Is group performance improved by evaluating task difficulty and by knowing about the differential effects of community?

Author(s):
Hansmann, Ralph; Crott, Helmut W.; Mieg, Harald A.; Scholz, Roland W.

Publication Date:
2003

Permanent Link:
https://doi.org/10.3929/ethz-a-004520885

Rights / License:
In Copyright - Non-Commercial Use Permitted
Is Group Performance Improved by Evaluating Task Difficulty and by Knowing about the Differential Effects of Conformity?

Ralf Hansmann, Helmut W. Crott, Harald A. Mieg, and Roland W. Scholz

April 2003
Is Group Performance Improved by Evaluating Task Difficulty and by Knowing about the Differential Effects of Conformity?

Ralf Hansmann, Helmut W. Crott, Harald A. Mieg, and Roland W. Scholz

Contents

Abstract ................................................................................................................................. 1

Is Group Performance Improved by Evaluating Task Difficulty and by Knowing about the Differential Effects of Conformity? ................................................................. 2

Method .................................................................................................................................. 5
  Basic Experimental Design ........................................................................................................ 5
  Participants, Experimental Setting and Tasks ........................................................................ 5
  Experimental Procedure ....................................................................................................... 6

Results .................................................................................................................................... 9
  Judgement Accuracy ............................................................................................................... 9
  Results of the Post-experimental Questionnaire .................................................................. 11

Discussion ................................................................................................................................ 12

References ............................................................................................................................. 14

Abstract

A field experiment investigated whether group performance is improved when the subjects a) individually conduct an evaluation of task difficulty before entering the process of group decision making and/or b) know about a result from group decision making research which suggests that conformity has a positive impact in easy tasks but negative impact in difficult tasks (Grofman, 1978; Thomas & Fink, 1961). Twenty-eight groups of three or four persons worked on judgmental tasks related to an environmental planning project in which they were participating for several months. Group judgements were significantly improved when subjects were advised to evaluate the task difficulty before entering group decision making compared to a set of control groups. The interpretation focuses on cognitive aspects of the evaluation procedure, which might counteract anchoring effects and enhance the generation of arguments. Though the trend was in line with the hypothesis, there was no significant result with respect to the information on the differential effects of conformity pressure in easy versus difficult tasks. A post-experimental questionnaire revealed that most of the subjects did not know about the differential effects of group conformity but considered a task difficulty analysis to be helpful and supportive for group decision making.

Key words: conformity pressure, group discussion, group technique, group judgement, judgmental accuracy, planning groups, task difficulty.
Is Group Performance Improved by Evaluating Task Difficulty and by Knowing about the Differential Effects of Conformity?

Environmental and organisational planning often takes place in small and medium sized groups that combine experts, politicians and citizens in round table discussions or focus groups (Kasemir et al., 1997; Krueger 1988). Group discussions mostly take place rather unstructured as free interaction processes. However, free interaction processes show some characteristics weakening their effectiveness. The following issues are of interest:

- **Dominance**: Single persons or particular topics often dominate the discussion in freely interacting groups (Delbecq, Van de Ven & Gustafson, 1975). This is especially harmful if the participation rate of group members does not correspond to their expertise or if important ideas and topics are suppressed (Janis, 1972, 1997; Steiner, 1972).

- **Disturbance**: In freely interacting groups, there is no real opportunity to work independently at a given task without being disturbed by others (Gustafson, Shukla, Delbecq & Walster, 1973; Witte, Sack & Kaufman, 1998).

- **Conformity Pressure**: Group norms, conformity pressures and tendencies of social desirability can exert uncontrollable influences on group members (Asch, 1956; Festinger, 1950, 1954, 1957; Janis, 1972, 1997). Opinion convergence can lead to a premature closure of a group discussion before all the information is exchanged. Conflict avoiding procedures like voting or averaging enhance this tendency (Hall & Watson, 1971).

In view of the societal relevance of group decisions, good designs for effective group processes are wanted. Various techniques for the improvement of group decision making have been developed such as the Delphi Technique (Dalkey & Helmer, 1963), the Nominal Group Technique (Van de Ven & Delbecq, 1971), and the Group Consensus Method (Hall & Watson, 1971). However, there exists contradictory empirical evidence with respect to the effectiveness of these methods (Erffmeyer & Lane, 1984; Hegedus & Rassmussen, 1986; McGrath, 1984; Roth, 1994).

