

Field experiment in psychology. Is there still room for conducting research in natural conditions?

BACKGROUND

This paper reflects on the status of the field experiment in contemporary psychology (especially its social branch). The subject of analysis is how frequently this method is used, especially in the context of the growing popularity of on-line research.

PARTICIPANTS AND PROCEDURE

The empirical part of the article presents an experiment revealing the uniqueness of the field experiment as a method allowing for the identification of subtle variables that can have a real impact on the results obtained. In the whole study 128 people were examined (half of the group were women). The obtained results show that some phenomena cannot be examined in any other way than by means of the field experiment.

RESULTS

The process of moving away from the study of real behaviours in the real environment, observed over the last

twenty years, has not stopped since the criticism of this phenomenon expressed about 20 years ago. On the contrary, it has intensified. The phenomenon described in the article is definitely harmful to the scientific discipline so strongly connected with the social context of research, namely social psychology.

CONCLUSIONS

Time will show whether the increasingly loud voices criticising this trend will translate into an actual change of methods used by social psychologists.

KEY WORDS

field experiment; research methods in psychology; replication crisis

ORGANIZATION – Faculty of Psychology, SWPS University of Social Sciences and Humanities, Wrocław, Poland

AUTHORS' CONTRIBUTIONS – A: Study design · B: Data collection · C: Statistical analysis · D: Data interpretation · E: Manuscript preparation · F: Literature search · G: Funds collection

CORRESPONDING AUTHOR – Tomasz Grzyb, Ph.D., Faculty of Psychology, SWPS University of Social Sciences and Humanities, 30 B Ostrowskiego Str., 53-238 Wrocław, Poland, e-mail: tgrzyb@swps.edu.pl

TO CITE THIS ARTICLE – Grzyb, T. (2020). Field experiment in psychology. Is there still room for conducting research in natural conditions? *Health Psychology Report*, 8(4), 428–434. <https://doi.org/10.5114/hpr.2020.97159>

RECEIVED 07.05.2020 · REVIEWED 31.05.2020 · ACCEPTED 04.06.2020 · PUBLISHED 10.07.2020

BACKGROUND

Perspectives on Psychological Science is one of the most influential psychological journals in the world. One of its features is the fact that, apart from articles supported by research results (or any other data), it allows for the publication of manuscripts that can be called “policy papers” (and those who remember history well could classify them as “what we are fighting for”). We are talking here about articles examining changes in psychology as a scientific discipline, analysing trends and directions of its development and diagnosing potential problems that limit its progress. The paper written by Robert Cialdini in 2009, entitled “We have to break up”, can be classified as one such paper (Cialdini, 2009).

The title suggested by Cialdini obviously refers to a kind of a “goodbye letter” written to a partner (a custom specific rather for teenagers) and based on quite characteristic key phrases: “it’s not your fault, but mine”, “we’ve been getting away from each other recently”, “yes, I’ve seen other people”, “I only want the best for you”. Cialdini uses exactly the same phrases, but addressing them to social psychology as a science. The author of the text points out three – as he claims, rather positive – fundamental changes that have taken place in psychology recently, pointing out, however, the specific “side effects” they have on this variant of experimental research that he likes to plan and carry out the most, that is, field experiments. These three changes are: the cognitive revolution (understood as focusing on cognitive variables as factors explaining human behaviour), the expectation that a series of studies will be published in the most important scientific periodicals (discussed by us in the section on replication), and the popularization of mediation analysis as a statistical tool for data analysis. In a very short text (the whole paper has only two pages), Cialdini describes why these three changes are making field experiments less popular, referring to what could be called “publishing economics”. If the chance to publish an article describing the research that has been carried out depends on the degree to which the author meets the expectations of reviewers and editors, and they expect (1) to take into account cognitive aspects in explaining behaviour, (2) to present a series of studies confirming the existence of the described phenomenon, (3) to perform analyses looking for mediators of the described phenomenon, then researchers carrying out field experiments are in this situation in a lost position. First of all, the most common dependent variable in field studies is behaviour. In addition, in the vast majority of cases it is treated dichotomously (someone threw money into a can, agreed to sign the petition, replaced the bulb with an energy-saving one – or did not). With the dependent variable defined in this way, it is difficult to incorporate the issue of veri-

