MSc, research student and Stefan Gustafson, PhD, senior lecturer, Department of Behavioural Sciences and Learning, Linköping University.

Thanks are due to the students in the two programs who participated in the study and to the referees of Scandinavian Journal of Psychology who gave helpful suggestions about the text.

Correspondence concerning this article should be addressed to Tomas Jungert, Department of Behavioural Sciences and Learning, Linköping University, Sweden. E-mail: tomas.jungert@liu.se
THERE IS MORE TO FAIRNESS IN TAXATION THAN FAIR TAXES: INTRODUCING A MULTI-FACETED FAIRNESS FRAMEWORK OF TAXATION

Ali Kazemi

Abstract
Taxation has attracted considerable research attention. In this article some preliminary steps are taken towards advancing a multi-faceted fairness framework of taxation according to which taxation fairness is not only related to the tax expenditure programme (distributive justice), but also to how tax decisions are made (procedural justice), how citizens are treated by the tax authorities/employees (interpersonal justice), and how tax decisions and the legislation are explained and justified to the citizens (informational justice). This framework further posits that taxpayers experience uncertainty in the context of taxation, and that fairness reduces uncertainty. It is also posited that fairness is a prerequisite for legitimacy of the tax system. The framework draws upon the literature on the social psychology of justice and applies it to the taxation context. Altogether, this theoretical framework provides an alternative account of tax attitudes and behavior in that it downplays the role of surveillance and sanctions and instead emphasizes the role of perceived fairness for citizens’ willingness to pay taxes and for tax system legitimacy.

Keywords: taxation, justice, fairness, legitimacy, uncertainty, trust

Three things are certain in life: death, taxes, and mankind’s unrelenting effort to evade both (Klepper & Nagin, 1989).

Introduction
The issue of taxation has attracted considerable interest and research due to its centrality in realizing the goals of the Welfare state in modern societies. Traditionally, research efforts were aimed at identifying factors affecting tax evasion (defined as deliberate and criminal acts of not paying tax liabilities) (Slemrod, 1992). This early research, guided by the rational choice theory emphasizing individual rationality and pursuit of self-interest as the dominant motive for human behavior (Allingham & Sandmo, 1972) was shown to be a too limited perspective. The social regulation and enforcement promoted by this perspective was the deterrence approach, primarily emphasizing employment of threat, coercion, and extrinsic material incentives. According to this line of thinking the tools that the tax agency can use to regulate citizen behavior are tax rates, perceived probability of detection in case of evasion, legal consequences, and severity of these legal consequences. However, a system that is largely based on surveillance and sanctions is likely to be too costly. Moreover, it does not always result in detecting evasion and may, in addition, produce
psychological reactance. Thus, such a system may lead to opposite and retaliatory behaviours, if it is perceived as illegitimate (Brehm & Brehm, 1981).

Instead of focusing on the problem of tax non-compliance, modern social and behavioral scientific theorizing stresses the role of norms, trust, justice, and morality in enhancing tax compliance which is a more positive focus (e.g., Cullis & Lewis, 1997; Falkinger, 1995; Kinsey, Grasmick, & Smith, 1991; Scholz & Lubell, 1998; Wenzel, 2002). In general, according to this line of research, there is an alternative approach, called the attitudinal or intrinsic approach to tax compliance and tax evasion. Tax payers are viewed as ‘voluntarily’ complying with their tax duties as a function of intrinsic motivation, i.e., internalized norms of conduct (e.g., reciprocity - taxpayers behave fairly toward the system if they are treated fairly). ‘Voluntarily’ in this context means that the citizens pay taxes without surveillance or threat of sanctions. Following this approach, this chapter presents a multi-faceted fairness framework of taxation. Perceived fairness of the tax system is assumed to play a paramount role in producing genuine acceptance of tax obligations and regulations. The theoretical framework proposed here derives from the attitudinal approach and is an attempt to explain and predict tax attitudes and voluntary tax compliance.

### A fairness approach to taxation

In this chapter, the concepts of justice and fairness are used interchangeably. The fulfillment of tax obligations is here proposed to be an issue of fairness. Several studies confirm this statement (Falkinger, 1995; Kim, 2002; Murphy, 2003; Wenzel, 2002, 2006). Taxation (or willingness to pay taxes) is a fairness issue because it involves an exchange process; the individual taxpayer wants to know what s/he receives in return for the taxes that s/he pays. However, fairness is a multi-faceted construct and does not only concern how societal resources are distributed among citizens but also how the expenditure programme is constructed (i.e., how each unit of paid taxes is distributed among different societal spheres including education, defence, police, Medicare, etc).

The distribution aspect of fairness (i.e., distributive justice - how societal resources like subsidies, education, Medicare, child care, etc are allocated among citizens), is presumably the dominant and most salient fairness aspect relevant to willingness to pay taxes. In this sense, tax payment can be conceived of as a social dilemma, defined as a situation of interdependence where the individual must decide whether to maximize own personal interest or the interest of society. The dilemma is embedded in the fact that, no matter how other fellow citizens choose, it is more desirable for each individual to choose an alternative that benefits her selfish interest (i.e., evading taxes) rather than society’s interest (i.e., fulfilling tax duties and contributing to the collective welfare). But if all citizens would behave selfishly, all will receive poorer outcomes than if everyone chooses to meet tax obligations (i.e., cooperation or acting for the
public good) (Dawes, 1980; Elffers, 2000; see also Kazemi & Eek, 2008 for different accounts of why fairness matters in social dilemmas).

The fairness framework of taxation proposed in this chapter posits that fairness and willingness to pay taxes is also related to taxpayer-authorities interaction. More specifically, fairness also concerns how tax decisions are made (procedural justice), how citizens are treated by the tax authorities/employees (interpersonal justice), and how tax decisions and the legislation are explained and justified to the citizens (informational justice).

Defining and relating the concepts of the framework
Based on recent theorizing and meta-analyses (Cohen-Charash & Spector, 2001 and Colquitt et al., 2001; see also Kazemi & Törnblom, 2008), the present fairness framework distinguishes between four types of justice aspects: distributive, procedural, interpersonal, and informational. Distributive justice refers to the perceived fairness of the final shape or outcome of a resource allocation event (e.g., a tax expenditure programme). Three allocation principles are usually discussed in the justice literature: equity, equality, and need. The equity principle (requiring proportionality between an individual’s inputs and outcomes should be proportional to his/her inputs) may be formulated in terms of (a) amount of expended effort, (b) ability, innate or achieved, and (c) productivity (i.e., the actual results accomplished). The principle of equality (for which inputs are irrelevant) may be conceived in terms of equality of (a) treatment, according to which everyone receives the same outcomes, (b) opportunity, according to which everyone has the same chance to receive, and (c) results, in which case everyone ends up with the same amount of resource in the long run, even though they may receive unequal shares at any particular occasion. The need principle can be conceptualized in terms of biological (means of survival like food, shelter, and air), basic (items and conditions which are considered a minimal standard of living in a given society), and functional (means and tools required to fulfill one’s function in society) (Törnblom & Jonsson, 1985). These conceptual distinctions help us in describing the tax system from a distributive fairness perspective. For instance, the Swedish tax system (like many other tax systems in other countries) demand income tax using an equity norm. Thus, the more you earn the higher the tax rates. But the governmental Medicare system that is financed by tax revenue distributes its resources equally or according to recipient needs. This may be perceived as an instance of mismatch between the norm of input and the norms of output and thus be a potential source to perceived unfairness.

While distributive justice refers to the perceived fairness of the final outcomes received, procedural justice pertains to how the decisions about the distributions are made. Procedural fairness may be assessed using multiple criteria: consistency (equal treatment across persons, e.g., the same rules of taxation applies to all citizens including governmental officials); bias
suppression (refraining from self-interest and preconceptions, e.g., the tax agency cannot make prejudiced tax decisions); accuracy (using all correct information available about the financial situation of individual taxpayers); correctability (opportunity for second opinions and modification of decisions, i.e., taxpayers should be provided an opportunity to get a second opinion if they believe that their tax return is faulty); and ethicality (compatibility of the tax legislation with universal moral standards and values) (Leventhal, 1980).

Whereas procedural justice targets the formal structural aspects of how tax decisions are made, interpersonal justice targets the informal communicative parts of justice and has to do with whether authorities are perceived as considerate and respectful, treating the citizens with dignity (which by the way do not incur any monetary costs, yet is quite effective in eliciting positive attitudes toward the tax agency and in the end for citizens’ willingness to pay taxes).

Informational justice focuses on the explanations for why certain procedures were followed or why certain decisions were made. Thus, it refers to the extent to which tax decisions are justified and to the explanations that the tax agency employees offer when citizens ask about the legislation, their tax shares, and penalties (cf. Greenberg, 1993).

Other concepts in this theoretical framework are uncertainty, trust, and legitimacy. When people are at the point of deciding whether to pay or evade taxes (i.e., to contribute or not to a public good), their decisions are affected by different types of uncertainties. Two notions of social and environmental uncertainty are used in the study of social dilemmas. Social uncertainty refers to the lack of knowledge about fellow citizens in society concerning their decisions whether to increase collective welfare or promote own selfish interests, while environmental uncertainty pertains to the vagueness of the defining features of a specific social dilemma. As people expect the tax expenditure programme to be fair (Alm, Jackson, & McKee, 1992, 1993), they are expected to seek information regarding the allocation of the revenue (i.e., gain environmental certainty). Environmental uncertainty is specifically related to: (i) what principles are used in distributing the revenue, (ii) what kind of resources (e.g., money in terms of different subsidies, services in terms of judicial system or Medicare, information in terms of educational efforts) that taxpayers receive in return for their paid taxes from the government, and (iii) the extent to which the citizens perceive that they can exert control over their share of the revenue. Thus, environmental uncertainty in the context of taxation pertains to the features of the exchange between the taxpayers and the government/authorities.

The citizens also experience social uncertainty when deciding to pay or evade taxes. The fairness framework also relates the notion of uncertainty to the notion of trust in that it proposes a bipartite social uncertainty notion consisting of horizontal and vertical trust (cf. Rothstein, 2005). Horizontal trust pertains to
trust in fellow citizens’ willingness to pay taxes as well as their actual behavior of paying taxes, whereas *vertical trust* pertains to trust in how authorities fulfill their own tax duties and how they use tax revenue for different purposes.

Trust is assumed to affect perceptions of fairness. When taxpayers know how the tax revenue is distributed, environmental uncertainty is reduced, and perceptions of fairness may increase. However, this is not a sufficient condition for increased perceptions of fairness. Citizens also have to perceive the distributions as fair, and that is likely to happen when they are informed about how decisions are made and when they can exert control over the allocations/tax expenditure programmes, or at least have a say in the debate.

The fairness framework posits moreover that perceptions of fairness lead to heightened *system legitimacy*. A legitimate agency/government is one that is considered to be a positive agency/government, acts in accordance with some agreed upon moral standards, and endorses the principle that people not acting in accordance with current duties deserve penalties even though these are authorities (Hegtvedt & Johnson, 2000; Walker & Zelditch, 1993). Social regulation is not possible unless the regulatory/system is perceived as legitimate. A crucial source of legitimacy is fairness. Thus, one implication for the tax administration is to identify and remove sources of unfairness in the system/regulations. Otherwise the system has a problem of legitimacy. The notion of legitimacy is also related to the notion of vertical trust in that to the extent that taxpayers are uncertain about whether, for instance, famous people and governmental officials meet their tax liabilities and how they are treated by the tax agency, affects the perceived legitimacy of the system and its perceived fairness.
Figure 1. Overview of the Fairness Framework of Taxation

Figure 1 shows how different sources of (un)fairness affect perceived (un)fairness of the system. Under each type of justice, some concrete instances are discerned. Other instances or concrete sources of (un)fairness under each justice type could easily be identified.

Personal utilization refers to the extent to which the individual taxpayer employs different services financed via taxes, such as Medicare, education, judicial system, etc. Personal contribution refers to the extent to which the individual pays taxes, i.e., the approximate sum of direct and indirect taxes paid in a year. The fairness framework predicts that the more they contribute to the public good (i.e., the collective welfare) by paying taxes, the more they expect in return from the system. The tax expenditure programme refers to how the revenue is distributed. There are different types of tax (e.g., work taxes, taxes on capital, taxes on goods and services, and corporation taxes). Furthermore, there are three general purposes for paying taxes: fiscal (providing revenue to foot the bill for the collective benefits offered by the state or community), allocative (taxation aiming at levelling wage and wealth differences among categories of citizens), and stabilizing (taxation to accomplish full employment and a stable cost of living). In addition, ‘targeted taxation’ is applied to raise prices in order to discourage consumption of unhealthy, environmentally toxic and other
dysfunctional items. An assumption is that the perceived fairness of different taxes varies with their different purposes.