Psychological research has revealed a great impact of normative influence on opinion formation (Asch, 1951, 1956; Deutsch & Gerard, 1955; Sherif, 1935; Stasser, Kerr & Davis, 1989). Often, it is supposed that conformity has a negative influence on group processes. Also Delphi and Nominal Group Technique are promoted by the argument that they achieve to reduce normative pressure in interaction processes (Delbecq, Van de Ven & Gustafson, 1975). However, it is not yet well understood when and why normative influence has a negative impact on group performance. Some insight into this question is provided by the assumptions underlying formal models concerning social influence within groups such as Latane’s Social Impact Theory (1981), the Social Interaction Sequence Model (Stasser and Davis, 1981), the JUS Model by Hastie, Penrod and Pennington (1983), the Norm Information Distance Model by Crott and Werner (1991, 1994), or the Probabilistic Model of Opinion Change Including Distance by Crott, Werner and Hoffman (1996). According to these models, the power of attraction towards an opinion is a function of the number of its advocates (Stasser, Kerr, & Davis, 1989). If there is a large sub-
group that prefers an incorrect answer, there is a high probability of opinion changes in favour of a false answer. If task difficulty is defined by the percentage of correct individual answers prior to any group discussion, then conformity pressure may be negative, particularly in difficult tasks where few persons advocate the correct answer. Simply by a statistical effect it is likely that a large faction accidentally agrees on false answers in difficult tasks and that a large faction or even a majority will have an incorrect solution at the beginning of a group process.

For tasks with only two possible answers, a correct and a false one, this relationship can be derived from binomial distribution. Table 1 shows this connection using the example of a five-person group. In the case of an easy task with a probability of a correct individual pre-discussion answer of \( p(\text{correct}) = .80 \), the group will make a correct decision with a probability of \( p = .94 \) if the group agrees with the majority. This probability is considerably higher than the probability of a correct individual decision and can be higher than the probability that a single expert, who is more skilled than the group members themselves, solves the task correctly. In a difficult task where an individual’s probability to prefer the correct answer is \( p(\text{correct}) = .20 \), the probability of a correct majority is only \( p = .06 \). Thus, large factions tend to have a misleading influence.

Table 1: Probability of 3 or More Group Members Having the Correct Answer in a 5-Person Group

<table>
<thead>
<tr>
<th>Opinion distribution [correct; incorrect]</th>
<th>( p(\text{individual correct}) = .20 )</th>
<th>( p(\text{individual correct}) = .80 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>[5; 0]</td>
<td>0.0003</td>
<td>0.3277</td>
</tr>
<tr>
<td>[4; 1]</td>
<td>0.0064</td>
<td>0.4096</td>
</tr>
<tr>
<td>[3; 2]</td>
<td>0.0512</td>
<td>0.9421</td>
</tr>
<tr>
<td>[2; 3]</td>
<td>0.2048</td>
<td></td>
</tr>
<tr>
<td>[1; 4]</td>
<td>0.4096</td>
<td>0.0512</td>
</tr>
<tr>
<td>[0; 5]</td>
<td>0.3277</td>
<td>0.9421</td>
</tr>
</tbody>
</table>

Note. \( p(\text{individual correct}) \) = probability of an individual preferring the correct answer; \( \sum p([x; y]) \) = sum of the probabilities for the opinion distributions which do represent vs. do not represent a majority for the correct answer.

For tasks that have multiple possible answers an analogous relationship can be derived from multinomial distribution. In such tasks, the negative effect of conformity is especially harmful if a particular false answer is very suggestive (Hansmann, 2001). Judgmental tasks like those of the present experiment can also be represented as tasks with a finite number of intervals as possible answers. Here, a negative effect of conformity results if the judgements have a strong systematic bias, and if few persons make appropriate judgements. Accordingly, in analysing the task-difficulty of estimation tasks requiring quantitative judgements, in addition to aspects like structure and complexity, characteristics of the distribution of individual judgements have to be considered.
Crott and co-workers (Crott & Hansmann, subm.; Hansmann, 2001) have designed the Improvement of Normative Functioning and Output of Groups intervention (INFO) technique which firstly informs the group members on the differential effects of conformity depending on task difficulty as has been described above. The INFO technique secondly proposes to conduct a task difficulty evaluation individually before entering group discussion. Thirdly, the decision makers are instructed to cope with conformity pressure correspondingly when acknowledging the task difficulty level and the information on the differential effects of conformity pressure.