fying the cognitive predictors of a given behaviour into the experimental procedure. Secondly, according to Cialdini, although the very expectation of self-replication of the effects obtained is in itself a thing worth praising, it is much easier to do it in the case of a questionnaire study than a field experiment. As Cialdini descriptively explains, the very time spent on obtaining the consent of the relevant ethics committee to conduct a field study takes as much time as a series of laboratory experiments, not to mention the duration of the research itself (which often – similarly to a small theatre performance – requires long preparations on the part of experimenters, confederates, troubleshooters and others involved in the experiment). It should be emphasized that the author of this paper does not criticise here the very idea of the need for replicating research – he only points out that it has completely different consequences for researchers working in the laboratory and those who do their research in the natural environment. The third element – the expectation of mediation analyses – is also very difficult to implement in the case of field research. The participants in such experiments (hotel guests, passers-by, customers buying buns in the supermarket) have no special interest in providing us with additional data enabling us to perform mediation analyses – they have already done their job (they used a towel in the hotel room for the second time, helped to collect scattered papers for a girl walking around with crutches, put a plastic glove on their hand while choosing bread). They would not, for example, fill in a questionnaire allowing us to assess their need for cognitive closure (Webster & Kruglanski, 1994), even if we really agreed with the idea of mediating the willingness to help with this variable.

One should reflect on the problematic consequences of abusing the self-reporting measures in psychology. These have already been reported by Richard LaPiere (1934) in his article on the well-known research into prejudice against the Chinese. LaPiere demonstrated that when asked by the researcher to picture their own behaviour, the subjects tended to present themselves in a more favourable light than in reality, e.g. by picturing themselves as individuals observing social norms. It is obvious though that the subjects do not have a full insight into the reasons behind their own behaviour (e.g. related to the emotions they experience) and by the same token, they are unable to fully picture them.

It should be noted that one of the reasons behind the shift away from the execution of studies in the natural environment is the ongoing process consisting in making our actions contingent upon the observance of the law (Grzyb, 2017). The experimenter’s assistants are subject to numerous risks (related to conducting their work in the natural environment – being hit by a car, conflicts with subjects unaware

*Field experiment
in psychology*

of their role, etc.). Such risk is eliminated in the case of laboratory studies, and this may be one reason behind the growing popularity of this form of study.

All these factors clearly affect the chances of publishing results from field experiments – Cialdini himself writes that he has not managed to publish any paper describing field research over the past fifteen years, in the key (one could say flagship) periodical presenting reports from social psychology – in the *Journal of Personality and Social Psychology (JPSP)*. Interestingly, he was the author or co-author of papers which were published there at the same time, yet they were devoted to research conducted using other methods (Griskevicius et al., 2007). Miles Patterson points out a similar phenomenon. He analysed another recognized journal – *Personality and Social Psychology Bulletin* – assessing what proportion of the articles published there contains data that can be considered as measurement of behaviour (Patterson, 2008). The results were absolutely unambiguous – this is because while in 1976 in about 70% of the presented studies any measurement of behaviour was considered to be a variable (dependent or independent), in 1996 and 2006 (Patterson used such study samples) it was no more than 25%. The conclusion is that we see a very similar phenomenon to the one described by Cialdini in his own example. In his paper, Patterson draws attention to another element – the decline in interest of researchers in the so-called “unfocused interaction”. It is defined as specific interactions between people (or, more broadly, the sender and the recipient of a message), which take place without being reflected in verbal or non-verbal communication between them (Goffman, 1963). Let us note that a significant proportion of measurements in studies carried out as field experiments concerns such interactions – for example, the studies on the effectiveness of road signs by James Fisher (1992) or the studies on the length of visual contact made by strangers (Kleinke, 1986).

The aforementioned article by Cialdini was published in 2009. It was actually a part of a series of manuscripts drawing attention to the problems of psychology (especially social) as a scientific discipline (Baumeister, Vohs, & Funder, 2007; Klebaniuk, 2012). A question arises then, whether something has

changed over the past few years – maybe even for the better?