As featured in Figure 1, procedural fairness has to do with the perceived quality of public services which refers to how citizens are treated by governmental agencies that are financed by tax money. Moreover, privatization of some public core services is assumed to affect citizens’ willingness to pay taxes for financing public governmental alternatives. This has to do with to what extent taxpayers/citizens are allowed to choose private alternatives to services that traditionally have been provided by the government from the tax revenue. This means that citizens could be given voice (a procedural justice criterion) about whether they want to have Medicare privately or via paying taxes. For instance, citizens who wish to take out private health care insurances could pay reduced tax rates. In this way, people are free to choose providers of societal services.

Interpersonal justice stresses taxpayers’ fundamental right to be treated with respect, even though they have deliberately or unintentionally not met their tax duties. Informational justice is related to the transparency of the system and as such it is a core democratic value. Wenzel (2006) maintains that informational justice is important as it: (i) serves as a mean for showing that the authority has made an accurate and objective decision about final tax liabilities (i.e., procedural justice), (ii) enhances the fairness of the tax system with regard to the tax expenditure programme and distribution of the tax revenue (i.e., distributive justice), and (iii) gives the impression that the authorities care about and respect the citizens as they spend time on providing citizens with sufficient information and explanations (i.e., informational and interpersonal justice, respectively).

In sum, Figure 1 identifies three primary predictors of tax attitudes and behavior (i.e., the primary dependent variable). The predictors are uncertainty (social and environmental), fairness (distributive, procedural, interpersonal, and informational), and legitimacy. The predictors are also proposed to affect each other in different ways. The fairness framework allows for a systematic examination of the independent, interactive, and joint effects of different types of justice on legitimacy and tax attitudes and behavior.

Assumptions and predictions of the taxation fairness framework
A first assumption of the fairness framework is the fairness motive assumption which states that taxation is in nature an issue of fairness, and that inherent in the human nature there is a quest for justice. However, this assumption does not imply that fairness is the only motive guiding behavior. There are other motives like self-interest.

Fairness is further assumed to be an important predictor of tax attitudes and behavior as it fulfills four different types of identity maintenance needs: (i) instrumental (material identity needs, Thibaut & Walker, 1975; Skitka, 2003),
(ii) relational (social identity needs, Lind & Tyler, 1988; Tyler & Lind, 1992),
(iii) protective (certainty or predictability needs, Lind & Van den Bos, 2002),
and (iv) moral (personal identity needs, Skitka & Mullen, 2002). How fairness
satisfies each need is briefly described in the following.

(i) Fairness matters in the context of taxpaying as people care about the
outcomes that they receive and the fact that individual and societal resources are
scarce. Citizens pay taxes (i.e., money) and in return expect to receive different
types of services, subsidiaries, and facilities. This account focuses on the
expected and actual outcomes that taxpayers receive, thus the name instrumental.

(ii) The relational account holds that taxpayers not only have material
identity needs, but also need to be treated in a way that makes them feel as
respected citizens with certain rights. The relational model further maintains that
judgments of tax fairness (e.g., decisions about tax duties and penalties for not
meeting tax liabilities) are relational in nature and that the issues of neutrality,
trustworthiness, and status recognition are central to understanding fairness.

(iii) According to the protective account, fairness is important as it is
conducive to reducing the uncertainty that citizens may experience in their
relationships with authorities and fellow citizens. Fairness is a norm, and
therefore functions to coordinate behavior. If I believe in fair play, and expect
others to do so, a kind of consensus effect results. Uncertainty management
theory (Lind & Van den Bos, 2002) posits that activation of fairness processes is
an indication of fairness judgments being utilized to settle some important social
or psychological issue. Lind and Van den Bos suggest that fairness serves
people in managing uncertain situations (e.g., are everyone else going to pay
his/her taxes?) by giving them “confidence that they will ultimately receive
good outcomes and because it makes the possibility of loss less anxiety-
provoking” (p. 195).

(iv) The moral account stresses the importance of personal moral stands and
positions as important determinants of how people reason about fairness.
According to the moral account if I strongly believe that, for instance, a
mismatch between norms of input (i.e., equitable tax payment) and norms of
output (i.e., equal or need based access to and utilization of Medicare and
educational services) is immoral and wrong, then I tend to perceive the system
as unfair regardless of procedural, interpersonal, and informational justice
considerations.

How fairness satisfies these needs makes sense when fairness is conceived
of as a multi-faceted issue/concept. The multidimensional fairness assumption
states that justice conceptions are pluralistic and multi-faceted in that they
cannot all be accounted for by reference to a single basic aspect or principle.
Even if people tend to think of fairness as a unity, this unity has different parts.
Differentiating between these parts is crucial as different parts sometimes have
differential effects on outcomes or modify each others effects on some outcome. A closely related assumption is the fairness interdependence assumption which maintains that different aspects/types/dimensions of fairness affect each other, i.e., perceptions of another aspect (for propositions regarding the effects of distributive and procedural justice on each other see Törnblo & Vermunt, 1999, see also Wenzel, 2006).

A final important assumption of the fairness framework presented here is the fairness availability assumption which highlights the heuristic function of fairness, and has to do with how salient aspects of justice influence the total justice judgment. This assumption derives from the fairness heuristic theory (Lind, 2001) which posits that people want to trust authorities, and to find out whether or not they can trust authorities they look at how fair they are. The theory further maintains that once fairness judgments are established, they serve as heuristics for the interpretation of subsequent events. In support of this reasoning it has been shown that when facing ambiguous information about final outcomes, people tend to give more weight to the procedural aspects of the allocation (Van den Bos, Lind, Vermunt, & Wilke, 1997). In the context of taxation, the interpersonal and informational fairness aspects are salient. This has to do with the fact that taxpayers remember how they have been treated by the authorities and specifically whether the authorities have treated them with respect and explained how decisions about penalties, tax duties, etc. are made. Own outcomes (how much I pay in taxes and how much I receive in return, i.e., distributive justice) are also known to taxpayers. However, taxpayers do not know about fellow taxpayers’ outcomes. The fairness framework suggests that to the extent that informational justice is in place, outcomes tend to be perceived as more fair (or less unfair), even though the outcomes are unfavorable. Moreover, the correctability and ethicality aspects of procedural justice are frequently known to taxpayers, whereas bias suppression and accuracy are not known to the same extent. In sum, the fairness availability assumption refers to the observation that when information about one justice aspect is lacking, people tend to rely on information available on other components. They use this information to assess the fairness of the whole event – an 'overall fairness evaluation' (cf. Törnblo & Vermunt, 1999).

Concluding remarks
Scattered research evidence seem to support the contention that perceived fairness of the tax system plays an important role for taxpayers’ attitudes toward the system and subsequently for their willingness to pay taxes and support the system (e.g., Falkinger, 1995; Kim, 2002; Kinsey, Grasmick, & Smith, 1991; Murphy, 2003; Wenzel, 2002, 2006). This earlier work has, however, not provided an integrated framework of: (i) the multi-faceted nature of fairness in the context of taxation, (ii) how fairness influences tax attitudes and behavior,
and (iii) why fairness matters in the context of taxation. The taxation fairness framework proposed here gives an overview of the multi-faceted nature of fairness and how it is related to trust and legitimacy which are two other key concepts in this context. As the fairness framework of taxation presented in this chapter has not been tested in its entirety, the predictions remain to be empirically tested by future research.

Theory and theoretical developments in social sciences should at its best in some sense serve the society and its institutions. The framework presented here attempts from a more general point of view to address the issue of social regulation, that is, how do we get citizens to follow the legislation. Thus, this framework has implications for other issues than taxes as it assumes fairness to be a fundamental motive in the social life.

If fairness is assumed to be important in the context of taxation (as it affects willingness to pay or evade taxes directly or via trust in authorities and system legitimacy, see Figure 1), what can decision and policy makers do to contribute to the perceived fairness of the tax system? There are two answers to this question. Decision makers should: 1) design a fair tax system (distributive and procedural justice), and 2) show that the tax system is fair (informational justice). Of course, authorities could force citizens to follow the regulations, but illegitimate forcing results in serious detrimental societal consequences, such as tax evasion resulting in an impoverishment of the collective welfare, less societal citizenship behavior, and emergence of hostility and conflicts between different social groups.

References


Author Note
Ali Kazemi, PhD, assistant professor of social psychology, School of Technology and Society, Social Psychology Unit, University of Skövde, P.O. Box 408, SE-541 28 Skövde, Sweden. E-mail: ali.kazemi@his.se
Thanks are due to Kjell Törnblom for valuable comments.
THE RELATIVE IMPORTANCE OF OUTCOME AND PROCEDURE FOR TOTAL JUSTICE JUDGMENTS

Kjell Törnblom & Ali Kazemi

Abstract
Contemporary justice theorists assume that a meaningful assessment of fairness in interpersonal encounters requires assessments of the outcome (the end result) as well as the procedure (the means) by which the outcome was accomplished. This study investigated the relative impact of four variables on the subjective importance of the outcome and procedure for total fairness evaluations, using a 2x2x2 factorial survey design: Type of offense (physical abuse vs. theft) x Severity of offense (moderate vs. serious) x Social relationship (particularistic vs. universalistic) x Status of the perpetrator relative to the victim (equal vs. superior). Results suggest that the outcome is considered more important than the procedure for fairness judgments of both offenses, regardless of their severity, relative status of the offender, and the social relationship within which the offense was committed. Furthermore, both outcome and procedure were viewed as more important when assessing the fairness of physical abuse as compared to theft. For the physical abuse offense, this was more likely to be the case when the status of the offender was superior to the status of the victim.

Keywords: Fairness, justice, outcome, procedure, physical abuse, theft

Introduction
The major focus of the present study concerns the perceived importance of the outcome distribution and the procedure by which outcomes were accomplished for evaluations of the overall fairness of an allocation event. The importance of these two components are assessed in various situations, where moderate and serious person and property offenses are committed within particularistic and universalistic relationships by a person of similar as well as a superior status (age) relative to the victim. In concrete terms, we investigated how important respondents think it is to consider (a) the ways in which physical abuse and theft are committed and (b) the nature and extent of the damage done, when they form an opinion about the (un)fairness of what happened. Contemporary justice theorists increasingly assume that a meaningful 'total fairness assessment' requires assessments of the distribution ('distribution', 'outcome', and 'ends' will be used interchangeably to refer to the endstate of a resource allocation event), the procedure (the means) by which the distribution was accomplished, and the manner in which the procedure was enacted interpersonally. In the ideal situation all three justice components are perceived
to be fair, i.e., when distributive, procedural, and interactional fairness prevail (See Törnblom & Vermunt, 1999, for more details and a model for integrating the two first mentioned allocation components).

Törnblom and Vermunt (1999) postulated that the perceived total fairness of a situation is a function of fairness assessments of both the distribution and the procedure, and when both distribution and procedure are salient, their fairness assessments are likely to be interdependent. Further, since salience is relative, a full understanding of the meaning of a given total fairness evaluation measure requires knowledge as to whether the distribution and the procedure were equally or unequally important or weighty for the assessment of total fairness, and under what conditions either possibility might be likely. Lind and Tyler (1988) proposed that procedural fairness assessments are at least as important as distributive fairness assessments for overall justice evaluations in legal and organizational settings. Almost a decade earlier, Leventhal (1980) suggested that procedural justice is most important when an organization is created, when people are dissatisfied with the distribution, and when distributive justice is violated. And distributive justice becomes most important when procedural justice is violated.

The purpose of the present study is to examine some factors that might affect the relative importance of the 'means' (i.e., the procedure) and 'ends' (i.e., the outcome distribution), respectively, for total fairness evaluations of resource allocation events. Where might we turn in our search for factors that might affect importance evaluations of the distribution and the procedure?