Crott & Hansmann (subm.) investigated the effect of the INFO technique on group performance in logical problems and knowledge questions (Hansmann, 2001). Group members in the INFO condition attained significantly more correct individual answers, even before the discussions began. At the end of the group discussions, the INFO groups yielded significantly more correct group decisions. In that study, both the control groups and the INFO intervention groups performed an individual task difficulty evaluation prior to the group discussion. In the experiment, Crott & Hansmann examined two conditions, which exclusively varied with respect to the providing of the information on the differential effects of conformity influences in easy versus difficult tasks. They could thus conclude that the positive effect of the INFO intervention on the quality of the outcomes was due to the information provided.

The present experiment will analyse whether the performance of a task difficulty evaluation itself improves group performance. Therefore, the two major components of the INFO technique, i.e. the advice on the differential effects of conformity pressure in easy versus difficult tasks and the evaluation of task difficulty, were separately manipulated and analysed in two experimental runs.
Method

Basic Experimental Design

The experiment was conducted in two separate experimental runs. In Run 1, those groups performing evaluations of task difficulty (experimental condition 1, EC1) were compared to control groups without performing such an evaluation (CC1). None of these groups received any information on the differential effects of conformity pressure depending on the task difficulty. In Run 2, groups that received the corresponding INFO advice and subsequently performed a task difficulty evaluation (EC2) were compared to groups that performed a task difficulty evaluation but did not receive any corresponding information (CC2). Table 2 shows the basic experimental design.

Table 2: Experimental Design

<table>
<thead>
<tr>
<th>Run 1</th>
<th>Experimental condition EC1</th>
<th>Control condition CC1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task difficulty evaluation</td>
<td>nGr = 8</td>
<td>nP = 28</td>
</tr>
<tr>
<td>Without task difficulty evaluation</td>
<td>nGr = 8</td>
<td>nP = 28</td>
</tr>
</tbody>
</table>

4 months of participation in the case study

<table>
<thead>
<tr>
<th>Run 2</th>
<th>Experimental condition EC2</th>
<th>Control condition CC2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task difficulty evaluation and INFO instruction</td>
<td>nGr = 6</td>
<td>nP = 23</td>
</tr>
<tr>
<td>Task difficulty evaluation without INFO instruction</td>
<td>nGr = 6</td>
<td>nP = 23</td>
</tr>
</tbody>
</table>

Note. Groups of three to four persons were working on seven tasks at each run of the experiment (nGr = number of groups; nP = number of persons). For Run 2 of the experiment new groups were formed out of the same pool of persons. In Run 1 there were 16 groups formed out of 56 persons. Due to organisational reasons, only 12 groups could be formed out of 46 persons in Run 2 of the field experiment.

Participants, Experimental Setting and Tasks

In the experiment, three- to four-person groups solved judgmental tasks concerning environmental planning. The participants of the experiment were fourth year male and female master students of Environmental and Natural Science at Eidgenössische Technische Hochschule Zurich (Swiss Federal Institute of Technology, ETH). These students participated in a project on the future of the Swiss Railways Company (SBB; Scholz, Bösch, Stauffacher, Oswald & Balmer, 2001). The project was a large scale embedded case study (see Scholz & Tietje, 2001). These case studies have a practical orientation and real impacts. They involve altogether more than 100 scientists, students, municipal authorities, citizens, and experts. The case study is a major part of the students’ curriculum. Each student is expected to work 320 hours on the project. Group work and group decision making is essential for the ETH-UNS case study (Crott,
Grotser, Hansmann, Mieg & Scholz, 1999; Mieg, 1996, 2000; Scholz, Bösch, Carlucci & Oswald, 1999; Scholz, Mieg & Weber, 1997). By combining practical application of knowledge, collaboration with experts, research, teaching and environmental problem solving, ETH UNS case studies establish a new type of instructional design (Gagne & Briggs, 1974) and are also denoted as transdisciplinary activities (Häberli, Scholz, Bill & Welti, 2000).

The judgmental tasks were related to the case study. Problems to be dealt with were on environmental or economical impact of rail traffic such as: What are the average costs (in Swiss Franks) to build a railway bridge with two railway lanes, with a height of 50 m and a length of 1 km in a valley (Land costs have to be ignored)?

