



## **Physical or Virtual Presence of the Experimenter: Psychological Online-Experiments in Different Settings**

Heike Ollesch, Edgar Heineken, Frank P. Schulte

*Universität Duisburg-Essen*

---

**Abstract:** Internet-based experiments are characterized by the physical absence of an experimenter and the difficulty to provide a standardized experimental arrangement. How does this affect internal validity and precision? Three experiments on different psychological questions, following studies by Heider and Simmel (1944), Morris and Stevens (1974), and Moyer (1973), were conducted in the virtual laboratory Lab.OR. To different degrees each of the selected experiments demanded commitment of participants and features measurements, which are usually considered susceptible to external disturbances. Each experiment was conducted as an Internet-based experiment and, using WLAN-notebooks, additionally in two other settings (laboratory rooms / public rest areas on campus), so that the degree of standardization of the experimental arrangement was varied. While the experimenter was virtually present in the Internet-based experiments, the kind of physical presence of the experimenter was varied in the laboratory and campus settings: either he was actively involved and gave oral instructions, or he just attended the experiment, but did not play an active part while all information was provided on the computer's screen.

Results show that all experiments provide equally valid and precise results, independent of setting or mode of presence of the experimenter. Taking into account the different psychological variables studied in the different experiments, some interesting effects were found: Participants' performance was affected by the kind of presence of the experimenter, when they had to write an elaborate report, but not when response latencies or recall rates were obtained. Lack of standardization of the environment did only affect response rates: performance in laboratory and Internet settings was superior to campus setting.

Implications for experimental online research and the potential of using the virtual laboratory Lab.OR in teaching experimental psychology to point out methodological specifics of online-experiments conducted in different settings are discussed.

*Keywords:* Internet-based experiments, laboratory experiments, web research methods, comparability, social presence

---

### **Methodological characteristics of Internet-based experiments**

Technical innovation has often brought about new possibilities for experimental research in psychology (Traxel, Gundlach & Zschuppe, 1986). Since the 1970s, specialized and expensive equipment for psychological research has been widely replaced by computers, which can support standardized, timely controlled presentation of stimuli, and provide objective, reliable and precise measurements of human behavior (Hopmann & Galley, 1986). Since the 1990s, carrying out experiments on the Internet gains interest in the field of psychological research (cf. Azar, 2000b; Banaji, Bruckman, Cohen & Couper, 2004; Birnbaum, 2000b, 2004a; Dillman, 2000;

---

Direct all correspondence to Heike Ollesch, Abteilung für Kognitionspsychologie, Institut für Kognition und Kommunikation, Universität Duisburg-Essen, Campus Duisburg, Forsthausweg 2, 47048 Duisburg, Germany. E-Mail: heike.ollesch@uni-duisburg-essen.de, Phone.: +49 203 / 379 2251, Fax: +49 203 / 379 1846.

Krantz & Dalal, 2000; Kraut et al., 2004; Musch & Reips, 2000; Reips, 2000, 2002). Internet-based experimenting promises considerable economical and methodological benefits: Among other advantages, it saves costs for experimenters, equipment and lab rooms. Large and heterogeneous samples can effortlessly be recruited, even special populations can easily be accessed, and objectivity of data collection is increased.

Current debates deal with methodological characteristics of Internet-based experiments in contrast to the traditional laboratory setting, focusing on potential benefits and limitations and on questions concerning validity or precision of experiments. It is discussed, how Internet-based experiments can meet the standards for experimental research in psychology compared to laboratory experiments. Methodological issues of experimental control and ethics in research involving human participants are considered lively (Azar, 2000a; Gosling, Vazire, Srivastava & John, 2004; Hewson, Laurent & Vogel, 1996; Krantz & Dalal, 2000; Kraut et al., 2004; Nosek, Banaji & Greenwald, 2002; Reips, 2000).

This paper focuses on two major specifics of Internet-based experiments: The physical absence of an experimenter and the difficulty to provide a standardized experimental arrangement. Therefore issues of the social psychology of experimenting and of control over the experimental arrangement and their effects on data quality have to be considered (Kraut et al., 2004; Nosek et al., 2002).

### **Social psychology of Internet-based experiments**

When experiments are conducted in a laboratory setting, participants are usually instructed and debriefed by an experimenter. The experimenter accompanies the participants during the course of an experiment, gives them orientation and informs them about their task. He can motivate them to follow instructions, to focus on the experiment and to make the required effort. Research on artifacts in psychological experiments focuses on unwanted effects deriving from this face-to-face-interaction between participants and experimenters (Bungard, 1980). Different artifacts have been reported, which result from expectations and situational beliefs of experimenters (e.g. Rosenthal, 1966) or participants (e.g. Orne, 1962).