Outcome distributions and allocation procedures may be evaluated in terms of various criteria, such as preference, acceptability, expediency, appropriateness, importance, desirability, efficacy, satisfaction, and fairness. Various factors determine what weights and values are assigned to each of these different criteria, as well as their relevance. For instance, the social psychological literature on distributive justice and fairness documents an abundance of theoretical and empirical attempts to map the factors that affect fairness evaluations and choices of principles that people think should govern distributions. These factors have been grouped into six general categories (Törnblom, 1992) on the basis of whether they characterize the actor (e.g., age and sex), the contribution (e.g., intentionality and valence), the social relationship (e.g., intimacy and competitiveness), the sociocultural and historical context (e.g., egalitarian and achievement orientation), the outcome (e.g., scarcity and particularism), or the outcome allocation (e.g., sign and way of accomplishment). Surprisingly, it appears that corresponding research on factors that affect fairness evaluations and choices of procedural as well as interactional, principles is largely missing.

It seems reasonable to assume that some of the factors which are known to influence justice evaluations of a distribution are also likely to affect such
evaluations of a procedure. Moreover, we assume that ratings of distributions and procedures in terms of their importance for total fairness assessments are also impacted by some of those factors. For instance, the social relationship between provider and recipient has been shown to be a significant determinant in a variety of studies in different areas (see Törnblom, 1992, for a review). Several studies on justice evaluations indicate that resource allocations according to the equity principle is considered most fair in competitive, distant and formal relationships, while the equality and need rules are endorsed for allocations within cooperative, close and caring contexts (e.g., Lerner, 1974; Törnblom and Jonsson, 1987). However, the impact of the social relationship between provider and recipient is likely to be modified by a number of other factors that affect justice ratings, such as the nature of the allocated resource (e.g., love vs. money), the valence of the resource (positive vs. negative), the magnitude of the valence, and the status (e.g., age) of the provider relative to the recipient.

Previous research has focused on the conditions under which distributive and procedural fairness interact, when one is a moderator of the other (i.e., how important or weighty procedural fairness evaluations are in the determination of distributive fairness, and vice versa; see Brockner and Wiesenfeld, 1996, for an extensive review). Justice researchers have also examined whether the distribution or the procedure is the most important factor in the determination of various outcomes other than justice and fairness (such as decisional satisfaction, organizational legitimacy and commitment, etc.). The 'total fairness model' proposed by Törnblom and Vermunt (1999) takes a different direction, in that predictions for fairness assessments of the whole situation are made. Based on the assumption that people generally perceive the allocation event in terms of a Gestalt, a wholeness, it follows that the 'figure' for fairness judgments is the entire entity, while its distributive and procedural aspects constitute the 'ground' (rather than the other way around), on the basis of their relative importance for the total picture. Thus, the distributive and procedural components (whether or not they interact) are not singularly taken to represent the fairness of a situation; neither are they necessarily assessed on the basis of how they affect each other. They are, however, evaluated in terms of their respective relative level of importance in the formation of a judgment of a situation's overall fairness.

Consistent with the approach suggested by the total fairness conception (Törnblom & Vermunt, 1999), the purpose of the present study is to examine the impact of some factors that might affect the relative importance of the procedure and the distribution in their capacity as components of total fairness evaluations of a resource allocation event. More specifically, we focus on the subjective relative weights of the distributive and procedural aspects of the total fairness assessments of different situations. Thus, as mentioned above, the major focus of the present study concerns importance ratings of the outcome distribution and
the procedure for overall fairness evaluations, when moderate and severe person and property offenses are committed within particularistic and universalistic relationships by a person of similar or superior status (age) relative to the victim.

**Importance of outcome and procedure for justice evaluations of moderate and severe offenses**

On the basis of what type of justice configuration between outcome and procedure do people react as a response to crimes or other offenses committed against oneself or others? As previously stated, we assume that assessments of justice include distributive as well as procedural components. However, would not these components likely be weighted differently and vary in their salience with different types of offenses? Even though this conjecture is intuitively reasonable, theoretically interesting, and of considerable practical importance, we have not been able to find any directly relevant research.

One way of thinking about these issues is that when a serious offense has been committed (e.g., physical abuse with visible injuries, destruction of valuable property, or theft of large financial resources), the results are immediate, salient and deeply felt. Thus, the outcome of a serious offense may very well be more figural (important) for assessment of the situation than the way in which the outcome was generated (i.e., the procedure). On the other hand, when the offense is mild or trivial, the victim’s primary attention is likely to be diverted away from the nature or extent of the offense, per se, due to a more urgent focus on the reason or intention behind the negative act – ”although what happened to me did no damage to speak of and is practically negligible, it still bothers and puzzles me that the offender wanted to hurt me” . Thus, the procedure might be more important than the outcome for a justice judgement of an offense below a given threshold. A negative procedure may convey negative intent which, in turn, is likely to be interpreted as an act directly aimed at and disrespecting the person. As such, she will feel that her moral rights to be treated respectfully are violated. As a consequence, she is likely to experience procedural and/or interactional injustice. Apparently, procedure (here conceived of as behavior that leads to an outcome) may (like most behavior) be experienced and interpreted in terms of intention, in which case analytical rigor requires clear distinctions. Based on above arguments, we make the following two predictions which will be tested for two different kinds of offense - physical abuse and theft:

**Proposition 1:** The outcome (O) is more important than the procedure (P) when evaluating the fairness of a serious offense (so).

\[ O_{so} > P_{so} \]

**Proposition 2:** The procedure (P) is more important than the outcome (O) when evaluating the fairness of a moderate offense (mo).
Importance of outcome and procedure for justice evaluations of person and property offenses, respectively

When a person offense (like physical abuse) is committed, both the harmdoer and the victim are usually present and aware during the act, completely enfolded by the malignant behavior as it evolves. The impact of such an offense is likely to be more direct and traumatic as compared to a property offense (e.g., theft), of which the victim is usually unaware during its commitment. Using Sarbin and Allen’s (1968) terminology in his characterization of role playing intensity, the participant’s organismic involvement is likely to be considerably higher in the context of physical abuse as compared to theft. Thus, it seems reasonable to suggest that both the procedure (the act by which the offense is committed) and the type and severity of the offense (i.e., the result or outcome of the act) are more focal and important for judgements of those offenses. Also, "In most instances of property harm, compared to physical or psychological harm, the harm can be rectified to a greater extent (Karniol and Miller, 1981). Therefore:

Proposition 3: The outcome (O) is more important when evaluating the fairness of physical abuse (pa) than when evaluating the fairness of theft (t).

\[ O_{pa} > O_t \]

Proposition 4: The procedure (P) is more important when evaluating the fairness of physical abuse (pa) than when evaluating the fairness of theft (t).

\[ P_{pa} > P_t \]

Impact of the offender’s status relative to the victim on the relative importance of outcome and procedure

Hogan and Emler (1981) noted that “...the more status one has, the more careful one must be not to arouse the resentment of the less unfortunate…” (p. 142). A person who capitalizes on her superior status to harm or take advantage of a person of inferior status is usually severely frowned upon, receive intense social disapproval, and is harshly judged. Identical offenses committed by an equal-status offender are likely to be viewed as less serious. Therefore, the acts of the superior-status offender are more carefully scrutinized. Both the outcome and the procedure that generated the outcome would likely assume greater importance for (in)justice assessments when the status of the offender is superior to the victim rather than equal.

Proposition 5: The outcome (O) is more important when evaluating the fairness of an offense committed by a person of superior status (ss) to the victim as compared to equal status (es).

\[ O_{ss} > O_{es} \]
Proposition 6: The procedure (P) is more important when evaluating the fairness of an offense committed by a person of superior status (ss) to the victim as compared to equal status (es).

\[ P_{ss} > P_{es} \]

Propositions 3 and 5 generate the following prediction:

Proposition 7: The outcome (O) is more important when evaluating the fairness of physical abuse (pa) than when evaluating the fairness of theft (t). This is more likely when evaluating the fairness of an offense committed by a person of superior status (SS) to the victim as compared to equal status (ES).

\[ (SS: O_{pa} > O_t) > (ES: O_{pa} > O_t) \]

Propositions 4 & 6 generate the following prediction:

Proposition 8: The procedure (P) is more important when evaluating the fairness of physical abuse (pa) than when evaluating the fairness of theft (t). This is more likely when the offense is committed by a person of superior status (SS) to the victim as compared to equal status (ES).

\[ (SS: P_{pa} > P_t) > (ES: P_{pa} > P_t) \]

Impact of the social relationship between the offender and the victim on the relative importance of outcome and procedure

One of the factors that has been shown to be a significant determinant of attitudes and behavior in a variety of studies in different areas is the social relationship between provider and recipient (e.g., Austin, 1980; Deutsch, 1985; Greenberg, 1978; Hassebrauck, 1984; Kidder, Fagan and Cohn, 1981; Lamm and Kayser, 1978; Lamm, Kayser and Schanz, 1983; Lerner, 1974, 1977; Mills and Clark, 1982). For instance, studies on justice evaluations indicate that resource allocation according to the equity principle is considered most fair in competitive and distant relationships, while the equality and need rules are endorsed for allocations within cooperative, close and caring contexts (e.g., Deutsch, 1975; Törnblom and Jonsson, 1987).

Another question, apart from which justice principle is viewed as most fair, is how relevant and important fairness considerations are in various situations and relationships. And when relevant, how important are the outcome and procedural justice components relative to each other? Intuitively, and as far as importance ratings in different types of social relationships are concerned, the procedure (i.e., way in which a benefit or harm is provided) will likely matter more than the outcome (i.e., the particular amount of benefit or harm) when assessing (first, intention and then) fairness within a particularistic (close,
intimate) social relationship. This is probably particularly true, when certain kinds of social resources (e.g., love, comfort, assistance, services, regard, and respect) are provided or withdrawn. For example, at any single occasion, the way in which I convey affectionate sentiments to my lover is likely to carry more weight than the outcome, the amount, of affection I give (although this may not be true in the long run). For other kinds of resources (e.g., money, goods, and information) the amount is probably more important than the way in which they were provided. Thus, two interesting propositions (not to be tested here) are that (a) the procedure is more important than the outcome when assessing the fairness of particularistic resource allocations, and (b) the outcome is more important than the procedure when assessing the fairness of universalistic resource allocations.

Further, the procedure might be even more weighty (especially within particularistic/close relationships) when negatively valent social resources (e.g., disservices, like person or property offenses) come our way. We are likely to be especially eager to find out how and why the unfavorable outcome did happen to us. (However, we care less about the hows and whys, i.e., the procedures, when luck smiles in our face.)

The connection between concerns about procedural justice aspects and one's status within the group has been discussed by Lind and Tyler (1988) and Tyler and Lind (1992). In the contexts of their 'group value model of procedural justice' and their 'relational model of authority', the importance of procedural justice is associated with concerns about the group member's social identity, inclusion, and group belongingness. When the group member's identification with the group is high, respectful treatment via fair procedures by others in the group is especially important and expected. As group identification and inclusion is likely to be particularly characteristic for intimate, close, or particularistic relationships, procedural justice concerns may be more important within such relationships as compared to more distant, universalistic ones. Bad treatment may indeed be harder to swallow if imposed within a close as compared to a distant relationship. Being the victim of acts of physical abuse and theft (i.e., negative procedures) committed by a family member or trusted friend is more tragic and represents a graver betrayal of trust, than when the offender is unknown or more distant. Thus:

Proposition 9: The procedure (P) is more important when evaluating the fairness of an offense committed within a particularistic (part) as compared to a universalistic (univ) relationship.

\[
P_{\text{part}} > P_{\text{univ}}
\]

We can find no obvious reasons as to why the outcome would be more or less important for either type of relationship. Therefore, we suggest that the outcome
will be equally important for fairness assessments within particularistic and universalistic relationships.

**Method**

**Design**
The study was designed as a $2 \times 2 \times 2 \times 2$ factorial survey: Nature of the offense (physical abuse vs. theft) $\times$ Severity of offense (serious vs. moderate) $\times$ Relative status/age position of the offender vis-à-vis the victim (superior vs. equal) $\times$ Social relationship (particularistic vs. universalistic). This may be expressed by a mapping sentence on the basis of which experimental conditions are created (see Figure 1).

<table>
<thead>
<tr>
<th>A. Nature of offense:</th>
<th>B. Severity of offense:</th>
<th>C. Social relationship:</th>
<th>D. Status of offender relative to victim:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1 physical abuse</td>
<td>b1 serious</td>
<td>c1 particularistic</td>
<td>d1 superior</td>
</tr>
<tr>
<td>a2 theft (money)</td>
<td>b2 moderate</td>
<td>c2 universalistic</td>
<td>d2 equal</td>
</tr>
</tbody>
</table>

on importance of (rated along 7-point scales)

**D. Justice aspect:**
- d1 outcome distribution
- d2 procedure

*when making total justice judgements.*

Figure 1. Overview of the study.