The experiment consisted of two runs. Seven tasks were used in Run 1 of the experiment, and another set of seven tasks in the Run 2. The appropriate answers for these tasks were obtained from official sources such as e.g. the Union International de Chemin de Fer (UIC) and experts of the Swiss Railway Companies (SBB).

**Experimental Procedure**

Each experimental run took place on one single day, and in both runs of the experiment the group sessions took place in the same rooms. 18 four-person groups and 10 three-person groups participated in the overall experiment. The group size was balanced between the two conditions (EC1 vs. CC1 and EC2 vs. CC2, respectively). Run 1 was conducted shortly after the students had started the project work. Eight four-person groups and eight three-person groups were formed. In Run 1 the task difficulty evaluation was only performed in the experimental group EC1, whereas the control group CC1 performed no evaluation. Run 2 was conducted about four months later, close to the end of the case study.

In Run 2, a new set of estimation tasks was used because the same participants took part in both runs. Ten four-person groups and two three-person groups participated in Run 2. The procedure in the experimental condition (EC2) consisted of informing the participants about the differential effects of pressure towards conformity depending on the difficulty of a task and of instructing them to evaluate the task difficulty. In the control condition CC2, the participants also had to evaluate the task difficulty, but received no additional information. One reason why all groups had to perform a task difficulty evaluation in Run 2 was that some group members could have spontaneously performed a task difficulty evaluation on the basis of their experience in Run 1. In the statistical analyses, the groups of Run 1 and Run 2 have been considered independent as the participants were assigned randomly to the groups in both runs.

---

1 The basis of the analyses that are described in this article are exclusively groups. It can be plausibly assumed that an entire and functionally independent group process took place in the groups of each experimental run because the groups were newly and randomly formed in the second run. Consequently, the chance probability of a group composition in experimental run 2 resembling a group composition in Run 1 was extremely low. Similar assumptions of statistical independence have been common practice in the analysis of group processes (e.g. Kerr, 1981, 1992; Kerr & MacCoun, 1985).
The following features were kept constant in both conditions of both experimental runs:

- The (seven) judgmental tasks selected for the corresponding experimental run and the order of presentation of these tasks
- The time spent on each task (the experimenter advised and controlled the groups to work approximately seven minutes per task)
- After the groups were formed, the experimenter guided his or her group to a separate room and read aloud the instructions of the corresponding condition. The experimenters were not allowed to interfere with the group process besides reminding the participants to fill in their experimental sheets and to keep the time constraints.
- Control groups and experimental groups always worked simultaneously in separate rooms
- The group members received a post-experimental questionnaire covering aspects of the previous group processes such as personal involvement in the discussions, the effectiveness of the group processes and personal satisfaction with the group processes.

The experimental procedure in the four conditions, which took about one hour, were:

**CC1**: The control groups in Run 1 followed the basic experimental procedure. After a task was presented, the participants worked individually on the task until all group members had reached an individual pre-discussion judgement. The group then discussed the task until a joint group judgement was obtained. After the group judgement, each group member once more noted his individual judgement for the task. Except for the time control and the phase of individual work prior to the discussions, CC1 was a free interaction process. The groups worked on the seven tasks following the same procedure for each task. Thereafter, the participants received the post-experimental questionnaire.

**EC1**: In Run 1, EC1 followed the experimental procedure of CC1. In addition, the participants had to individually evaluate the difficulty level of each task prior to the discussion as the first step in the individual working phase. The group members rated the task difficulty level on a scale reaching from 1 = very easy to 5 = very difficult. For this judgement they were supposed to consider the structure and complexity of a task, as well as the range and distribution of individual judgements. To ensure that the latter aspect would be taken into account in the evaluation, the participants had to denote an estimate of the within-group range of the individual pre-discussion judgements (prediction of the highest, and lowest within group judgement). After all group members had denoted their individual task difficulty rating and their individual pre-discussion judgement for the task itself, the group discussion started.

At the end of this discussion the groups had to decide on a common group judgement for the task-difficulty level, and subsequently for the task itself.
CC2: In Run 2, the basic experimental procedure resembled that of Run 1. However, both control groups and experimental groups performed an individual difficulty evaluation of each task prior to the corresponding group discussion, and made a collective judgement on the task difficulty level before the collective judgement for the task itself was formed. Hence, the procedure in the control condition CC2 was similar to that of EC1, the experimental condition of Run1.