In the Internet setting, the whole course of experiments is completely automated and standardized, there is no face-to-face-communication between experimenters and participants. Information and motivation of participants is obtained exclusively by written information. Therefore, some researchers hope that artifacts originating from the interaction between experimenter and participants do not occur in Internet-based experiments (Birbaum, 2004b; Reips, 2000, 2002). Despite the physical absence of an experimenter the elimination of these artifacts can not be guaranteed: a "virtual" experimenter shows up in instructions and participants' tasks in the experiment. This information may as well induce a set of "demand characteristics" (Orne, 1962). Participants will behave according to their idea of the experiment's objectives – if this idea corresponds to the hypothesis, it may cause experimenter effects; if this idea is a misconception, it may cause participants to produce artifacts based on their assumptions. Even new problems may arise: Lack of social control in online-experiments may reduce participants' commitment. This may result in participants ignoring instructions, cheating, or in participating more than once.

#### *Standardization of Experimental Arrangements: Problems of Control in Internet-based Experiments*

In the laboratory setting, experimenters care about ideal conditions to make effects of the experimental treatment clearly observable. This includes standardization of participants' environment and minimization of uncontrollable influences that can impair the effects of experimental treatment. If experiments are conducted on the Internet, it is assumed that higher error variances result from lack of the control of unsystematic error variables, heterogeneity of settings during participation and different computer equipment of participants. This is supposed to be easily compensated by larger numbers of participants and resulting stronger statistical power (see Reips, 2000, 2002).

### **Comparing Internet-based and laboratory experiments: Problems and solutions**

Studies comparing data collected in Internet-based experiments with data gathered in laboratory settings or with expectations derived from theoretical models show high consistence of findings in different settings (e.g. Birbaum, 2000a; Krantz, Ballard, & Scher, 1997; McGraw, Tew, & Williams, 2000; Metzger, Kristof, & Yoest, 2003; Senior, Phillips, & Barnes, 1999). Krantz and Dalal (2000) reported some studies, where results of Internet-based experiments match with results of corresponding experiments conducted in the laboratory („convergent validity“) or to theoretically deduced expectations („construct validity“). They appraise these findings as "surprising" similarity of the results of online-experiments to the expectations drawn from the

traditional psychological research in the laboratory (Krantz & Dalal, 2000, p. 56), but hesitate to speak of a general equivalence of psychological research in online and laboratory settings. Yet these comparative studies have covered only a small sample of the wide range of experimental research in psychology: Krantz and Dalal (2000, p. 38) state: „The Web allows basically any type of research that is possible using computers as equipment and humans as subjects, [...] the validity of Web-based research may vary for the different psychological variables studied.“ Further studies have to be carried out to explore the respective advantages and constraints of experimenting in the lab and on the Web, taking into account different tasks and measurements.

Because laboratory and Internet settings differ in many ways, the causes for resemblances and differences in the results usually remain obscure. It is necessary to take a closer look, how the specifics of Internet-based experimenting affect validity and precision. A current development of computer technology supports mobility: Using the Internet is no longer restricted to stationary computers at home or at the workplace. WLAN-enabled notebooks or other communication devices and access points allow connection to the Internet at more and more public locations. This provides new research opportunities for psychologists: Experimenters can approach participants and conduct Internet-based experiments at places outside the psychological laboratory (Ollesch, Heineken, & Schulte, 2003; Ollesch, Schulte, & Heineken, 2004). Using WLAN-devices the experiment will “come to the participant”, and the experimenter can choose to “come to the participant” as well and establish a face-to-face-contact, which is not possible in “classic” Internet-based experiments.

Considering the mode of presence of the experimenter and the potential to provide a standardized experimental arrangement and shield participants from external disturbances as critical variables, WLAN-enabled notebooks can be used to create experimental arrangements which resemble more or less common Internet-based or laboratory settings: The experimenter can define time and place for conducting an experiment – this may be a laboratory room, where the experimenter has full control over the experimental arrangement as well as a public location, where he can not provide a standardized environment and shield the participants from external interferences. Mode of presence of the experimenter can be varied as well: Internet-based experiments always comprise a “virtual” experimenter and there is definitely no face-to-face-communication between experimenter and participants. When experiments are conducted in the laboratory or in public locations, an experimenter will be involved, but giving oral instructions and debriefings is dispensable. It is also possible to give all information necessary in written form like in an Internet-based experiment.

When experiments are conducted in different settings (Internet, laboratory, public locations) and with different mode of presence of an experimenter (“virtual experimenter”, physical presence of an experimenter / written instructions; physical presence of an experimenter / oral instructions), the following questions can be investigated:

- Internal validity of experiments: Do expected effects show in all different settings?
- Precision of experiments: Are there differences in variability of measures in different settings?
- Does commitment of participants (efforts made, obedience to instructions) depend on the different settings or on the mode of presence of the experimenter?
- What kind of impact does face-to-face-communication between experimenter and participants have on the experiment’s results?