**Respondents**
146 male and 170 female undergraduate students at a Swedish university volunteered to participate in the study during ordinary class meetings. Their age ranged from 19 to 63 years ($M = 25$ years). Each respondent was randomly assigned to answer one of 16 different versions of the questionnaire described below.

**Instrument**
The design of this study generated 16 conditions each of which was represented by a vignette (see below), i.e., a scenario describing a situation in which one of two types of offense took place. Eight of the vignettes (1 - 8) described incidents of physical abuse (a1) and the remaining eight (9 - 16) theft of money (a2). The
moderate (a1b2) and serious (a1b1) physical abuse offenses consisted of a slap in the face (vignettes 5 – 8) and incurring face wounds requiring stitches as well as kicks in the stomach (vignettes 1 – 4), respectively. The moderate (a2b2) and serious (a2b1) thefts consisted of SKR 600 = $70 (vignettes 13 – 16) and SKR 1,200,000 = $140,000 (vignettes 9 – 12), respectively. The social relationship between the perpetrator and the victim was either particularistic (c1) (intimate, closely related, i.e., a wife/daughter victim, the offense taking place at the victim's home) or universalistic (c2) (they did not know each other, were not related, i.e., a colleague/child victim, the offense occurring at a day-care centre). Relative status position of the offender was operationalized as the age of the offender vis-à-vis the victim, i.e., superior (d1) (adult offender and child victim) or equal (d2) (adult offender and adult victim). The 16 experimental conditions and the numbers of the corresponding vignettes:

| a1b1c1d1 | 1 | a2b1c1d1 | 9 |
| a1b1c1d2 | 2 | a2b1c1d2 | 10 |
| **a1b1c2d1** | 3 | a2b1c2d1 | 11 |
| a1b1c2d2 | 4 | a2b1c2d2 | 12 |
| a1b2c1d1 | 5 | a2b2c1d1 | 13 |
| **a1b2c1d2** | 6 | a2b2c1d2 | 14 |
| a1b2c2d1 | 7 | **a2b2c2d1** | 15 |
| a1b2c2d2 | 8 | a2b2c2d2 | 16 |

A sample of four (in bold) of the experimental conditions - 3, 6, 10, and 15 follows:

**Vignette 3: (a1b1c2d1 - Physical abuse, Serious offense, Universlistic relationship, Superior status)**

A male staff member has just been hired at six year old Ida’s day-care center. One day when the children are sent outside to play, Ida refuses to join them. The new staff member tries to persuade Ida to put her coat on, but Ida screams and tries to get away. The staff member loses his temper, grabs Ida and shakes her violently. This only makes Ida scream even louder, whereupon he hits Ida several times in the face and kicks her in the stomach. Ida is brought to the hospital in an ambulance where the doctor puts five stitches in her face.

**Vignette 6: (a1b2c1d2 - Physical abuse, Moderate offense, Particularistic relationship, Equal status)**

Lena and her husband have just returned home from their jobs when they start arguing about whose turn it is to make dinner. After a while Lena’s husband loses his temper and slaps her face once.
Vignette 10: (a2b1c1d2 - Theft, Serious offense, Particularistic relationship, Equal status)
Lena had just celebrated her birthday at which time she was lucky to win $140,000 at a lottery. Lena’s husband decides that he needs that money more than Lena, and he manages in an unknown way to transfer all the money to his own bank account.

Vignette 15: (a2b2c2d1 - Theft, Moderate offense, Universalistic relationship, Superior status)
Six-year old Ida spent the morning with her granny who accompanies Ida to her day-care center in the afternoon. Granny tells the new male staff member that she has put some money in Ida’s purse for her parents to deposit in Ida’s bank account. When Ida is outside playing with the other children the staff member opens Ida’s purse and finds $140. Because he is short of money, he decides to steal $70 from Ida. He knows that he will not be able to return then money to Ida, and he therefore hopes that Ida will not notice that they are missing.

The vignettes were constructed with the following considerations in mind, aimed at increasing the likelihood that a sense of injustice will be aroused (Deutsch, 1985; Karniol and Miller, 1981):

- Acts are described as intentional; such acts are more likely to arouse a sense of injustice as compared to unintended acts.
- The perpetrator of the offense has no moral or legitimate authority to cause harm. Thus, the procedure/behavior is unethical.
- Similar harm is not (typically) inflicted on other people who are considered similar to the victim.
- For illegitimate property access (theft), there is no intention to return the stolen resources.
- For physical offense, there is no stated intention to ask for forgiveness or provide compensation for the harm done.

A number of questions that followed the vignettes generated data for the dependent variables reported here: (a) relative importance ratings of the outcome and (b) relative importance ratings of the procedure for evaluations of total fairness. Ratings on these variables were done from the viewpoint of the respondent.

**Results**
The means and standard deviations for the perceived importance of outcome and procedure in fairness judgements for each of the 16 experimental conditions of the present study are shown below in Table 1.
<table>
<thead>
<tr>
<th>Experimental conditions</th>
<th>Importance</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Outcome</td>
<td>Procedure</td>
<td>Outcome</td>
<td>Procedure</td>
<td>Outcome</td>
<td>Procedure</td>
</tr>
<tr>
<td>a1b1c1d1</td>
<td>6.4</td>
<td>1.3</td>
<td>5.1</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a1b1c1d2</td>
<td>5.4</td>
<td>2.4</td>
<td>3.1</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a1b1c2d1</td>
<td>6.3</td>
<td>1.2</td>
<td>4.8</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a1b1c2d2</td>
<td>6.5</td>
<td>0.7</td>
<td>5.1</td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a1b2c1d1</td>
<td>6.4</td>
<td>1.0</td>
<td>4.5</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a1b2c1d2</td>
<td>5.8</td>
<td>2.2</td>
<td>3.4</td>
<td>2.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a1b2c2d1</td>
<td>6.3</td>
<td>1.6</td>
<td>4.0</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a1b2c2d2</td>
<td>6.3</td>
<td>1.0</td>
<td>3.6</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a2b1c1d1</td>
<td>5.7</td>
<td>1.9</td>
<td>2.6</td>
<td>1.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a2b1c1d2</td>
<td>5.6</td>
<td>1.8</td>
<td>4.3</td>
<td>1.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a2b1c2d1</td>
<td>5.9</td>
<td>1.6</td>
<td>4.2</td>
<td>1.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a2b1c2d2</td>
<td>4.3</td>
<td>2.4</td>
<td>3.2</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a2b2c1d1</td>
<td>5.9</td>
<td>1.7</td>
<td>1.9</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a2b2c1d2</td>
<td>6.3</td>
<td>1.0</td>
<td>3.0</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a2b2c2d1</td>
<td>4.9</td>
<td>2.8</td>
<td>2.2</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a2b2c2d2</td>
<td>5.5</td>
<td>1.9</td>
<td>3.4</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A = type of offense (a1 = physical abuse, a2 = theft),
B = severity of offense (b1 = serious, b2 = moderate),
C = social relationship (c1 = particularistic, c2 = universalistic),
D = status of offender relative to victim (d1 = superior, d2 = equal)

A 2 (Severity of offense: serious vs. moderate) x 2 (Importance: outcome vs. procedure) analysis of variance (ANOVA) with repeated measures on the last factor was used to test propositions 1 & 2. This analysis yielded a significant
main effect of importance, $F(1, 315) = 239.08, p < .001, \eta^2_{\text{partial}} = .43$, which was qualified by a two-way interaction, $F(1, 315) = 11.11, p < .01, \eta^2_{\text{partial}} = .03$. The main effect of importance indicates that outcome was perceived as more important ($M = 5.8, SD = 1.8$) than the procedure ($M = 3.6, SD = 2.1$) when evaluating the fairness of an offense regardless of its severity. The interaction effect qualifies the finding that outcome is more important than the procedure in that it indicates that this effect is larger when evaluating the fairness of a moderate offense as compared to evaluating the fairness of a serious offense. This finding provides support for Proposition 1. However, Proposition 2 stating that procedure would be perceived as more important than the outcome in evaluation of a moderate offense was not supported (see Table 2).

### Table 2
**Relative Importance of Outcome and Procedure in Fairness Evaluation of Severe and Moderate Offenses**

<table>
<thead>
<tr>
<th>Security of offense</th>
<th>Outcome</th>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Moderate</td>
<td>5.9</td>
<td>1.7</td>
</tr>
<tr>
<td>Serious</td>
<td>5.8</td>
<td>1.8</td>
</tr>
</tbody>
</table>

Moreover, to test Propositions 3-9 two sets of ANOVAs (see A and B below) were performed with type of offense, severity of offense, social relationship, and status as independent variables and importance of outcome and procedure as dependent variables in each set of analysis.

(A) The ANOVA on perceived importance of outcome revealed a main effect of type of offense, $F(1, 301) = 11.71, p < .01, \eta^2_{\text{partial}} = .04$; and a type of offense and social relationship interaction, $F(1, 301) = 7.21, p < .01, \eta^2_{\text{partial}} = .02$. The main effect of offense supports Proposition 3 in showing that outcome is more important when evaluating the fairness of physical abuse ($M = 6.2, SD = 1.5$) than theft ($M = 5.5, SD = 2.0$). Proposition 5 stated that outcome is more important when evaluating the fairness of an offense committed by a person of superior status ($M = 6.0, SD = 1.7$) as compared to equal status ($M = 5.7, SD = 1.8$). This difference was in the predicted direction but was not statistically significant. Proposition 7 predicted a type of offense and status interaction that was not supported. Specifically, the results showed that the difference in
perceived importance of outcome for evaluation of fairness of physical abuse and theft was the same across the unequal and equal status conditions \((Ms = 6.3 vs. 6.0\) for physical abuse and \(Ms = 5.6 vs. 5.4\)). The type of offense and social relationship interaction indicates that within a particularistic relationship outcome is perceived as equally important for fairness evaluation of physical abuse and theft \((Ms = 6.0 vs. 5.8)\), whereas within a universalistic relationship outcome is perceived as more important in evaluating fairness of physical abuse than theft \((Ms = 6.3 vs. 5.1)\). No proposition was stated for this exploratory finding.

(B) The ANOVA on perceived importance of procedure revealed a main effect of type of offense, \(F(1, 301) = 25.34, p < .001, \eta^2_{\text{partial}} = .08\); a main effect of severity, \(F(1, 301) = 13.56, p < .001, \eta^2_{\text{partial}} = .04\); and a type of offense and status interaction, \(F(1, 301) = 13.74, p < .001, \eta^2_{\text{partial}} = .04\). The main effect of offense supports Proposition 4 in showing that procedure is more important when evaluating the fairness of physical abuse \((M = 4.2, \, SD = 2.2)\) than theft \((M = 3.1, \, SD = 1.9)\). The main effect of severity indicates that the procedure is perceived as more important for fairness evaluation of serious offenses \((M = 4.0, \, SD = 2.1)\) as compared to moderate offenses \((M = 3.2, \, SD = 2.1)\). (Also, recall that the tests of Propositions 1 and 2 showed that the outcome was more important than the procedure for evaluations of both moderate and severe offenses.) Proposition 6 stated that procedure is more important when evaluating the fairness of an offense committed by a person of superior status \((M = 3.6, \, SD = 2.2)\) than equal status \((M = 3.6, \, SD = 2.1)\). As is evident from the means, procedure was perceived as equally important. Thus, Proposition 6 was not supported.

Proposition 8 predicted a type of offense and status interaction that was statistically confirmed. The pattern of means indicate that the interaction is related to the observation that procedure was perceived as more important in fairness evaluation of a physical offense \((M = 4.6, \, SD = 2.1)\) than theft \((M = 2.7, \, SD = 1.8)\) only in the unequal status condition. The corresponding difference was marginal in the equal status condition \((Ms = 3.8 vs. 3.5)\).

Proposition 9 was not confirmed in that the data were in the opposite direction to the prediction. Thus, procedure was perceived as less important when evaluating the fairness of an offense committed within a particularistic \((M = 3.5, \, SD = 2.2)\) as compared to a universalistic relationship \((M = 3.8, \, SD = 2.1)\). This difference was however not statistically significant.