EC2: In addition to the experimental procedure of CC2, before the presentation of the first task, the participants in EC2 received information on the differential effects of majority influences on the quality of group decisions in easy versus difficult tasks, revealing the reasoning described in the introduction. Accordingly, the necessity to evaluate the difficulty of a task by analysing its characteristics was pointed out. The information generally encouraged the group members to concentrate on the logic of argumentation. Moreover, if they considered a task to be difficult, they were instructed to listen especially carefully to the argumentation of minority factions. After the experimenter had read this information aloud, a poster containing the basic principle of the differential effects of group conformity and the corresponding INFO advice was presented to the subjects (see Figure 1).

![Diagram](Figure 1: Poster summarising the basic principle of the INFO advice)
Results

The analyses focus on the effects of task difficulty evaluation, which was experimentally varied in Run 1, and of the awareness of the differential effects of conformity pressure (Run 2). We firstly analyse the accuracy of the individual and group judgements. Subsequently, the answers of the participants in the post-experimental questionnaires will be considered.

Judgement Accuracy

The data of the answers to the 14 tasks were transformed using the T-Transformation (McCall, 1939) to construct a quality index which allows for comparisons of all tasks. For each task the individual pre-discussion judgements, the group judgements, the individual post-discussion judgements, and the expert judgement were jointly rank-ordered. These ranks were then transformed to T values. After this transformation, the data were normally distributed with a mean of $M = 50$ and standard deviation of $s = 10$. Thus, the absolute deviation of each individual and group judgement’s T-value from the T-value of the expert judgement was taken as (in)accuracy measure.

An ANOVA on the absolute deviation of the (T-transformed) individual pre-discussion judgements, of the group judgements, and of the individual post-discussion judgements from the expert solution was calculated. Each of these three measures was then averaged over the seven tasks a group had to solve. The resulting mean deviation (D) from the expert judgement served as the dependent variable for the analysis of the accuracy of the judgements. The ANOVA was performed with the independent variables experimental run (2 levels: Run 1 vs. Run 2), and condition (2 levels: experimental vs. control), and the repeated measurement variable assessment (3 levels: individual pre-discussion judgement, group judgement, individual post-discussion judgement).

The ANOVA provided a significant main effect of the repeated measurement variable assessment, $F(1.88, 45.01) = 60.94$, $p < .001$. Table 3 shows the mean deviation from the expert judgement for all three judgements, separated for the two experimental runs and the corresponding experimental and control condition. Considering all groups, the mean deviation from the expert judgements was $M_{D_{i1}} = 14.35$ for the individual pre-discussion judgements, $M_{D_{Grp}} = 12.31$ for the group judgements, and $M_{D_{i2}} = 12.93$ for the individual post-discussion estimates. The corresponding simple contrasts revealed that the individual pre-discussion judgements were significantly less accurate than the individual post-discussion judgements, $F(1, 24) = 40.48$, $p < .001$, and than the group estimates, $F(1, 24) = 102.27$, $p < .001$. These results are in accordance with previous research, which has shown that group judgements are usually more accurate than individual pre-discussion estimates (see Gigone & Hastie, 1997; Hastie, 1986). Additionally, in our study, the accuracy of the group judgements proved to be significantly higher than the accuracy of the individual post-discussion estimates, $F(1, 24) = 22.57$, $p < .001$.

---

2 Because of a significant violation of the sphericity assumption (Mauchly Sphericity Test, $p < .05$), the degrees of freedom were corrected according to the Huynh-Feld-Epsilon ($\eta = .938$).
Table 3: Mean Absolute Deviation of the Individual Pre-discussion Judgements. Group Judgements and Individual Post-discussion Judgements From the Solutions. Separated for the two Conditions in Each Experimental Run