### **The impact of mode of presence of the experimenter and degree of standardization of experimental arrangements on validity and precision in three selected experiments**

Three experiments on different psychological questions were selected to be conducted under the different conditions described above. The experiments follow studies by Heider and Simmel (1944), Morris and Stevens (1974), and Moyer (1973). The latter examined whether mental comparisons follow the same principles as comparisons of perceived objects. Morris and Stevens (1974) compared the efficacy of different strategies for memorizing lists of words: repetition, developing mental images for single words, and developing complex mental images for a sequence of three words. They doubted the general superiority of strategies involving mental imagery against simple repetition. They could prove their assumption, that superior recall rates can only be achieved by strategies that involve imagery and organization of items at the same time. Heider and Simmel (1944) investigated how participants perceive the events in an animated movie with simple geometrical objects. Participants' reports were recorded and analyzed, regarding subjective organization of the perceived events and attribution of human traits to the geometrical objects. To different degrees each of the selected experiments demands commitment of participants and features measurements, which are usually considered susceptible to external disturbances.

The three experiments were implemented in the virtual laboratory Lab.OR (Heineken, Ollesch & Schulte, 2003), and conducted as Internet-based experiment, and – using WLAN-enabled notebooks – in laboratory rooms and public areas on the campus. Presentation of stimuli and mode of participants’ reactions were standardized, as the

Lab.OR-environment served in all settings for delivering stimuli and registering data on computers connected to the Internet. Experiments were designed carefully in consideration of coherence, simplicity, involvement and consistency (Aronson, Ellworth, Smith & Gonzales, 1990). A list of requirements was presented to participants at the beginning of each experiment: They were asked to close other programs, reserve enough time to complete the experiment without interruptions, and avoid external disturbance. While in the Internet setting participants themselves had to arrange for suitable conditions, in laboratory and public location setting experimenters were responsible for meeting those requirements. Information on the experiment, instructions and debriefing were designed carefully to fit the needs of participants even if only a virtual experimenter was involved in the Internet setting. Details on the experimental design of the three studies will be given in the next section. The study materials can be accessed from <http://heineken.uni-duisburg.de/onlineexp1/>.

Table 1  
*Synopsis on conditions for conducting the experiments*

Setting	Mode of presence of an experimenter		
	<i>actively involved (oral instructions, exploration and debriefing)</i>	<i>attendant (written instructions, exploration and debriefing)</i>	<i>"virtual" experimenter</i>
<i>Laboratory-setting</i>			
<i>Public-area-setting</i>			
<i>Internet-setting</i>			

Each experiment was conducted in the following settings to vary the standardization of experimental arrangements (see Table 1): via the Internet, on different public locations on campus (public rest areas inside and outside campus buildings, provided with banks and tables) via WLAN-connected mobile notebooks provided by the experimenters, and in laboratory rooms of the institute. Participants of the Internet-based experiments were recruited by the student experimenters by postings in different mailing lists. In "public-area"- and laboratory settings experimenters recruited participants by addressing them on campus and asking them to take part in an experiment. While in the Internet setting an experimenter appeared only "virtually", e.g. in written instructions or debriefing, mode of presence of the experimenter was varied in laboratory and "public area" setting: A face-to-face contact between experimenter and participants was established during the recruitment. During the course of an experiment conditions were twofold: Under one condition the experimenter was actively involved in the experiment and maintained face-to-face communication with the participants by giving oral instructions, exploration and debriefing. Under the other condition the experimenter attended the experiment, but did not play an active part; participants got all instructions and information they needed solely on the computer's screen.

As it is neither practicable to prevent participants fully from face-to-face contact with the experimenter in laboratory and "public area" setting, nor to establish a face-to-face communication between experimenter and

participants in a “classic” Internet-based experiment, a full factorial design could not be implemented. In the reported exploratory studies different aspects of “classic” laboratory and Internet-based experiments can only be approached by varying settings and mode of presence of the experimenter.

**Method and results of the experiments**

*Experiment 1: Mental Comparisons (Moyer, 1973)*

It was investigated whether mental comparisons follow the same principles as comparisons of perceived objects. In the implementation of this experiment in the virtual laboratory, participants were asked to indicate as quickly as possible which one of the two animals named on the screen was taller. Ordinal distance between size of the animals was varied as a within-subjects-factor and response latencies were registered. According to Moyers’ findings, response latencies should decrease when size difference between the compared animals increases.

The comparison of the data gathered under different conditions shows whether the measurement of response latencies – which is usually considered sensitive to external disturbances – is actually affected by the different standardization of experimental arrangements in the different settings, and whether participants’ commitment (efforts made, showing in the overall response latencies) depends on the mode of presence of the experimenter.