Kai Sassenberg and Lara Ditrich (2019) tried to answer this question. They analysed the whole year-books of the flagship journals on social psychology in order to compare the nature of the research in 2009-2011 and 2016-2019. They examined their characteristics, such as sample sizes, the way they were selected, and the research methods used. The results which they obtained are not optimistic. Although it may of course lift the spirits that increasingly larger samples are analysed as a part of research (though one may wonder whether the research carried out on a sample of 1,353,075 people can still be treated as being significantly carried out on a “sample”), the direction in which the methods used have changed perfectly confirms the diagnosis made by Cialdini. Sassenberg and Ditrich have shown that the number of tests carried out using self-reporting measures is increasing, and they have noted a significant rise in the number of tests carried out on-line (which may explain the surprising number of tests mentioned several lines above).

This problem was addressed in another article (Anderson et al., 2019) that pointed to the process of the “mTurkification” of social psychology. Anderson et al. analysed important scientific journals in the field of social psychology (*Journal of Personality and Social Psychology*, *Personality and Social Psychology Bulletin*, *Journal of Experimental Social Psychology*) and determined what percentage of studies carried out in particular years was based on on-line tools (five-year intervals were used). The precise data are presented in Table 1.

It is therefore clear that an increasing proportion of research in the field of social psychology is based on self-reporting and is increasingly often being carried out on-line on various platforms, like the popular mTurk. It should be noted that this does not bring only negative consequences. On-line research simplifies the life of the researchers, allowing them to collect large amounts of data in a relatively short time. It also makes it possible, thanks to its technical capabilities, to carry out even complex experiments in an intergroup system. In one of the studies carried out on the Ariadne platform (the Polish equivalent of mTurk), we even managed, together with Dariusz Doliński, to replicate partially Stanley Milgram’s research on obedience to authority (Grzyb, Doliński, Trojanowski, & Bar-Tal, 2018). Naturally, this had to be organised in a way that significantly affected the experimental realism, as the subjects were only supposed to imagine that they were participating in such an experiment, but the data obtained basically coincided with those we obtained in a traditional laboratory experiment. This shows, therefore, how great the possibilities for researchers to carry out experiments on the Internet are.

Table 1

Number of on-line and off-line studies in selected social psychology journals

	2005	2010	2015
On-line research	7	50	153
Off-line research	213	197	155

Note. Author’s own review based on Anderson et al., 2019.

However, it is worth asking ourselves what consequences this entails – and therefore what price science pays for such a clear departure from research carried out in natural conditions in an environment known by the subjects and familiar to them. Is it sometimes not the case that by moving away from field experimentation as a research method, we lose the chance to capture the delicate variables that influence the results, which cannot be measured by other means?

Let us note that most often field experiments are performed on streets, squares, university campuses – generally in places where many people circulate – and the performance of the research itself is facilitated from this perspective (the time between successive interactions with the subjects is minimized). Rarely, however, do the descriptions contain precise information on what facilities are located in the vicinity of the square or street where the experiment is conducted. And as it turns out, the type of these facilities can have a significant impact on the results of the experiment. The French psychologist Nicolas Guéguen is a scientist analysing the role of the research site and its influence on the results achieved (especially from the perspective of altruistic behaviour, but not only). His studies from 2012 (Guéguen, 2012) showed, among other things, that young women are more likely to give their phone number to a man when he asks for it near a florist's shop (compared to requests made in close proximity to a confectionery or shoe shop). In other studies carried out by Lubomir Lamy, Jacques Fisher-Lokou, and Nicolas Guéguen (2015), the researchers examined how the proximity of the place where the respondent could provide help to the person in need influenced his decision to do so. One hundred and ninety-two individuals participated in the study (half of them were women). The experimental procedure was related to the situation in which the participant could “volunteer” to help a young woman. This woman, with a visible orthosis put on her leg, was walking along the pavement, leaning with one hand on a crutch and holding a pack of sweets in the other. At some point (around 5 m from the future study subject, who had been previously selected from among passers-by), the woman would drop the sweets and try to pick them up while balancing to avoid falling down. The researchers wanted to determine the probability of her receiving help (picking up the sweets or verbally offering to pick them up) if the fall happened in one of four possible locations: a control point (a street away from public buildings), a hospital, a church and a florist's. The researchers also evaluated what impact the gender of the person examined had on the decision to provide assistance.