<table>
<thead>
<tr>
<th>Run 1</th>
<th>Individual pre-discussion judgement</th>
<th>Group judgement</th>
<th>Individual post-discussion judgement</th>
<th>Overall Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC1</td>
<td>16.22</td>
<td>16.45</td>
<td>16.11</td>
<td>16.26</td>
</tr>
<tr>
<td>EC1</td>
<td>14.64</td>
<td>13.64</td>
<td>14.07</td>
<td>14.12</td>
</tr>
<tr>
<td>Difference</td>
<td>1.58</td>
<td>2.81*</td>
<td>2.04*</td>
<td>2.14*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Run 2</th>
<th>Individual pre-discussion judgement</th>
<th>Group judgement</th>
<th>Individual post-discussion judgement</th>
<th>Overall Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC2</td>
<td>13.64</td>
<td>8.84</td>
<td>10.38</td>
<td>10.95</td>
</tr>
<tr>
<td>EC2</td>
<td>12.19</td>
<td>8.48</td>
<td>9.72</td>
<td>10.13</td>
</tr>
<tr>
<td>Difference</td>
<td>1.45</td>
<td>0.36</td>
<td>0.66</td>
<td>0.82</td>
</tr>
</tbody>
</table>

Note. *p < .05. **p < .01. ***p < .001. two-tailed.

The main effect of condition was only marginally significant, $F(1, 24) = 3.45, p < .1$. If all three judgements were considered jointly, there was a higher accuracy of the judgements in the two experimental conditions EC1 and EC2 with a mean deviation of $MD_{EC} = 12.04$ from the expert judgements than in the two control conditions CC1 and CC2 with a mean deviation of $MD_{CC} = 13.98$.

Simple contrasts for the effects of experimental condition showed that in Run 1 the mean judgement accuracy was significantly higher in the experimental condition than in the control condition ($MD_{EC1} = 14.12, MD_{CC1} = 16.26, F(1, 25) = 4.29, p < .05$). In Run 2, there was no significant difference between the accuracy of judgements in the experimental condition and control condition averaged over the three assessments ($MD_{EC2} = 10.13, MD_{CC2} = 10.95, F(1, 25) = 0.41, p = .53$).

There was a significant 3-way interaction between experimental run, condition and assessment, $F(1.88, 45.01) = 3.68, p < .05$. As Table 3 shows, the task difficulty evaluation in experimental Run 1 had a positive effect on the accuracy of the group judgements which provided a significant contrast ($MD_{EC1} = 13.64$ vs. $MD_{CC1} = 16.45, F(1, 25) = 6.69, p < .05$). In addition, there was a marginal significant effect of task difficulty evaluation on the accuracy of the individual post-discussion judgements, $F(1, 25) = 3.36, p < .1$. The mean post-discussion judgements were more accurate ($MD_{EC1} = 14.07$) in the experimental condition than in the control condition ($MD_{CC2} = 16.11$). There were no significant contrasts of condition in experimental Run 2.

There was also a significant main effect of the variable experimental run, $F(1, 24) = 33.88, p < .001$. The mean deviation from the expert judgements was higher in Run 1 than in Run 2. This result seems to indicate that the students gained judgmental expertise during case study participation (Mieg, 2001). This conclusion has to be considered with reservation, as the tasks of Run 1 were different from those of Run 2.
Results of the Post-experimental Questionnaire

All participants received a post-experimental questionnaire. Questions had to be answered on a scale of one (1 = very negative) to five (5 = very positive) with three being the neutral point. In both runs and all conditions, the participants considered the tasks as rather clearly formulated (MQ1 = 3.45), were engaged in the group discussions (MQ2 = 3.26), were highly satisfied with the emotional climate in their group (MQ3 = 4.54), and judged the effectiveness of the group discussions positively (MQ4 = 3.68). There were no significant main effects or interactions. In addition to the general questions, which were similar in both experimental runs, at the end of Run 1 the participants were asked if they believed that the performance of a task difficulty evaluation, which analyses the complexity of a task, could possibly improve group discussions concerning judgmental tasks. The mean answers of the students on the five point rating scale was MQ5,Run1 = 3.5. In tendency, this suggests that it appears helpful to students to evaluate the difficulty and complexity of a task as part of a judgmental group process. In the experimental condition (MQ5,EC1 = 3.65), where the groups had actually performed such a difficulty evaluation, the answer to this question tended to be slightly more optimistic than in the control groups (MQ5,CC1 = 3.39). The corresponding difference was however not significant.

As a manipulation check of the INFO technique, the participants of Run 2 were asked whether conformity pressure would have a positive effect on the quality of group decisions in easy tasks rather than in difficult tasks. In EC2, 87.5% of the participants answered this question in accordance with the INFO rationale, as compared to only 36.1% in the control condition, $t(10) = 2.95, p < .05^3$.