In total, fifty persons participated in this study, ten in each of the different conditions for conducting the experiment. Under all conditions response latencies increased with decreasing size differences between the compared animals (see Table 2). Mean response latencies for pairs of animals with little size difference were 1235 ms (*SD* = 346), for pairs with medium size difference 1112 ms (*SD* = 219) and for pairs with large size difference 1054 ms (*SD* = 216). Standard deviations (displayed in Table 2) show no systematic effect of conditions on precision of measurements. In an analysis of variance with the within-subjects-factor „size difference between compared animals“ the anticipated effect proved itself as a highly significant, strong effect under all conditions ( $\eta^2$  between 0.52 and 0.69; see Table 2).

Table 2  
*Results of the experiment following Moyer (1973): response latencies for mental comparisons of pairs of animals with different size differences under different conditions*

Setting	Mode of presence of the experimenter	N	Size difference between compared animals (mean response latencies and SD)			F	$\eta^2$
			Small	medium	large		
Internet	"virtual experimenter"	10	1403 409	1190 253	1145 267	$F(2, 18) = 13.22,$ $p < .01.$	0.60
Public area	experimenter present, written instructions	10	1287 360	1099 267	1071 267	$F(2, 18) = 10.47,$ $p < .01.$	0.54
Public area	experimenter involved, oral instructions	10	1214 253	1032 116	976 80	$F(2, 18) = 9.91,$ $p < .01.$	0.52
Lab	experimenter present, written instructions	10	1513 393	1201 184	1047 191	$F(2, 18) = 13.12,$ $p < .01.$	0.59
Lab	experimenter involved, oral instructions	10	1208 241	1037 223	1030 225	$F(2, 18) = 19.83,$ $p < .01.$	0.69

When in addition to the within-factor „size difference between compared animals“ the factors „setting“ or „mode of presence of the experimenter“ were included in the analysis, no effects besides the reported effect of the within-factor „size difference between compared animals“ were found.

In this experiment, the anticipated effect was found, even with small sample sizes, independent of the different conditions under which the experiment was conducted. Although, due to small sample size, statistical power was relatively low, the findings indicate that the different standardization of experimental arrangements in the different settings and the different mode of presence of the experimenter did not affect the experiment’s outcomes: Precision of measurements was not affected by different standardization of environmental conditions in the different settings. Commitment of participants was unaffected by the mode of presence of the

experimenter – efforts of participants, showing in their response latencies, did not depend on the experimenter being actively involved, or only physically or „virtually“ present.

*Experiment 2: Mnemotechniques in Serial Learning (Morris & Stevens, 1974)*

In the implementation of this experiment in the virtual laboratory, each participant was asked to follow three different mnemotechnical instructions when trying to memorize a list of 24 concrete nouns: memorizing, developing mental images for single items, and developing complex mental images for three items, which followed another in the list. Sequence of mnemotechnical instructions was counterbalanced over participants. Recall rates were recorded.

A comparison of results from the different conditions, under which this experiment was conducted, shows whether recall rates are affected by different standardization of experimental arrangements in the different settings, and whether the mode of presence of the experimenter influences participants' efforts in memorizing the presented items, and their commitment in following the mnemotechnical instructions even in the physical absence of an experimenter.

Fifty-two persons participated in the study, 16 in the Internet setting and 9 participants in each of the other four conditions. Under all conditions the best recall rates were obtained, if participants were asked to include three following items in one mental image. The mean recall rates under the different conditions are presented in Table 3. When participants were instructed to memorize items by repetition, they recalled at an average 10.29 items ( $SD = 4.20$ ). When they were instructed to form a mental image of each single item, the mean recall rate is 11.60 items ( $SD = 3.19$ ). When they were asked to combine three following items in one mental image, they recalled 16.58 items ( $SD = 4.75$ ). Except for the “public area” setting, when the experimenter was not involved and only written instructions were given, the anticipated effect was significant. There is no systematic impact of conditions during the course of the experiment on precision of measurement, as shown by the standard deviations presented in Table 3.

Table 3  
Results of the experiment following Morris and Stevens (1974): recall rates with different mnemotechnical instructions under different conditions

			<b>Mnemotechnical instruction (mean recall rates and SD)</b>				
<b>Setting</b>	<b>Mode of presence of the experimenter</b>	<b>N</b>	<b>Repeating</b>	<b>Imagining single items</b>	<b>Imagining three following items</b>	<b>F</b>	<b><math>\eta^2</math></b>
Internet	"virtual experimenter"	16	11.88 5.57	11.87 3.07	17.38 3.98	$F(2, 30) = 14.21,$ $p < .01.$	0.49
Public area	experimenter present, written instructions	9	9.11 3.22	10.33 2.69	13.11 4.01	$F(2, 16) = 2.59,$ n.s.	0.24
Public area	experimenter involved, oral instructions	9	8.11 2.42	11.78 3.70	15.22 5.09	$F(2, 16) = 12.97,$ $p < .01.$	0.62
Lab	experimenter present, written instructions	9	10.89 3.18	11.89 3.82	17.11 5.42	$F(2, 16) = 6.58,$ $p < .01.$	0.45
Lab	experimenter involved, oral instructions	9	10.22 3.93	11.89 3.06	19.44 4.00	$F(2, 16) = 16.21,$ $p < .01$	0.67