**Discussion**

*Importance of outcome and procedure for justice*

*Evaluations of moderate and serious offenses*
As predicted by Proposition 1, and as suggested by our data, when people who are victims of a serious offense (in this case serious physical abuse) and evaluate such an incidence in terms of fairness, their conclusions are based more heavily on the outcome (i.e., the damage done) than on the way in which the damage was inflicted (i.e., the procedure). This prediction was tested in the context of two different kinds of offense: Serious physical abuse that was operationalized as violence to the victim’s face and stomach resulting in injuries requiring hospital care, and a $140,000 theft. Our data yielded the same result for moderate offenses (a slap on the face and a $70 theft). Thus, although statistically significant, the prediction in Proposition 2 (that the procedure would be more important than the outcome for evaluations of moderate offenses) was reversed. Intuitively, it seems reasonable to expect a focus on the damage done when making justice judgements about the incidence.

A word of caution against generalizing our findings is prudent here, as physical abuse may take several forms. The outcome might not always be more salient than the procedure when assessing this type of offense. Imagine a situation in which the victim was mercilessly tortured. Would not the way in which such a serious crime was carried out (i.e., the procedure) be the most focal and horrific aspect, even though the physical and mental outcomes certainly would be traumatic and life threatening for the victim? Further, the damage done (i.e., the outcome) may also vary in severity. Thus, for example, a small amount of financial loss incurred via petty theft may be of little significance for fairness assessments and assume less weight and importance than the procedure. The act of theft, per se (i.e., the procedure), might be very negatively experienced as disrespect, an invasion of my privacy, a violation of my rights, etc.. A petty theft may also be performed with great cruelty in which case the procedure would, again, be more salient than the outcome.

The point is, that if conclusions about the relative importance of outcome versus procedure for judgements of an offense is to be adequately meaningful, we certainly need further studies involving a variety of offenses, as categorized according to their nature (calling for a typology), and according to their varying degrees of severity. Further, nature and severity must be analytically disentangled to avoid that either be confounded by the other. Physical abuse, for example, is normally considered a more serious type of crime than offenses involving property. And physical abuse may, in turn, be characterized in terms of its severity or intensity. Thus, two types of evaluations of severity are apparently involved, making conceptual clarity indispensible.

The distinction between outcome and procedure is also crucial in all real life situations where attributions for behavior are made. The resulting treatment from others for an offense committed by a person may frequently vary, depending on whether the outcome/damage of the offense or the way in which it was committed is deliberately or undeliberately made focal. In a court context
the judge is required (and is likely to be trained) to consider both aspects when determining a just penalty for a crime. However, to what extent are members of the jury vulnerable to attributional biases due to possible unawareness of the two aspects?

**Importance of outcome and procedure for justice evaluations of person and property offenses, respectively**

As physical abuse is usually more condemned and reproachable than property offenses (e.g., theft), it should not be surprising if judgements on the basis of both the outcome and the procedure were more forthcoming and considered more important in the former than in the latter context. Accordingly, the predictions stated in Propositions 3 and 4 were confirmed by our data. Both the procedure and the outcome were considered more important for fairness judgements of physical abuse than of theft.

An observation with regard to the notion of procedure is noteworthy. Our study focused on how important it is to consider the way in which outcomes are accomplished, (the procedure relative to the outcome itself) when forming an opinion of how just or fair an event is. Thus, we were not asking subjects to rate the fairness of the procedure, per se, according to Leventhal’s (1980) formal criteria, for instance (i.e., ethicality, correctability, consistency, bias suppression, accuracy, representativeness), or according to interactional criteria (Bies and Moag, 1986) (i.e., the way procedures are enacted: truthfulness, justification, respect, propriety). In the event subjects’ ratings of importance were confounded by thinking about how fair they considered the way in which the outcomes (i.e., physical and property damages and losses) were accomplished, the only useful and meaningful formal and interactional criteria would be ethicality, and justification and respect, respectively. But these qualities would surely enter most types of behavioral evaluations (in terms of fairness, importance, acceptability, desirability, etc.). Future studies may need to be carefully designed to avoid this possible type of confounding bias.

**Impact of the offender’s status relative to the victim on the perceived relative importance of outcome and procedure**

A person of superior status who harms or takes advantage of others who are more vulnerable due to their inferior status positions usually receive intense social disapproval. Offenses committed by equal-status persons are likely to be viewed as less serious. Propositions 5 and 6 predicted that both the outcome and the way in which the outcome is generated will assume greater importance for justice assessments when a superior-status person is the offender as compared to when the status of the offender and victim are equal. However, these predictions were not statistically confirmed, although the data concerning the relative importance of outcome were in the expected direction.
In all experimental conditions status was operationalized in terms of the offenders' age relative to that of the victim. The equal-status vignettes described offenses by one male adult against a female adult, while in the superior-status conditions a female child was victimized by a male adult. Thus, there is a possibility that both age and sex were operational and had different kinds of confounding effects on the results. For example, perhaps age had a much greater impact in the adult-child physical-abuse conditions than in the theft conditions, while the impact of sex was negligible and similar in all conditions. On the other hand, age may very well be of little importance in the adult-adult conditions, while the male sex status of the offender is very focal and significant when the victim is a same-age female.

Recall that both the outcome and the procedure were considered more important for fairness judgements of physical abuse than for theft (according to Propositions 3 and 4, respectively). And this held true for both superior-status and equal-status conditions. In addition, we predicted that the difference between the importance of the outcome (Proposition 7) and of the procedure (Proposition 8) for fairness judgements of physical abuse and theft would be greater when the offender is of superior as compared to equal status to the victim. Only the latter proposition was confirmed. Thus, the results suggest that when the age statuses of the offender and the victim are equal, the difference between the importance of the way in which physical abuse and theft are inflicted (i.e., the procedure) is smaller than the same difference when the offender’s age status is superior. This was not the case when the outcome was focalized.

**Impact of the social relationship between the offender and the victim on the relative importance of outcome and procedure**

We mentioned that a significant number of studies have shown that the social relationship between provider and recipient appears to be a significant determinant of justice related attitudes and behavior, such as the endorsement of distributive justice principles. Another side of the justice coin is the procedure by which the distribution (outcomes) is accomplished. Both aspects usually enter in our views about how fairly we or other individuals and groups are treated. We hypothesized that the procedural aspect would matter more than the outcome aspect when assessing fairness within a particularistic social relationship, particularly when particularistic resources (the values of which are determined by the identity of the provider) are given or withdrawn (e.g., love, comfort, assistance, services, regard, and respect). The way affectionate sentiments are conveyed to my sweetheart (e.g., via kisses) would thus carry more weight than the amount of the provided outcome (i.e., the number of kisses). However, the amount of money received or lost is presumably more important than the way in which it was provided or withdrawn (taken away,
withdrawn). Thus, we made two predictions for future investigation - the procedure is more important than the outcome for fairness assessments of a particularistic resource allocation, and the outcome is more important than the procedure when the fairness of a universalistic resource allocation is determined. We previously argued that offenses (in this study, being beaten as well as robbed) committed by a family member or trusted friend are probably experienced as more painful and flagrant betrayals of trust, than when the offender is unknown or more distant. Further, this would presumably be particularly true when it comes to the way in which the offense was carried out. However, this prediction, as stated by Proposition 9, was not confirmed by our data. In fact, although not statistically significant, the direction of our prediction was reversed, hinting at the possibility that people would instead be more concerned with the procedure within the context of a distant relationship. The corresponding data for the outcome (for which no prediction was stated) yielded no difference between the two types of social relationship between the offender and victim.

In sum, the outcome (the damage done regarding the physical injury and the amount of stolen money) was considered more important than the procedure (the way the damage was incurred) for fairness judgments of both offenses regardless of their severity, the relative status of the offender, and the social relationship within which the offense was committed. Further, both outcome and procedure were viewed as more important when assessing the fairness of physical abuse as compared to theft. This was, for the physical abuse offense, more likely to be the case when the status of the offender was superior to the status of the victim.

References


**Author Note**

Kjell Törnblom, PhD, professor of social psychology and Ali Kazemi, PhD, assistant professor of social psychology, School of Technology and Society, Social Psychology Unit, University of Skövde. Correspondence should be addressed to Kjell Törnblom, ITS, University of Skövde, 54128 Skövde. E-mail: kjell.tornblom@his.se
GROUP DYNAMICS, PROFESSIONAL STEREOTYPES AND DOMINANCE – THE PERFORMANCE OF INTERDISCIPLINARY TEAMS IN HOSPITALS.

Endre Sjøvold & Anna-Catharina Hegstad

Abstract
This paper discusses group development in real life naval groups and groups of physicians in hospitals. In this study groups facing identical challenges under comparable environmental conditions are analyzed. The study covers groups from two different settings: military groups on maneuvers and medical groups during important strategic planning processes in hospitals. The setup for observation, two maneuvers for the military groups and strategic decisions with high impact and tight time schedules in the medical groups, emphasize the groups' ability to make the best decisions under rapidly changing conditions and/or high levels of stress. Shifts in role structure, predominant interpersonal behavior, and emergent dynamic patterns are discussed in the light of the SPGR theory and method. SPGR is an acronym for "Systemizing the Person-Group Relation" (Sjovold, 2007). The paper attempts to explain the dynamics of the social field observed. Both direct observation and peer reporting are used to sample the data.

Keywords: Group-dynamics, interdisciplinary teams, SPGR, health care, military groups

Introduction
The aim of this study is to identify the effect of dominant members on group dynamics, and whether or not the likelihood of dominance in interdisciplinary teams stems from professional stereotypes. The study has three distinct parts: The effect of dominance on group dynamics; the effect of dominant members entering a group; and the existence of professional stereotypes. We have chosen to focus on professions from the military and health care system, which are traditional and strictly hierarchic organizations, with clearly defined and dominating professions exposed to working tasks that in most cases can only be performed by interdisciplinary teams. The military and health care system are both designed to take care of assignments that vary from routine work in a calm environment with little need of leadership, to potential “life-death” situations where the stress levels are very high and demand clearly defined and strong leadership. In this paper we present findings from a study of change in dynamics in groups of naval cadets during one of their most challenging practical maneuvers and a medical group during an important strategic meeting. The two groups differ in the sense of professional homogeneity. The navel group consists of one profession with sub-specializations, while the medical team consists of at least two professions, physicians and nurses, where both professions are strictly
sub-specialized with their own distinct professional identity. In the hospital setting we have also investigated existing professional stereotypes, and we discuss how such constructions may influence the dynamic of interdisciplinary teamwork. Our findings are discussed in the light of the SPGR (Systematizing the Person-Group Relation) theory. The single most important factor influencing dynamic change in our groups seems to be the existence of dominant members.

The introduction of Maneuver Warfare and its implications for military leadership

Like most countries in the western world Norway has adapted the concept of maneuver warfare (Lind, 1985; William S. Lind, The Maneuver Warfare Handbook) for its armed forces. The reason is the expectation of more international operations and the challenges experienced in asymmetric warfare. Embedded in the concept is a distinct change from centralized leadership by lines of commands to a situation where the importance of mutual trust and decentralized decisions are stressed, or according to John Boyd:

” …to improve our capability to shape and adapt to unfolding circumstances, so that we can survive on our own terms, or improve our capacity for independent action”

(Boyd, POC, 1987, p. 58).

One practical consequence of introducing maneuver warfare is that we need to develop robust teams with a dynamic that fosters mutual understanding between its members and where all members may equally contribute.

Creating such dynamics in a military context is difficult due to its very nature and if created, such dynamics may appear to be fragile. A single person’s traits and behavior may completely change a team’s dynamic (Barry & Stewart, 1997; Stewart & Barrick, 2004; Williams & Sternberg, 1988). Williams & Sternberg found that even one overly zealous or domineering member in a group significantly inhibited the quality of that group’s performance. Due to the traditional stereotype of a military leader which is still alive, such situations are more than likely to happen.

The hospital organization and interdisciplinary teamwork

During the last ten years hospital organizations in Norway have faced increasing demands due to organizational and technological changes and claims to be more cost and time effective. This development coincides with increasing demands from society concerning both quantity and quality of hospital services, less tolerance for mistakes and more detailed political and economical control.

Many industries have successfully met such challenges by reducing organizational bureaucracy and converted to a more team-based work design (Mohrman et al. 1995). This is an approach that recently has also become
popular in hospital organization. This is especially true when introducing Lean production in hospitals (Ben-Tovim et al. 2007; Furman and Caplan 2007; Karvonen et al. 2007; Nelson-Peterson and Leppa 2007; Sunyog 2004; Long 2003; Herfarth 2003). Team-based work design is characterized by cooperation by people from different professions where every individual input is valued according to its contribution to task-solving, and not to the person’s status or profession.