The obtained results showed that gender did not influence the offer of help (or its provision). However, there were clear differences in the reactions of the examined persons in different places. Most of the

respondents helped in the vicinity of the hospital (91.6% of the study participants) and in the vicinity of the florist's (87.5%). It was different in the vicinity of a church (75%) and in a street far away from these places (68.7%). The researchers explained the differences obtained with the associations that people have with the research sites. In the pilot study they showed that most people associate the florist's with love and hospital with helping, which may have triggered a pattern of helping a person in need. However, because associations with these objects may vary depending on the culture, we decided to conduct a slightly modified version of this experiment in Polish conditions.

PARTICIPANTS AND PROCEDURE

We made a number of modifications in the experimental design itself. First of all we changed the dependent variable into a verbal request for financial assistance in a troublesome, random need. The second modification was to add a place where the respondents should (according to the hypotheses) be less willing to give such help. We considered such a place to be the vicinity of a grocery discount shop, mainly due to the fact that it is precisely around discount stores that people trying to ask passers-by or shop clients for “some change” appear quite often. The third modification was the introduction of a second person asking for help – depending on the experimental group, it was a woman or a man in our study. The last element differing from the original French one was to record several demographic variables: we noted the exact age, education, marital status and – of course – gender of the people surveyed. The survey with these questions was handed out after the experiment itself had been completed and the participants were explained that they had participated in a psychological study.

The survey was conducted by two experimenter's assistants: a woman and a man, each aged 25. They wore neat, decent clothes: the woman was wearing a blue, knee-length dress and cornflower jacket with short sleeves, and blue sandals with small heels. She had a ponytail. The man was wearing beige trousers, a jacket, shoes and a blue shirt. He had short hair, smoothly combed backwards. The research took place in four locations in Wrocław:

- by St. Joseph's Care Church at Olbińska St.,
- by T. Marciniak Hospital on Traugutta St.,
- by the Biedronka discount store on Królewiecka St.,
- by a florist's at Solny Square.

The experimenter's assistants approached every fifth randomly passing person who had not seen the previous interaction and asked for help (assignment to a particular group was based on a previously created randomization). They made the following request:

Sorry to bother you, but I have a problem. My money and documents have been stolen and I have no way to get home. I don't want to go without a ticket, so please give me a ticket or three zlotys so I can buy one.

When the person to whom the request was addressed pulled out the money or the bus ticket, he or she was approached by the supervisor experimenter, who thanked the person very much for his/her help and informed the person that he/she was taking part in a psychological experiment. He asked for a short survey on age, education and place of residence. He also stated that the above data would be used only for scientific purposes. Each of the experimenter's assistants asked 16 people for help in each location.

RESULTS

In the whole study 128 people were examined (half of the group were women). The number of individuals willing to help in particular locations is presented in Table 2, taking into account the gender of the respondents.

In Table 3 the numbers of people helping in particular conditions are shown, taking into account the gender of the experimenter's assistant.

Statistical analysis showed a strong effect of the place where help was asked for – most frequently re-

spondents decided to help in the vicinity of a church (78.1%) or a hospital (71.9%), while significantly fewer people offered help in the vicinity of a florist's shop (56.3%) and the least in the vicinity of a grocery discount store (9.4%). These differences were statistically significant – $\chi^2(3) = 37.32, p < .001$, Cohen's $d = 1.28$. The influence of the gender of the investigator's assistant and the gender of the respondent on the willingness to help was also assessed – both effects were found to be statistically insignificant. However, the interaction between them was found to be significant – the women surveyed were more or less equally eager to help the asking man and the asking woman. However, the men surveyed were more willing to offer help to the asking man than to the asking woman: $F(1, 124) = 5.47, p = .021$, Cohen's $d = 0.42$. The remaining variables (age, education, marital status) did not affect the willingness to help.