An analysis of variance with the within-subjects-factor "mnemotechnical instruction" and the between subjects-factor "setting" showed – besides the expected effect of the factor "mnemotechnical instruction",  $F(2, 98) = 45.64, p < .01; \eta^2 = 0.48$  – another significant effect of the factor "setting",  $F(2, 49) = 4.48, p < .05; \eta^2 = 0.15$ . When the experiment was conducted in public areas on the campus less items were recalled correctly ( $M = 11.27; SD = 2.30$ ) than in the Internet- ( $M = 13.71; SD = 3.35$ ) or in the laboratory setting ( $M = 13.57; SD = 2.44$ ).

In an analysis of variance with the within-subjects-factor "mnemotechnical instruction" and the between-subjects-factor "mode of presence of the experimenter", only the expected effect of the factor "mnemotechnical instruction" showed as significant,  $F(2, 98) = 45.34, p < .01; \eta^2 = 0.48$ ; the factor "mode of presence of the

experimenter" did not have a significant impact on recall rates. When the experiment was conducted in the Internet setting with a virtual experimenter averagely 13.71 items were recalled correctly ( $SD = 0.72$ ). When conducted in laboratory and "public area" setting mean recall rate was 12.07 ( $SD = 0.68$ ) when instructions were given in written form and 12.78 ( $SD = 0.68$ ) when the experimenter was actively involved in the experiment.

Independent from the conditions during the course of the experiment, internal validity was given - even if the anticipated effect was not significant in one setting, results showed the anticipated tendency. Participants' lower performance in the "public area" setting may be due to lack of standardization of environmental conditions: participants' concentration may be impaired by external disturbances. Despite lack of standardization of the environment, no impairment of performance was observable in Internet setting compared to lab setting. Presumably participants arranged for themselves undisturbed conditions when participating in the experiment.

Mode of presence of the experimenter did not affect commitment. Despite the relatively low statistical power, results indicate that this factor did not affect the experiment's outcomes: Neither did participants make stronger efforts in memorizing items in laboratory and "public area" setting, when an experimenter was actually attending the experiment or was actively involved, compared to the Internet setting with a "virtual experimenter", nor did they take advantage of the physical absence of an experimenter in the Internet setting and cheated to improve their recall rates. Participants followed the different mnemotechnical instructions no matter if the experimenter was physically or just virtually attendant, nor if instructions were given orally by the experimenter or in written text provided by the computer.

### *Experiment 3: Phenomenal Causality and Social Perception (Heider & Simmel, 1944)*

In this experiment, it was investigated how participants perceive the events in an animated movie in which simple geometrical shapes were moving across the screen. Participants were asked to write a short report of the observed movie. Heider and Simmel (1944) reported that participants organized the perceived events in a way that implies a coherent and meaningful structure by interpreting the movements of the geometrical shapes as actions in a social setting and attributing motives to the shapes. Therefore, it is assumed that reports on the movie include anthropomorphisms regarding the geometrical shapes and attributions of motives of the shapes for their "social behavior".

For the implementation of the experiment in the virtual laboratory, two animated movies were created. One of them was similar to the original movie by Heider and Simmel (1944), the other inverted the size-relations of the geometrical objects. Each movie was presented to one half of the participants. After watching the movie two times, participants were asked to write a short report on the events they had seen in the movie, and were asked some questions about the actions of the geometrical objects. It was analyzed how the different size-relations of the objects affected the attribution of motives to the objects.

Altogether 105 persons participated in the experiment, between 19 and 23 under each of the different conditions (see Table 4). Differences in the attribution of motives to the geometrical objects between the two movies were found as expected; this issue is not to be addressed further in this paper. To compare the data collected under different conditions during the course of the experiment, it was analyzed whether the conditions affect the written reports regarding the following questions: Do participants report anthropomorphisms of the objects under all conditions? Does the condition affect the grade of elaboration of the reports?

Each report was examined by an expert rater, whether anthropomorphisms could be found (e.g. descriptions of the shapes' movements as intentional actions, like "pursuing", "running away", "hiding", "teasing", attribution of roles to the shapes like "child", "parent", "thief", attribution of motives or emotions to the shapes) or not (neutral descriptions of the shapes' movements). Anthropomorphisms of objects were found in 71.4% of the reports. Table 4 shows that this was independent from the conditions during the course of the experiment. Grade of elaboration of the written reports, quantified by average number of words, ranges from 71.05 words (lab setting; experimenter involved, oral instruction) to 109.40 words (Internet setting).