In spite of several successful introductions of interdisciplinary teams in hospital organizations proving their efficiency (Freedman & Berger, 2004; Heineman & Zeiss, 2002; Warren et al. 1998), many authors cite difficulties when introducing interdisciplinary teamwork in hospitals (Phillips et al., 2002; Gerrish, 1999; Øvretveit, 1997). Explanations for these difficulties often point to the highly specialized health professions. According to Glouberman and Mintzberg (2001) a hospital is best described as consisting of four separate professional worlds: the world of “Cure” is inhabited by physicians, the world of “Care” by nurses, the world of “Control” by managers and mercantile professionals, and the world of “Community” by trustees. Members of the four worlds differ in their perceptions of the same task. The physicians enter the operation site to do their job and then disappear to their next case/patient, while the nurses take part in both the cure and the follow-up care. Compared to the physicians, nurses strive for a more holistic understanding of the patient and have a more local orientation characterized by strong loyalty to norms and regulations of their workplace. Furthermore, there are studies that show strong disagreement and even aggression between the professional groups in operating departments (Coe and Gould, 2008; Undre et al., 2006).

The similarities among the military groups are striking; both groups are supposed to change their dynamics to a more democratic way of working in a system and frame of reference where this is less likely to happen.

**Professions**

Professions are formed to provide a public service. In general terms a profession can be defined as an aggregate of people who find/establish an identity in sharing values and skills absorbed during formalized courses and intensive training, through which they all have passed. One important distinction between professions and other occupations lies in its demand and control of technical skills:

“The profession bases its claim by its position on the possession of a skill so esoteric or complex that non-members of the profession cannot perform the work safely or satisfactorily and cannot even evaluate the work properly”.

(Freidson, 1975 p. 45)
Another important distinction or characteristic is autonomy. Professional people have the special privilege of freedom from the control of outsiders, combined with the obligation of self-regulation. Its autonomy is both justified and tested by self-regulation (Freidson, 1975 p. 137). Autonomy is associated with both status and privileges. Professional autonomy and degree of freedom and control have a strong impact on the employee’s self-esteem and contentment at work (Gecas, Burgke 1995).

An important feature of professional autonomy is control. Professions do not only serve as instruments for achieving objective external goals as health care. The professions also have internal goals. They are also an interest/marketing group that seeks to protect their territory or monopoly, to preserve technical, intellectual and organizational dominance in relation to other professions (Sarfatti, Larson 1977). Abbot describes the professional system as a relational system referred to as the process of professionalization. The structuring of this system will thus be of vital importance when it comes to efficiency of the professions.

“It is control of work that brings the professions into conflict with each other and makes their histories interdependent. It is differentiation of work that often leads to serious differentiation within the professions. – The central phenomenon of professional life is thus the link between a profession and its work, that I shall call jurisdiction. To analyze professional development is to analyze how this link is created in work, how it is anchored by formal and informal social structure, and how the interplay of jurisdictional links between professions determines the history of the individual professions themselves”

(Abbot, 1988, 19-20)

There is a hierarchy among the different professions, which is closely linked to autonomy. Among the healthcare professions, medicine as a profession has the highest degree of autonomy, and thus the highest status – which primes them for a dominant role. Many healthcare professions are linked and touch upon the medical profession, but none of them can replace the physicians.

“Paramedical occupations, of which nursing is perhaps the most prominent example, are clearly in a markedly different position than medicine, for while it is legitimate for them to take orders from and be evaluated by physicians, it is not legitimate for them to give orders to and evaluate physicians. Without such reciprocity we can hardly consider them the equals of physicians”

(Freidson, 1975 p. 76)

Professions play different roles with different goals, and they have different value standards. The professions’ interpretation of their role is influenced by their tasks, perspectives and surroundings. Within one working environment, as in a hospital, where most of the professions are committed to the same task –
treated patients, you will find a wide range of interpretations of professional roles, which both complement and compete with each other.

**Professional stereotypes**

We have indirectly revealed/identified established stereotypes as one challenge to effective teamwork. In hospitals, professionals’ specific approaches to the same task seem to be reinforced as health care work is interpretative and typified by the need to constantly react to contingent events (Timmermans and Berg, 2003). In addition status and power are often associated with professions which often inhibit effective sharing of knowledge and experience in interdisciplinary teams.

In some cases professional stereotypes may be so distinct that they provide their members with a set of values and practices which are accommodated into the culture of organizations. As such strong professional cultures may be a strong hindrance for smooth interdisciplinary cooperation, since their interpretation of the problem at hand and superior goals may be different. On the other hand the existence of professional cultures may reinforce cooperation since different perspectives are crucial when a group faces complex tasks.

**Group development, dynamics and the SPGR theory**

We will base our further discussion on the SPGR (Systematizing the Person-Group Relation) theory on group development (Sjøvold 1995, 2006, 2007) and will therefore give a brief overview of the most important components of the theory.

**Four basic group functions**

The idea that the predominant behavior of a group differs in the course of its existence and when the group meets different challenges is fairly well established (Bion, 1961; Tuchman and Jensen, 1977; McGrath, 1991; Chidambaram and Bostrom 1996; Jern and Hempel 1999; Poole and Hollingshead 2004 or Sjøvold 2006). In SPGR terms we will say that the group mobilizes different *functions* to meet different challenges. The construct functions in SPGR are quite similar to Parsonian thinking (Parsons 1951, 1953).

The four basic group functions according to the SPGR model are: ‘Control’, ‘Nurture’, ‘Opposition’, and ‘Dependence’. The basic idea is that a group activates the function best suited to meet the specific problem they face. If the problem at hand is instrumental, then the Control function is activated; if the problem is relational, the Nurture function is activated and so on. When one of the functions is activated the predominating behavior of the group members reflects that active function. When the Control function is active, analytical, task-oriented or even autocratic behavior dominates; when the Nurture function is active, caring, empathic or even spontaneous behavior dominates; if the
Opposition function is active, critical, assertive or even self-sufficient behavior dominates; and when Dependence is active, passive, conforming, and obedient behavior dominates. Since an active group function is always reflected in group behavior, systematic observation of behavior is an efficient tool to investigate these phenomena. This is the approach used to study the groups referred to in this paper.

Group functions, balance and maturity

The construct ‘group constitution’ is defined as the balance of basic group functions. A group may activate one function to solve a specific problem and activate another to solve another problem. On the other hand a group may be stuck in one function even though that function is not adequate to meet the challenge the group actually faces. This phenomenon is similar to what Bion (1961) refers to as basic assumption groups in contrast to his high-performance ‘work group’. The SPGR parallel to Bion’s ‘work group’ is a mature group (Sjøvold 1995, 2005, 2006).

A mature group is a group capable of rapidly activating the group function best suited to meet any challenge at hand. To achieve such flexibility all members of the group need to be capable of performing behavior that supports all of the four functions. In less mature groups members tend to take on roles according to their zone of comfort, and limit their behavior to support one basic function. In such groups one member may be the caring person (Nurture), another person the achiever (Control) and so on. However, in a flexible group communication between members needs to be both distinct and rapid. Perceptions of the situation need to be shared, evaluated, decided and acted upon in a very short period of time. When all members are capable of recognizing and performing behaviors that support all four functions, this process is almost instant. The ‘one person–one role’ group will be less flexible, since each member perceives the situation and other member’s actions as well through the eyes of his or her role. A lot of negotiation needs to be done before the group is able to act.

Balance is an important concept of the SPGR model. ‘Balance’ is, however, not equivalent to the concept of equilibrium like Bales (1953,1955) describes as a homeostatic controlled status quo. ‘Balance’ is a constant shift and polarization between active group functions. Then members of a group free themselves from fixed roles and they become capable of performing behavior that supports all functions. A well-performing or well-balanced group can be compared to a gyroscope. It is the speed of rotation that makes it stable and robust. The ‘one person–one role’ group may balance the group function by having an equal number of supporting roles, but such a group will respond very slowly and be vulnerable to environmental change.

There are two more aspects of group constitution used in our analyses. First: Synergy and Withdrawal. Synergy appears in groups where the basic group
functions are well balanced and characterized by engagement and constructive goal-oriented teamwork. At a lower level of maturity where members still commit to their initial role preferences, they tend to restrict themselves from contributing to the common group work, which in turn results in passive behavior and resistance, i.e. Withdrawal. Second: Influence and Passivity. Group-members may differ in influence. In a synergy group members tend to have equal influence while in a withdrawal group dominant members are more apt to appear.

**Group tasks and effectiveness**

Group maturity, as defined here, is closely related to role structure. The more specific roles group members assume, the less flexible and responsive the group will be. The interdependence of individual and group development is also obvious. As members expand their behavior repertoire and skills, the group also becomes a better arena for learning. The individual needs the group to develop, and the group will only develop through its members (Mills 1984). Innovative groups have a high capacity to learn and are, in our terminology, mature.

All groups do not need to be innovative or mature to be effective. Group effectiveness is a highly flexible concept (McGrath 1991; Gersick 1988; Hackman 1983, 1992, 2002). In this paper we define *group effectiveness* as how well group resources are mobilized to solve a specific task. Groups may be effective even when operating on a low level of maturity. Effectiveness is always related to group task and context. The more complex tasks are, and the more unpredictable the context is, the more mature the group needs to be for success. If the task is simple or dividable and the context is fairly structured, a fixed role structure may be more effective and more forceful due to its ability to be focused. No creative noise distracts them from fulfilling their task. A team of surgeons and nurses who perform their specialized tasks in a strictly coordinated manner under the senior surgeon’s command is a good example.

Groups tend to adjust to the level of maturity best suited for the task at hand. However, the higher levels of maturity require members with well-developed personal skills. If the group faces challenges beyond their competencies, they may end up in dysfunctional reactions like positive or depressive groupthink (Janis, 1972; Granström and Stiwne 1998), polarization and conflict.
Method

Participants and design

This study consists of three distinct parts:

<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups with members of equal status</td>
<td>Naval cadets during a challenging maneuver</td>
<td>Identify changes in group dynamics over a period of time in groups with and without dominant members</td>
</tr>
<tr>
<td>- Peer ratings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interdisciplinary teams of members with unequal status</td>
<td>Physicians and nurses during a 5 hour strategy meeting</td>
<td>Identify how status differences and dominance influence interaction in situ</td>
</tr>
<tr>
<td>- Direct observation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional stereotypes</td>
<td>Physicians and nurses across units in three different hospitals</td>
<td>Identify implicit theories of interacting professions</td>
</tr>
<tr>
<td>- Survey</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Peer ratings from groups of naval cadets:
Data from the naval cadets were gathered before, midway, and after finishing what is supposed to be the most challenging maneuver during their first year at the naval academy. Starting at the academy the cadets were distributed in eight-person groups which were maintained during the first year including this maneuver. The cadets knew their group members well having worked together for several months. In some way these groups resemble a laboratory setting since all groups perform identical tasks in identical contexts. For the cadets, however, the group is very real since their performance as team members may to a high degree influence their future as cadets and officers. Their performance as team members were constantly monitored and evaluated.

Observation of interdisciplinary teams in hospital.
The teams observed consisted of both physicians and nurses who worked together for several years with the exception of a couple of members. In the meeting observed the topics of discussion were of strategic importance and had both personal and emotional impact on the individual.
Survey for professional stereotypes in hospitals
The respondents for the survey on stereotypes were physicians and nurses from three different Norwegian hospitals and throughout different hospital units.

Instrument and procedures
The SPGR instrument:

The Systematizing Person-Group Relations (SPGR) instrument reflects the constructs of the theory and consists of four dimensions labeled Control - Nurture (C-N), Opposition - Dependence (O-D), Withdrawal - Synergy (W-S) and Influence – Passivity (I-P). The pairs of the first two dimensions represent what is defined as basic group functions. The pairs of the third dimension are indicators of a group’s maturity, and the fourth dimension reflects the degree of influence or dominance in the group.

A group’s constitution along these dimensions are identified either by direct observation using the SPGR category system for observation or the SPGR scales for self and peer ratings of behavior in groups (Sjøvold, 1995, 2002). Both procedures are used in this study.

The SPGR behavior scale consists of 24 items describing behavior in a set of three word or phrases:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Cautious, reliable, willingly assumes duties</td>
</tr>
<tr>
<td>7</td>
<td>Caring, supportive, encouraging</td>
</tr>
<tr>
<td>8</td>
<td>Effective, self-confident, dares to take the lead</td>
</tr>
</tbody>
</table>

The respondent is asked to rate according to whether the behaviors never or seldom occur (1), sometimes (2), or often or always occur (3) by the subject in question.