---

3 This t-test was accomplished on the basis of group means. For the calculation of these group means, individual answers in accordance with the INFO rationale were assigned the value 1 (= true), and answers inconsistent with the INFO rationale were assigned the value 0 (= false).
Discussion

The experiment investigated the effects of a task difficulty evaluation and of knowing about negative and positive effects of group conformity in easy versus difficult judgmental tasks. In order to establish ecological validity of real world planning and decision making, the experiment was conducted in the context of environmental planning. Altogether, 28 groups have been investigated in the two runs of the overall experiment. In Run 1, groups evaluating the task difficulty were compared to control groups that did not evaluate the task difficulty. In Run 2, all groups performed a task difficulty evaluation, but only the experimental groups were additionally informed about the differential effects of group conformity.

In both runs, the individual pre-discussion judgements, the group judgements and the individual post-discussion judgements in the experimental condition tended to be more accurate than those in the corresponding control condition. The overall effect was marginally significant.

Concerning Run 1 the analysis revealed a significant positive effect of performing the task difficulty evaluation on the accuracy of the group judgements, and a marginally significant positive effect of this evaluation on the accuracy of the individual post-discussion judgements. The positive effect of evaluating the task difficulty on the group process in Run 1 resulted even though the participants were not informed of the differential effects of group conformity pressure. Presumably, a preceding task difficulty evaluation enhances a vigilant and careful processing of the tasks and effectively counteracts judgmental biases. The latter might be interpreted in terms of the following mechanisms:

The analysis of the difficulty of a task as performed in the present study made it necessary to consider the range of plausible answers, which possesses an upper and a lower bound. These two boundaries obviously represent two anchors for the subsequent judgements. This split of anchoring might counteract biases in judgement, which can be caused by the effects of single anchors (Kahneman, Slovic & Tversky, 1982). Moreover, the pre-discussion prediction of the range of individual judgements within the group might be connected to the generation of arguments for corresponding high and low judgements. As Burnstein and Vinokur (1975) have shown, "... knowledge about others’ preferences can be a sufficient condition for revising preferences to the extent that it leads one to think of arguments in support of the courses of action others have selected, arguments which previously had not come to mind (1975, p. 423)." It might well be the case that the prediction of others’ preferences and judgements, respectively, has similar effects. In this case, the generation of arguments for both, high and low judgements might a) generally enhance the informational basis of the judgements, and might b) serve to balance the search for arguments favouring high versus low judgements.
In **Run 2**, only the experimental groups were informed about the differential impact of conformity in easy versus difficult tasks, whereas the pre-discussion task difficulty evaluation was conducted in all groups. Here, as compared with the control groups, the accuracy of the judgements in the experimental groups was slightly, but not significantly higher in all three assessments. A potential explanation of this would be that the subjects in both conditions of **Run 2** shared the common knowledge about the differential effects of group conformity. However, the answers to the post-experimental questionnaire are in contradiction with this interpretation. We rather assume that the non-significant effects could be due to a ceiling effect, as the group judgements in **Run 2** showed by far the lowest deviation from the appropriate values (see Table 3).

The final question of **Run 2** asked the participants if conformity pressure would tend to have a positive effect in easy tasks rather than in difficult tasks. Most of the participants in the control condition responded that effects of conformity pressure tend to be helpful in difficult tasks and not in easy tasks. Accordingly, the INFO advice seem to contain information which is counterintuitive. This might be related to the need for social orientation, which – according to Festinger (1954) – should be stronger in difficult tasks as compared to easy tasks where the informational cues are clearer. The knowledge of the differential group conformity principle might thus motivate group members to differentially deal with conformity pressures during group discussions, and to perform a corresponding analysis of task difficulty.
References


Muss man den Formalismus beherrschen, um die Formalisten zu schlagen? [Environmental research between formal science and comprehension: is command of the formalism necessary for beating the formalists?] In A. Daschkeit & W. Schröder (Eds.), Umweltfor- schung zwischen Formalwissenschaft und Ver- ständnis. [Environmental research thought laterally: perspectives on integrating environmental research and teaching] (pp. 309–328). Berlin: Springer.


Macht man die Formalisten zu Verrückten, um die Verrückten zu schlagen? [Environmental research thought laterally: perspectives on integrating environmental research and teaching] (pp. 309–328). Berlin: Springer.