An analysis of variance regarding elaboration of reports with the factor "setting" did not show a significant effect. In an analysis of variance with the factor "mode of presence of the experimenter" a significant effect arose,  $F(2, 104) = 3.72, p < .05$ . In the Internet setting, when an experimenter appeared only "virtually", participants wrote the most elaborate reports (109.39 words;  $SD = 48.16$ ). When an experimenter was actively involved in the course of the experiment and gave oral instructions in laboratory and "public area" setting, the reports consisted only of 75.71 words ( $SD = 44.18$ ). This difference proved significant in paired comparisons according to Scheffé. No difference was found between those two conditions, and the laboratory and "public area" setting, when the experimenter attended the experiment, but was not actively involved. Under this condition the reports consisted of 90.93 words averagely ( $SD = 51.55$ ).

Table 4

Results of the experiment following Heider and Simmel (1944): Grade of anthropomorphisms referring to objects seen in the movie and grade of elaboration (number of words) of written reports under different conditions

Setting	Mode of presence of the experimenter	N	Anthropomorphisms of objects found in report	Elaboration of report (number of words)
Internet	"virtual experimenter"	23	73% of participants	109.39 (SD = 48.16)
Public area	Experimenter present, written instructions	21	80% of participants	80.47 (SD = 46.28)
Public area	Experimenter involved, oral instructions	22	63% of participants	79.95 (SD = 52.65)
Lab	Experimenter present, written instructions	19	63% of participants	102.47 (SD = 55.77)
Lab	Experimenter involved, oral instructions	20	75% of participants	71.05 (SD = 33.02)

As in the original study by Heider and Simmel (1944) a predominant part of participants – regardless of conditions during the course of the experiment – perceived the events in the movie as a social situation, in which the animated objects acted purposely. Therefore, internal validity was given under all conditions. The different settings did not significantly affect the efforts of participants in writing reports – although reports in the “public area” setting were comparatively shorter (see Table 4). The involvement of an experimenter affected participants efforts in writing reports: When the experimenter was not physically present as in the Internet setting, participants wrote more elaborate reports as when the experimenter was attending the experiment. The least elaborate reports were given when the experimenter was actively involved in conducting the experiment.

### Discussion

Results of the three experiments show that it is important to take the respective psychological variables studied into account, when discussing validity and precision of experiments in the Internet and the laboratory setting. It proved to be a valuable approach not only to compare results from experiments conducted in Internet and lab setting, which differ in many ways, but to take a closer look by varying the standardization of the experimental arrangements and the mode of presence of the experimenter in a more systematic way. Although - independent of the different conditions under which they were conducted - all experiments yielded the expected results and can therefore be considered internally valid, differences in participants' performance were found dependent on the standardization of experimental arrangements or on the mode of presence of the experimenter.

In general, the results speak in favor of the validity of Internet-based experiments. When measures like response latencies or recall rates were obtained, which are usually considered susceptible to external disturbances that might lead to higher error variances, different standardization of the experimental arrangements in laboratory, “public area” and Internet setting did not affect the precision of measurements in a systematic way.

Participants' performance in the mental comparison task was not impaired due to the lack of shielding from external disturbances in Internet and “public area” settings. This corresponds to the findings of McGraw et al. (2000). When participants were asked to recall previously memorized items, poorer recall rates were obtained in “public area” setting compared to lab and Internet setting. Presumably, participants were distracted from their task in the experiment in the “public area” setting. But apparently, these disturbances do not affect participants' performance in general.

It is promising that no differences in performance between lab and Internet setting were found: Presumably, participants themselves arrange for an undisturbed environment when participating in an Internet-based experiment. When experiments were implemented in the virtual laboratory, efforts were made to warn participants that external disturbances might jeopardize the serious intentions of the experiment. Instructions informed participants about the anticipated duration and the goals of the experiment and motivated them to take



full concentration during the whole course of the experiment. Results indicate that in the majority of the cases participants would commit themselves to those demands.

Experiments involving assessment of memory performance are often considered inappropriate for being conducted via Internet, because participants may "cheat" by taking notes or by participating repeatedly to improve their performance (Metzger et al., 2003). Obviously, this is not a problem in the reported experiment on mnemotechniques. Performance in Internet setting was not noticeably superior to "public area" and laboratory setting, where the physical attendance of the experimenter makes cheating less likely to occur. Obviously, high commitment can be obtained in well-designed experiments, when participants are instructed appropriately.