The twelve categories for direct SPGR observation are listed in Table 1. The vector code indicates which dimension it belongs to; Control vectors are labeled C1 and C2 and so forth. I-P scores are indirectly tapped from the scores in these twelve observation categories.
Table 1. The SPGR observation categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Code</th>
<th>Typical behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task-orientation</td>
<td>C1</td>
<td>Controlling, autocratic, attentive to rules and procedures</td>
</tr>
<tr>
<td>Task-orientation</td>
<td>C2</td>
<td>Analytical, task-oriented, conforming</td>
</tr>
<tr>
<td>Relation</td>
<td>N1</td>
<td>Taking care of others, attentive to relations</td>
</tr>
<tr>
<td>Creativity</td>
<td>N2</td>
<td>Creative, spontaneous</td>
</tr>
<tr>
<td>Loyalty</td>
<td>D1</td>
<td>Obedient, conforming</td>
</tr>
<tr>
<td>Acceptance</td>
<td>D2</td>
<td>Passive, accepting</td>
</tr>
<tr>
<td>Criticism</td>
<td>O1</td>
<td>Critical, opposing</td>
</tr>
<tr>
<td>Assertiveness</td>
<td>O2</td>
<td>Assertive, self-sufficient</td>
</tr>
<tr>
<td>Resignation</td>
<td>W1</td>
<td>Sad appearance, showing lack of self-confidence</td>
</tr>
<tr>
<td>Self-sacrifice</td>
<td>W2</td>
<td>Passive, reluctant to contribute</td>
</tr>
<tr>
<td>Engagement</td>
<td>S1</td>
<td>Engaged, inviting others to contribute</td>
</tr>
<tr>
<td>Empathy</td>
<td>S2</td>
<td>Showing empathy and interest in others</td>
</tr>
</tbody>
</table>

Peer ratings from groups of naval cadets:
The cadets were asked to rate their own group using the 24-item SPGR behavior scale. The questionnaire was distributed on paper just before, midway and just after the maneuver. We obtained a total of 66 complete ratings (a response rate of 84%). Using the pre and post measures we identified teams with and without dominant members and analyzed for changes in dynamics in the two groups of teams.

As a measure of changes in team dynamics we used changes in scores along the C-N, S-W and I-P dimensions of the SPGR behavior space. To identify teams with dominant members we used the difference between the median of the particular group and its most dominant member. Teams where this difference was higher than the median for the complete sample were defined as a “team with dominant members”. Teams having a median equal to the median for the complete sample were not included.

Observation of interdisciplinary team in hospital:
The teams were both observed in situ and videotaped. The videotapes were later used for more detailed observations. The SPGR twelve categories for direct observation were used by two observers. Both observers had passed a standardized inter-rater reliability test. The ratings were performed using a tablet computer with a touch-screen and the SPGR observation software.

The team we used as an illustration for this paper had a total of eleven members of which there were 6 physicians and 5 nurses.

Survey for professional stereotypes in hospitals:
In this survey respondents were asked to describe the “typical” nurse and physician using the 24-item SPGR behavior scale. The survey was distributed through a tested web solution. The response rate was 51% with a total of 68 complete ratings. Of the 68 respondents in this sample, 25 were physicians.

Results

The effect of dominance on group dynamics in groups with members of equal status
Table 2 summarizes findings from our study of Norwegian naval cadets during their most challenging maneuver. All data are/represent peer ratings using the SPGR behavior scale. Figure 1 and 2 exemplifies our findings of two of the groups in the sample:

<table>
<thead>
<tr>
<th>Group type</th>
<th>n</th>
<th>Variance: ΔPre – post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without dominant member</td>
<td>26</td>
<td>3.61*</td>
</tr>
<tr>
<td>With dominant member</td>
<td>32</td>
<td>2.86*</td>
</tr>
</tbody>
</table>

*P< 0.010 total n = 66

In Table 2 the variance between pre and post measures in groups with and without dominant members are compared. From the table we see that groups with dominant members tend to have a more fixed role structure (less variance from pre to post measure) than groups with no dominant members. In groups with a fixed role structure group members tend to stick to their position, or role, in the group and maintain their initial opinions. This gives a stable and predictable dynamic in the group. Operating in a fairly stable context such groups tend to be effective when their task is relatively simple, but tend to fail confronted with unexpected events. Quite contrary groups with a flexible role
structure tend to be more creative and able to cope with complex tasks and contexts.

To illustrate our data findings we have used the printouts of the SPGR analyses. Group members in these displays are plotted as circles in a three-sector template (see for instance Figure 4 for a full illustration). In the top (Control) sector behaviors that supports structure and task orientation are plotted. In the lower right (Nurture) sector, caring and empathic behaviors are plotted, and in the lower left (Opposition) sector behaviors identified with criticism, withdrawal and even rebellion are plotted. The size of the circle indicating each person in the group expresses the influence the person has on the group--the larger the circle, the higher influence. In figure 1 and 2 we have used a subset of the SPGR diagram showing the upper right quadrant.

![Figure 1: Team B Member with equal influence and flexible rolestructure](image)

Figure 1 illustrates how a group with no dominant members (all circles are equal) displays a highly flexible role structure. The figure shows the role structure prior to, midway (under) and right after the MLS maneuver. Even though there are some recognizable patterns, the members have experienced their group dynamics quite differently depending on their most recent experience. In practice this expresses group members ability to take on different roles and perspectives depending on the needs of the situation. In such groups the potential contribution from each member is optimally utilized. This group was able to solve the most complex and challenging mission to which they were confronted. Another characteristic of this group was their appearance; they
solved their task low-key with extensive exchange of non-verbal transactions between all members.

---

**Figure 2: Team A Group with one dominant member and fixed role structure**

Figure 2 illustrates how a group with a dominant member displays a fixed role structure. In this group the role structure prior to, midway and after the MLS maneuver displays minor and no changes. The dominant person runs the game and all other members assume submissive roles. Confronted with well-known tasks that could be successfully met by drilled solutions, this group was extremely effective. Confronted with complex tasks in an unforeseen context, they always failed. In contrast to the group in Figure 1 (Team B), this group was fairly noisy with extensive command-like type of communication. They also tend to brag about their field experience, something very unlikely for team B.

*Interdisciplinary teams of members with unequal status. The effect on group dynamics when a dominant high status member enters the group.*

Figure 3 show a situation where a senior physician enters a strategy meeting where both physicians and nurses took part. The illustration was selected from a series of structured observations of similar meetings in Norwegian hospitals. The twelve SPGR category observational system is used.
In situation A we see a group of people concerned with a task of strategic importance and with severe practical and emotional impact for the individual. All members operate in the task-oriented sector of the behavior space and their influence is fairly/quite evenly distributed, except for the person at the bottom of the diagram who is totally withdrawn who is sitting laid-back, showing no interest in the discussion.

The meeting has been going on for about twenty minutes when the senior physician enters the room. Situation A is a snapshot just before this event and situation B occurs a few minutes later. What we see is a rapid and dramatic shift in the group’s dynamic. In situation B some members of the group surround the senior physician while others (the majority) withdraw. This change is towards a situation where less members contribute with their opinions, knowledge and skills. These members’ skills and experience may have contributed to better preparation and better performance of the team, but are no longer available for the group.

A typical pattern is that those members who withdraw in this manner are nurses. The nurses’ general behavior demonstrates that they depend on the physicians and the physicians typically dominate the team dynamics. Although communication appears to be fairly open on the surface, an apparent fear of criticism easily leads to misunderstandings, a tendency to oversee deviations and lack of feedback during group action.

Figure 3: Structured observation: a significant shift in group dynamics appears when a senior physician enters a meeting. A: just before he enters and B a few minutes after his entrance.
The existence of professional stereotypes and their effect on group dynamics. In Figure 4 the difference in professional stereotypes between physicians and nurses is diagrammed. The stereotypes were measured by a survey in three hospitals and a variety of departments using the SPGR value scale.

Figure 4: Difference in professional stereotypes

The stereotype of physicians is described as dominant and leans towards the autocratic sector of the behavior space, while the stereotype of nurses is distinctly more submissive and leans towards the nurture sector.

Sorting our findings by professions (Figure 5), we find they always tend to describe the other profession towards a less collaborative direction in the space compared to their description of their own. This probably expresses norms of team behavior that are distinctly different between nurses and physicians. We observe that while nurses are highly conscientious, the physicians enjoy a significantly higher degree of freedom.
Figure 5: Physician rates physicians (A) and nurses (B), and Nurses rates physicians (C) and nurses (D)

Professional stereotypes act as implicit maps for acceptable behavior toward other groups of professionals. Their mission is to establish professional identity and functional-role expectations. However, they may be effective hindrances when a group confronts organizational change or unpredictable situations. Strong professional stereotypes as identified in our sample are strong reinforcements for conserving status quo.
Looking in more detail into the descriptions of stereotypes using the twelve SPGR categories, we find systematic differences illustrated in Figure 6 (the value for each category in the plot have a range from 0 to 9). The nurses describe the physicians as extremely task-oriented and dominant, while themselves as caring and extremely loyal. The physicians recognize the nurses descriptions although they describe themselves as more caring and loyal than the nurses do. When physicians describe the nurses, they find them submissive, nagging and complaining in addition to their caring. This description is quite different from the nurses description of themselves.

Given that professional stereotypes act as maps that guide joint cooperation, it is obvious that the differences identified in our sample may have severe negative effects on the performance of interdisciplinary teams.

**Discussion**

In this paper we present findings from a study of the effect of dominant members on group dynamics, and whether or not the likelihood of dominance in interdisciplinary teams stems from professional stereotypes.

*The effect of dominance on group dynamics in groups with members of equal status*

Our findings indicate that dominant members may influence the group dynamics in a dysfunctional way. Dominant members in groups tend to act as point attractors freezing the structure of the group. This phenomenon is illustrated in our sample by less change in a member’s position in the SPGR space in groups with dominant members, even after stressful events that are particularly designed to induce a new consciousness of the group dynamics. Static dynamics
indicates that members tend to use more attention and energy on preserving their role in the group than releasing their resources in joint collaboration. In SPGR terms we would say that the distortion of balance along the influence dimension regresses the group to a less mature level. The group is no longer able to utilize its resources in an optimal way.

According to the principles of maneuver warfare such groups are quite contrary to what is wanted. What is wanted is groups where members have equal influence when contributing to task-solving and are able to understand the superior mission in a way that they are able to take the correct action even if it means disobeying orders in a given situation. Such groups are characterized by a high level of trust between members of the group, but also a high degree of trust in their superiors and the system supporting them. To be able to create a shared understanding of a superior mission and how their superiors think, they need a high degree of interaction on an equal basis.

You may think that the energy of strong leadership could be channeled to mobilize mutual sharing and induce a well-functioning group dynamics. In our sample this is not the case. In the groups of cadets there were no formal leaders nor anyone with a status higher than other members. In the hospital groups were different.

**Interdisciplinary teams of members with unequal status. The effect on group dynamics when a dominant high status member enters the group.**

Having identified how dominant members may negatively influence equal status group’s dynamics, this paragraph concerns the effect of dominant members in groups with clearly defined differences in status.

Our observations showed an immediate and dramatic shift in the dynamics towards a lower level of maturity when a high status member of the group, a senior physician, enters the room. This phenomenon may have several explanations. Physicians have traditionally played a dominating role. Physicians have been looked upon as the most prominent among members of the generally recognized profession (Freidson, 1970). The present results support this assumption. The physicians are outnumbered by the nurses, but retain their control. The dominating role of the physicians can be explained by several factors. The medical profession has a long tradition and they exhibit exclusive technical skills. The majority of senior physicians are male, and practicing medicine is looked upon as highly prestigious work. The medical profession furthermore have strong unions. Political, legal and economic changes during the last decades favor devaluation of the traditional dominance of the physician:
their autonomy and control has been limited. There social status has been reduced; most hospitals and their departments are run by nurses and financial experts. In hospitals they are enrolled as employees, just as everyone else, without any formal leadership. Still the physicians seem to manage to maintain their territory. The most important reason for this is that, when it comes down to patient treatment, which is the main task in the health system, the physician has the last word. They are in control of the key knowledge and skills to perform patient treatment. The nurses serve this purpose, but cannot replace the physicians’ role in performing this task.