DISCUSSION

The obtained results can be considered from the perspective of the field experiment methodology in at least two ways. The first is to show the specific problems that may be posed by replication of research carried out in other countries (and, more broadly, in other cultures). It can be seen that the area of florists,

Table 2

Numbers of individuals willing to help in particular locations

Site	Church	Respondent	Woman	Result		
				Help	No help	Total
			Man	13	3	16
			Man	12	4	16
		Total		25	7	32
	Hospital	Respondent	Woman	14	2	16
			Man	9	7	16
		Total		23	9	32
	Florist's	Respondent	Woman	10	6	16
			Man	8	8	16
		Total		18	14	32
	Discount shop	Respondent	Woman	0	16	16
			Man	3	13	16
		Total		3	29	32
Total		Respondent	Woman	37	27	64
			Man	32	32	64
		Total		69	59	128

Table 3

Numbers of people helping in particular conditions (including the gender of the experimenter's assistant)

Site				Result		
				Help	No help	Total
Church	Assistant	Woman	11	5	16	
		Man	14	2	16	
		Total	25	7	32	
Hospital	Assistant	Woman	12	4	16	
		Man	11	5	16	
		Total	23	9	32	
Florist's	Assistant	Woman	13	3	16	
		Man	5	11	16	
		Total	18	14	32	
Discount shop	Assistant	Woman	2	14	16	
		Man	1	15	16	
		Total	3	29	32	
Total	Respondent	Woman	38	26	64	
		Man	31	33	64	
		Total	69	59	128	

*Field experiment
in psychology*

which in the French research was found to increase the willingness to help, is no longer so “effective” when the research is conducted in Polish conditions. The opposite is true of the church: the original French studies in the area near this place did not show a big difference in the number of helpful behaviours in comparison with the control group – in Polish studies it was by the church where the majority of people decided to help (although, naturally, some of the differences may be due to another operationalisation of the dependent variable). Therefore, we may dare to claim that replication of experiments carried out in different cultures entails the need to take into account cultural specificity also at the stage of constructing an experimental scheme and the very operationalisation of variables.

The second important result obtained in the study is an extremely low (nearly meeting the criteria of the so-called “floor effect”) willingness to help in the area of discount grocery stores. As it seems, such a result may be caused by the frequent meeting of people asking for money in certain places in the city – in the vicinity of railway stations, main city squares and precisely discount stores and supermarkets. One possible explanation for this phenomenon is the heuristics of representativeness (Kahneman, 2016; Kahneman & Tversky, 1972; Sotirovic, 2016) – if we meet a person asking for money in the vi-

cinity of certain places, we tend to categorise him/her as someone trying to extort cash from us, and we will not particularly care about the reasoning that the person is using. In many studies carried out in Wrocław, we have noticed that two of the worst sites for field research are Wrocław Market Square and Solny Square – two places of high pedestrian traffic located near to each other. It turned out that in these particular locations the default answer to all possible requests (or even an attempt to open one's mouth to make such a request) is “no”.

CONCLUSIONS

To sum up, it is worth noting that the process of moving away from the study of real behaviours in the real environment, observed over the last twenty years, has not stopped since the criticism of this phenomenon expressed about 20 years ago. On the contrary, it has intensified. As we have tried to show in this article, it is a phenomenon that is definitely harmful to the scientific discipline so strongly connected with the social context of research, namely social psychology (see Brzeziński & Siuta, 1991). Time will show whether the increasingly loud voices criticising this trend (Doliński, 2018) will translate into an actual change of methods used by social psychologists.