Mode of presence of the experimenter did not affect performance when response latencies or recall rates were recorded. Participants followed instructions and made comparable efforts in fulfilling the tasks assigned to them in those experiments. The involvement of an experimenter affected participants' performance only when they were asked to give a written report. Mode of presence of the experimenter showed a clear impact on the elaborateness of reports: Participants' efforts in writing extensive descriptions of events seen in a movie were not so strong when the experimenter was actively involved in face-to-face communication with the participants during the course of the experiment. In the Internet setting, in which the experimenter was involved only "virtually", reports were the most elaborate. One reason may be that the task of writing an extensive report is perceived more reasonable, when the person who this report is addressed to is not involved in the course of the experiment. It might be irritating to *write* a report addressed to a person sitting next to the author and observing him. Generally, a greater willingness for self-disclosure is reported for Internet-based studies, where anonymity of participants is given, compared to other settings, where personal contact between researchers and participants is established (e.g. Joinson, 1999). Evidently, the social situation in an experiment can affect participants' efforts, depending on the given task.

The results of these comparative studies suggest that further research is necessary. While lack of standardization of the experimental arrangements in Internet-based experiments did not affect the results in the experiments presented, the different social situation in Internet and the laboratory setting had an influence on participants efforts at least in one case. Therefore it seems especially interesting to take a closer look at the impact of different aspects of the social situation in Internet-based experiments on different psychological variables, e.g. in experiments on perceptual defense (Bruner & Postman, 1942; Freeman, 1954), on distortions of memories (Loftus & Palmer, 1974; Jacoby, Woloshyn, & Kelley, 1989), or on the impact of group standards on level of aspiration (Festinger, 1942). Comparative studies may show how the anticipated effects are affected by the involvement of an experimenter: Can prerequisites for "perceptual defense" be produced, if an experimenter is only "virtually" attendant? Can memory distortions be elicited or fictitious group standards be induced persuasively by a virtual experimenter just like in the laboratory?

Comparative studies like the reported ones are not only valuable for research but for the enhancement of quality in teaching methods of experimental psychology as well. Since Internet-based experiments embody methodological characteristics, it is required that students of psychology do not only prepare for experimental research in the laboratory, but also meet the special requirements involved in the design of Internet-based experiments (Kraut et al., 2004; Reips, 2002). It would be ideal if students could directly experience the specifics of both the Internet and the lab setting. By using the virtual laboratory Lab.OR in teaching experimental methods (Heineken et al., 2003; Ollesch et al., 2003; Ollesch et al., 2004), students can conduct their own comparative studies, supervised by their tutors. Lab.OR offers the unique opportunity for students to plan, create, and conduct Internet based studies *without any programming* (cf. Kraut et al., 2004). By conducting the same experiments in the Internet and laboratory setting as well as in public areas on campus, they can directly experience the characteristics of research in different settings. Methodological issues of standardization of environmental conditions and problems of the social psychology of experimenting clearly become visible.

## References

Aronson, E., Ellsworth, P. C., Carlsmith, J. M., & Gonzales, M. H. (1990). *Methods of research in social psychology*. New York: McGraw-Hill.

Azar, B. (2000a). A web of research: They're fun, they're fast, and they save money, but do Web experiments yield quality results? *Monitor on Psychology*, 31, 42-47.

Azar, B. (2000b). A Web experiment sampler. *APA Monitor*, 31, 46-47.