A high degree of medical order and dominance is both needed and appreciated in order to ensure a patient’s safety and efficient medical treatment, especially in acute situations as when dealing with multi-traumatized patients. The chief surgeon should be an authoritarian leader of the trauma team. He or she should be in total control and everyone should pay attention to his or hers moves and decisions. Extensive autonomy in one setting will, however, easily be transferred to settings where is not needed and not efficient. The strategic meeting in the medical department in the present study does not call for an authoritarian approach. In strategic planning, it would be more efficient if everyone contributed. Physician dominance then becomes a negative factor that inhibits the group process. The nurses withdraw, and their knowledge and skills are not utilized. The negative dominance of the physician was imposed by shifting/changing the focus by talking about medical questions that the nurses were not competent to express their opinions. The temperature of discussion was raised characterized by rapid transactions using medical terms and rough language. The physicians turned their attention towards each other without paying attention to the nurses using their dominance to exclude others. The atmosphere was tense and cold without any smiling. In this meeting the group functioned on a very low level of maturity. There was a large gap between the task at hand and the appropriate level of maturity on which the group was operating.

According to SPGR theory, groups may change their way of operating and confronting different situations and tasks. In a different meeting, with the same group discussing the same topic, we saw quite another dynamic. We experienced the positive effect of physician dominance. The physician used less medical terms and reinforced contribution from all group members, including the nurses. The nurses took part in the discussion on an equal level as the physicians responded positively to the nurse’s opinions. The temperature of the discussion was still high, but in a way that raised the positive energy creating an atmosphere where different opinions were appreciated. In this setting the physician recognized that the task they were obligated to perform could not be completed in time without the nurses. The climate in the group changed when one of the senior physicians followed up one of the nurse’s suggestion on a
strategic priority. The same suggestion was presented in an earlier meeting, but then did not catch any of the physicians’ attention, and was thus not followed up. This exemplifies how dominance from high status members may induce positive energy in a group.

The existence of professional stereotypes and their effect on group dynamics. Our findings of differences in stereotypes between nurses and physicians may explain some of the difficulties experienced in interdisciplinary work in hospitals. Even though professions are an effective way of organizing work, it may as well hinder effective interdisciplinary work if institutionalized stereotypes value one profession over another. Professional autonomy and degree of freedom and control have a strong impact on the employee’s self-esteem and contentment at work (Gecas, Burgke 1995), but when seen as a privilege or universal right to do whatever you want or to dominate in a negative way, it may be fairly destructive.

Most teamwork labeled as Interdisciplinary in hospitals is best described as coordinated activities with shared responsibility and tasks between professions. The operation team is a good example. Every member contributes with their specialized skills and knowledge, in sequence and under strong supervision of the senior surgeon. If everything works according to the plan and predefined procedures, this is the most efficient way of working with this task in that context, even if we would characterize such a team as operating on a low level of maturity.

Even if effective under normal conditions, teams used to operating on a low level of maturity may have problems shifting to a higher level when unforeseen events appear. If the team is not able to change their perspective to adequately meet the new situation, they often tend to induce more structure and control, or in other words, do more of what does not work. Most of them reported erratic behavior in similar situations (Coe and Gould, 2008; Cole and Crichon, 2006; Thomas et al. 2006) which may be due to malfunctioning group dynamics.

Our findings suggest that this is more than probable. They also suggest that it is less probable that an experienced operational nurse will raise her voice to a more senior physician who performs a erroneous action. It is also probable that if she did, the physician would interpret her comments as nagging and not meant seriously (Undre et al 2006; Mills et al 2008).

It is obvious that the different professional levels of autonomy among health care professions in our limited sample may be in conflict. There is a tremendous potential for improving interdisciplinary communication. To unleash this potential may place considerable demands on the physician’s ability to change and practice their implicit leadership in the hospital organization.

**Practical implications for interdisciplinary teamwork**
Due to their knowledge and skills the physicians will necessarily play a dominant role in the hospital settings. This fact has often been commented on in a negative way and can, as shown in the present study, have negative effects on the performance of interdisciplinary teams. However, instead of trying to reduce the physician’s dominance, relative to the nurses, is it possible to turn the energy it represents into something positive for interdisciplinary teamwork. While the physicians are expected to constantly upgrade their medical knowledge and technical skills, there is less focus on their non-technical skills. There is, however, an increasing demand for ways to develop non-technical skills among physicians, even in acute situations (Mills et al. 2008; Mishra et al. 2008; Cole and Crichton, 2006; Makarv et al., 2006; Thomas et al., 2003). If these demands also include the ability to facilitate group dynamics, we may also be able to expand the role of physicians from being an expert and controller into taking on responsibilities on the relational level. The physicians will still perform the medical treatment, and nurses the patient care, but both professions will have the responsibility of imposing non-technical skills. In addition to analyzing the jurisdictional professional links, we will resolve and perhaps reform these links in order to create more efficient teams, and thus advance Abott’s professionalization process one step further. When this dominance is used positively, it will improve the dynamics and thus the efficiency of interdisciplinary health care teams.

One of the main barriers for such a mental shift may be the difference in values and norms among the physicians and nurses. The physicians’ ethical code has its origin in the crossing of traditional teleological and deontological frame(work)s, which correspond to the Control sector in the SPGR space. Nurses take on individual, relational care ethics belonging to the Nurture sector in the SPGR space. The greatest challenge to this change is that the old privileges and status of physicians have to be removed. On the other hand they will acquire a new position and status that will better support their role as treating the whole patient, and create an environment of both effectiveness and care, which enhances their professional legitimacy.

References


psychology (2nd ed), 3, 199-297, Palo Alto CA: Consulting Psychologists Press


Sjøvold, E. (2002). *The SPGR manual, Oslo*


Author Note
Assoc professor MD, PhD Anna-Catharina Hegstad, Department of Cancer Research and Molecular Medicine. Faculty of Medicine Norwegian University of Science and Technology, NTNU, Trondheim, Norway
Assoc professor PhD Endre Sjøvold
Institute of industrial economics and technology management, Norwegian university of science and technology, NTNU, Trondheim, Norway.
Correspondence concerning this article should be addressed to Endre Sjøvold Institute of industrial economics and technology management, Norwegian university of science and technology, NTNU, N-7491 Trondheim, Norway.
eMail: endre.sjovold@iot.ntnu.no
Summarizing the impressions of the 2008 GRASP conference is not easy. Group research in Sweden has expanded from quite a modest start in the beginning of the 1970s when I wrote my licentiate thesis, a social psychological analysis of two psychotherapy groups. In order to get a good picture of group research, I traveled to Harvard and many other places and met Robert F. Bales, Ted Mills and Paul Hare who were all very well-known names in group research. At that time, group research mainly concentrated on studies of groups in laboratory conditions. Questions were characterized by the fact that the design chosen had to answer the questions asked. To what extent results could be generalized to ordinary groups, where people depended on each other even outside the frame of the project, and where group work was of real importance, was not found interesting.

My interest in what happens in ordinary working groups got the better of me and I came in contact with the, at that time, quite extensive working life research conducted mostly within different working life institutes but also at universities. There was a small group of working life researchers well-known internationally. The differences in approaches between the laboratory group studies compared to the studies made within the frame of ongoing activities within the industry by the help of so-called natural groups were huge. Efforts to study working group processes systematically by using video recordings and interviews was a way for me to find a bridge between these two approaches. This became my doctoral thesis, which was later on followed by many studies within different areas.

My experiences make me regard this year’s theme as both important and educational. It is also possible to notice that today’s use of laboratory conditions has undergone an exciting development, even if it will take time to convert the results into practice. There is a wide range of studies of naturalistic groups, among which there are also some interesting studies, with a high level of generalization, but also some, which would benefit from learning from the experiences offered by today’s laboratory studies. Judging from some of the studies sent in for reviews in international journals there are also some studies opting for the opposite that is what systematic laboratory studies can learn from more naturalistic studies. In my opinion a creative and open exchange of thoughts and experiences between these two traditions is an important step in the development of group research in the future.

The international contribution to this year’s GRASP conference was professor Susan Wheelan (USA), a well-known lecturer and author of several books on group processes and teamwork in different settings. She started the conference with a lecture on Researching Work Groups in Natural Settings. She focused on
the importance of systematic studies of work groups in natural settings and described her model of group development, used in several settings. Various examples of how to handle the difficulties connected to this particular kind of studies were given.

Her second lecture was a more thorough presentation of the model she had given an overview of. The title was: Helping Work Groups Be More Effective: A Research-based Approach. The conference participants had the possibility to try to estimate the effectiveness of a group they belonged to with the help of the form for group development established by Wheelan et al. This vitalized the discussion in an exciting way and gave room for questions regarding how one can prove that various groups differ in their development. A very positive aspect of her participation was that she was present during both days. It gave a feeling of continuity. I cannot remember any of the former keynote speakers having done that.

When I look back at the parallel seminars and the discussions they raised, I feel that the overall aim of the conference, to be an open forum for discussing methods, results and problems concerning group and organizational research was achieved in a very sympathetic way. In addition, the dinner outdoors at a restaurant in Lund was an inspiring experience, both culinary and socially with many new contacts and a renewal of many older contacts. It actually says something about continuity and renewal of the GRASP conference.

The theme of this year’s conference was to analyze how different research methods can cooperate and inspire each other, which was meant to be mirrored in the title “Dynamics within and outside the lab”. It is always difficult to find a precise theme for a conference. There has to be a theme and it has to have a broad approach in order to awaken the interest of as many researchers as possible. The theme title has to work as an intersection or a meeting point, but at the same time point at something that may simplify the choice of focus for those who want to present their research.

All in all 27 projects were presented in parallel sessions. In relation to the theme presented in the invitation, my impression is that several of these studies compared different research methods, but not in the way the arrangers had hoped for, that is the choice between an experimental design and a naturalistic approach. As has been accounted for in the introduction, the most common methods were questionnaires, observations and interviews and in some cases the discussion connected to a presentation discussed which type of information different methods had gained. However, I think it might have been interesting to problematize the choice of method. The suggestion of a round table discussion presented in the invitation might have contributed to lift this aspect of methodical questions in a new way.

The focus was rather on the choice of areas of application. My view is that the majority of the studies built on data from areas such as primary care, maternity
care, health care, psychiatry, home help services, education and occupational health service with a focus on student health service. Two more classic areas of research within the field, which attracted some participants, were leadership in working life and in child groups, and inter-group phenomena exemplified by studies of interdisciplinary groups and the cooperation between customer and performer. Destructive processes in small groups were exemplified by studies of bullying in child groups and by what was called a collective collapse in a football team. There were also studies focusing on organization, learning and identity as well as analyses of interdisciplinary teams. More classical social psychological areas were represented by studies of justice, attribution and social effects as well as by some studies of cross-cultural and ethno-cultural studies.

Compared to earlier conferences I can see some new angles. One of these is economics and psychology illustrated by two studies of investments and majority influence in the stock market. Another one concerns companies and entrepreneurship. There were also studies discussing sports, both in its real interpretation where focus was on the practice of team sports and one discussing demonstrations related to competitions. A third area concerned traffic research and risk behaviour and how one through a method of reflection can improve traffic safety in young male drivers.

A general impression is that the GRASP conference fills an important social and collegial aim. My view is that the conference offers a forum where you quite freely can present a project within group and organization. Through the discussion connected to the presentation and through the contact with other participants one is able to get constructive remarks on how to move on. Many projects seem to be on going, and the conference gives a possibility to raise specific questions and to get feedback. It is important to stick to the rule that those attending the conference should have an on-going project. Everyone should not have to present every time, but they need to be in an on-going research process and to be open to other theoretical approaches and methodological considerations.

It is evident that many researchers return and even if it is common that the projects they present are parts of their doctoral studies, there are also projects characterized by development work within a certain area of application. For some the conference is also a forum where one can try some more playful projects, which may lead to stimulating discussions. It is also easy to see that a large part of working life research focused on leadership within organizations, organizational development, entrepreneurship, family businesses, non-profit organizations, working life and health is generally not represented here. When represented, it is mostly doctoral students who view the conference as a good opportunity to test their theories and methods. Other more specialized conferences, both national and international, fill that void.
The Nordic contribution isn’t very pronounced. Every year, there are a small number of participants from Norway or Denmark, which might be enhanced by getting information out earlier. However, I do believe that this is due to personal contacts on different levels. One category attending is the rather limited group doing group- and organizational research in Sweden with a starting point in more group dynamic oriented theories and models. There are, however, similar research traditions both in Norway and Denmark, which could be developed.

Stockholm, April 2009

Siv Boalt Boëthius
Professor Em.
Department of Education
Stockholm University