REFERENCES

- Anderson, C. A., Allen, J. J., Plante, C., Quigley-McBride, A., Lovett, A., & Rokkum, J. N. (2019). The MTurkification of social and personality psychology. *Personality and Social Psychology Bulletin*, *45*, 842–850. <https://doi.org/10.1177/0146167218798821>
- Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, *2*, 396–403. <https://doi.org/10.1111/j.1745-6916.2007.00051.x>
- Brzeziński, J., & Siuta, J. (1991). *Spoleczny kontekst badan psychologicznych i pedagogicznych: Wybór tekstów* [Social context of psychological and pedagogical research: Selection of texts]. Poznań: Wydawnictwo Naukowe UAM.
- Cialdini, R. B. (2009). We have to break up. *Perspectives on Psychological Science*, *4*, 5–6. <https://doi.org/10.1111/j.1745-6924.2009.01091.x>
- Doliński, D. (2018). Is psychology still a science of behaviour? *Social Psychological Bulletin*, *13*, 1–14. <https://doi.org/10.5964/spb.v13i2.25025>
- Fisher, J. (1992). Testing the effect of road traffic signs' informational value on driver behavior. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *34*, 231–237. <https://doi.org/10.1177/001872089203400208>
- Goffman, E. (1963). *Behavior in public places*. New York: The Free Press.
- Griskevicius, V., Tybur, J. M., Sundie, J. M., Cialdini, R. B., Miller, G. F., & Kenrick, D. T. (2007). Blatant benevolence and conspicuous consumption: When romantic motives elicit strategic costly signals. *Journal of Personality and Social Psychology*, *93*, 85–102. <https://doi.org/10.1037/0022-3514.93.1.85>
- Grzyb, T. (2017). Obtaining informed consent from study participants and results of field studies. Methodological problems caused by the literal treatment of codes of ethics. *Polish Psychological Bulletin*, *48*, 288–292. <https://doi.org/10.1515/ppb-2017-0032>
- Grzyb, T., Doliński, D., Trojanowski, J., & Bar-Tal, Y. (2018). Cognitive structuring and obedience toward authority. *Personality and Individual Differences*, *133*, 115–120. <https://doi.org/10.1016/j.paid.2017.08.032>
- Guéguen, N. (2012). “Say it... near the flower shop”: Further evidence of the effect of flowers on mating. *The Journal of Social Psychology*, *152*, 529–532. <https://doi.org/10.1080/00224545.2012.683463>
- Kahneman, D. (2016). Heuristics and biases. In R. Sternberg, S. Fiske, & D. Foss (Eds.), *Scientists making a difference: One hundred eminent behavioral and brain scientists talk about their most important contributions* (pp. 171–174). Cambridge: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: a judgment of representativeness. *Cognitive Psychology*, *3*, 430–454. [https://doi.org/10.1016/0010-0285\(72\)90016-3](https://doi.org/10.1016/0010-0285(72)90016-3)
- Klebaniuk, J. (2012). Profesor Stapel na doping. O upiększaniu psychologii społecznej [Professor Stapel on dope. On the beautifying of social psychology]. *Psychologia Społeczna*, *3*, 213–217.
- Kleinke, C. L. (1986). Gaze and eye contact: a research review. *Psychological Bulletin*, *100*, 78–100. <https://doi.org/10.1037/0033-2909.100.1.78>
- Lamy, L., Fischer-Lokou, J., & Guéguen, N. (2015). Places for help: Micro-level variation in helping behavior toward a stranger. *Psychological Reports*, *116*, 242–248. <https://doi.org/10.2466/21.PR0.116k14w6>
- LaPiere, R. T. (1934). Attitudes vs. actions. *Social Forces*, *13*, 230–237. <https://doi.org/10.2307/2570339>
- Patterson, M. L. (2008). Back to social behavior: Mining the mundane. *Basic and Applied Social Psychology*, *30*, 93–101. <https://doi.org/10.1080/01973530802208816>
- Sassenberg, K., & Ditrich, L. (2019). Research in social psychology changed between 2011 and 2016: Larger sample sizes, more self-report measures, and more online studies. *Advances in Methods and Practices in Psychological Science*, *2*, 107–114. <https://doi.org/10.1177/2515245919838781>
- Sotirovic, M. (2016). Heuristic information-processing. In G. Mazzoleni (Ed.), *The International Encyclopedia of Political Communication* (pp. 1–5). Hoboken, NJ: Wiley.
- Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, *67*, 1049–1062. <https://doi.org/10.1037/0022-3514.67.6.1049>