- Birnbaum, M. H. (2000a). Decision making in the lab and on the Web. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 3-34). San Diego, CA: Academic Press.
- Birnbaum, M. H. (2000b). *Psychological experiments on the Internet*. San Diego, CA: Academic Press.
- Birnbaum, M. H. (2004a). Human research and data collection via the Internet. *Annual Review of Psychology*, 55, 803-832.
- Birnbaum, M. H. (2004b). Methodological and ethical issues in conducting social psychology research via the Internet. In C. Sansone, C. C. Morf, & A. T. Panter (Eds.), *The Sage handbook of methods in social psychology* (pp. 359-382). Thousand Oaks, CA: Sage.
- Bruner, J. S., & Postman, L. (1947). Emotional selectivity in perception and reaction. *Journal of Personality*, 16, 69-77.
- Bungard, W. (Ed.). (1980). *Die "gute" Versuchsperson denkt nicht – Artefakte in der Sozialpsychologie*. [The "good" experimental subject does not think. Artifacts within social psychology.] München: Urban & Schwarzenberg.
- Francis, G., Neath, I., & Surprenant, A. M. (2000). The cognitive psychology online laboratory. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 267-283). San Diego, CA: Academic Press.
- Freeman, J. T. (1954). Set or perceptual defense? *Journal of Experimental Psychology*, 48(4), 283-288.
- Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust Web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist*, 59(2), 93-104.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, 57, 243-259.
- Heineken, E., Ollesch, H., & Schulte, F. P. (2003). Experimentalpsychologische Ausbildung im virtuellen Labor: Das Laboratorium für Online-Research (Lab.OR). [Experimental psychological education in the virtual lab: The Laboratory for Online Research (Lab.OR).] In G. Krampen & H. Zayer (Ed.), *Psychologiedidaktik und Evaluation IV*. Bonn: Deutscher Psychologen Verlag.
- Hewson, C. M., Laurent, D., & Vogel, C. M. (1996). Proper methodologies for psychological and sociological studies conducted via the Internet. *Behavior Research Methods, Instruments, & Computers*, 28, 186-191.
- Hopmann, G., & Galley, N. (1986). Microcomputer in der Psychologie. [Micro computers in psychology.] In R. Brickenkamp (Ed.), *Handbuch apparativer Verfahren in der Psychologie* (pp. 423-468). Göttingen: Hogrefe.
- Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences on memory produced by dividing attention. *Journal of Experimental Psychology: General*, 118(2), 115-125.
- Joinson, A. (1999). Social desirability, anonymity, and Internet-based questionnaires. *Behavior Research Methods, Instruments, & Computers*, 31(3), 433-438.
- Krantz, J. H., Ballard, J., & Scher, J. (1997). Comparing the results of laboratory and World-Wide-Web samples on the determinants of female attractiveness. *Behavior Research Methods, Instruments, & Computers*, 29, 264-269.
- Krantz, J. H., & Dalal, R. (2000). Validity of Web-based psychological research. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet*. (pp. 35-60). San Diego, CA: Academic Press.
- Kraut, R., Olson, J., Banaji, M., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online. Report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. *American Psychologist*, 59(2), 105-117.
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior*, 13, 585-589.

- McGraw, K. O., Tew, M. D., & Williams, J. E. (2000). The integrity of Web-based experiments: Can you trust the data? *Psychological science*, *11*, 502-506.
- Metzger, M. M., Kristof, V. L., & Yoest, D. J. (2003). The World Wide Web and the laboratory: A comparison using face recognition. *CyberPsychology, & Behavior*, *6*(6), 613-621.
- Meyer, H. A. (2002). Webbasierte Experimente [Web based experiments]. In: D. Janetzko, M. Hildebrandt, & H. A. Meyer (Eds.), *Das Experimentalpsychologische Praktikum im Labor und im WWW* (pp. 113-127). Göttingen: Hogrefe.
- Morris, P. E., & Stevens, R. (1974). Linking images and free recall. *Journal of Verbal Learning and Verbal Behavior*, *13*(3), 310-315.
- Moyer, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception and Psychophysics*, *13*(2), 180-184.
- Musch, J., & Reips, U.-D. (2000). A brief history of web experimenting. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 61-88). San Diego, CA: Academic Press.
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). E-research: Ethics, security, design, and control in psychological research on the Internet. *Journal of Social Issues*, *58*(1), 161-176.
- Ollesch, H., Heineken, E., & Schulte, F. P. (2003). Das Labor im Rucksack – mobile computing in der psychologischen Grundlagenausbildung. [The lab in the backpack – mobile computing in the psychological basic education.] In M. Kerres & B. Voß (Eds.), *Digitaler Campus. Vom Medienprojekt zum nachhaltigen Medieneinsatz in der Hochschule*. Münster: Waxmann.
- Ollesch, H., Schulte, F. P., & Heineken, E. (2004). Experimentalpsychologische Grundlagenausbildung in einem hybriden Lernarrangement – das virtuelle Lab.OR als Cognitive Tool. [Experimental psychological education in a hybrid tutorial arrangement – the virtual Lab.OR as a cognitive tool.] In M. Kerres, M. Kalz, J. Stratmann, & C. de Witt (Eds.), *Didaktik der Notebook-Universität*. Münster: Waxmann.
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*, 776-783.
- Reips, U.-D. (2000). The Web experiment method: Advantages, disadvantages and solutions. In M. H. Birnbaum (Ed.), *Psychological experiments on the internet*. (pp. 89-114). San Diego, CA: Academic Press.
- Reips, U.-D. (2002). Standards for internet-based experimenting. *Experimental Psychology*, *49*(4), 243-256.
- Rosenthal, R. (1966). *Experimenter bias in behavioral research*. New York: Appleton-Century-Crofts.
- Senior, C., Phillips, M. L., Barnes, J. (1999). An investigation into the perception of dominance from schematic faces: a study using the World-Wide-Web. *Behavior Research Methods, Instruments, & Computers*, *31*, 341-346.
- Traxel, W., Gundlach, H., & Zschuppe, U. (1986). Zur Geschichte der apparativen Hilfsmittel in der Psychologie. [On the history of technical aids in psychology.] In R. Brickenkamp (Ed.), *Handbuch apparativer Verfahren in der Psychologie* (pp. 1-22). Göttingen: Hogrefe.