Original Papers

Polish Psychological Bulletin
2016, vol. 47(2) 233–235
DOI - 10.1515/ppb-2016-0027

Tomasz Grzyb[*]

Why can't we just ask? The influence of research methods on results. The case of the "bystander effect"

Abstract*: The article discusses the issue of the departure from examining real behaviours in a real environment, a trend in social psychology which has been observed going back several years, and the impact of this phenomenon for social psychology as a scientific discipline. The article presents two studies on the well-known and explored "bystander effect" (Darley, Latane, 1968). This phenomenon is examined in two ways – once by way of a "traditional" field experiment conducted in natural conditions, and once through a survey. As it turned out, the results generated by the two studies were diametrically opposite, and only in the field experiment were we able to achieve a pattern of results consistent with those in the original studies.*

Key words*: field experiment, "Bystander effect", duffusion of responsibility*

In his memoirs "Not by Chance Alone. My Life as a Social Psychologist" (2010), Eliott Aronson relates a conversation he had with Gordon Allport which took place at Harvard in 1960. Allport, listening to the descriptions of experiments presented by Aronson, asked him – a bit connivingly – why he conducted those "deceitful experiments" rather than just asking people how they would behave. To his surprise, Aronson discovered that the eminent psychologist and specialist in personality psychology and social psychology had absolutely no idea about conducting research in the social world. It would seem that, 55 years after that conversation, many psychologists around the world are capable of repeating the exact same assumptions made by Allport.

Why have social psychologists simple abandoned the idea of conducting field studies? There are several specific reasons for this, and one general – the most important, which can be referred to as economics of publication. It seems almost self-evident that the number of articles to be found in respected scientific publications is a proxy for the quality of the modern scientist. If we then consider that field experiments are time-consuming and require the labour of entire teams of people, and their results are always uncertain owing to the large number of interfering variables – it is easy to understand those who feel that a better route for conducting studies is to distribute questionnaires over the Internet. This allows for results to be generated quicker, easier and at far less cost, and it is also possible to modify and replicate such a study essentially at once. It should also be kept in mind that every study requires the approval of an Ethics Committee – it is far easier to receive approval for survey studies than experimental ones (if only due to the issue of informed consent on the part of study participants, which is vexing in the case of street experiments). Another difficulty is that of statistically preparing results – we have a wide range of advanced statistical methods available for survey studies (path analysis, mediation analysis, structural equation modelling). Street experiments generally provide qualitative data (someone agrees to a request or not, puts up a poster or not, etc.). Treatments of such results appears quite primitive from the perspective of advanced statistics, and also reduces the chances of results being published in a respected journal.

We can thus observe that there are many reasons why researchers are inclined to avoid performing field experiments, preferring to focus on other methods.

What, however, does this state of affairs cost us? Are results generated by survey studies of equal value to those obtained during field studies? To find an answer to this question, a study was performed to explore the well-known and well-described "bystander effect". In a series of experiments examining real-life behaviours, Darley and Latane (1968) demonstrated the dependence between the number of witness of an interaction and the chance that help will be received. They showed that the more people who hear or see that someone is in need of help, the smaller the chances that

[*] **Uniwersytet Opolski, Instytut Psychologii, pl. Staszica 1, 45-015 Opole**

person will in fact receive assistance. Darley and Latane, as well as those who followed in their footsteps (meta-analysis of Fischer et al., 2011) conducted numerous natural and field studies which demonstrated the durability and constancy of that effect. However, would they have received similar results if they had followed the advice of Gordon Allport and simply asked people about their most likely reaction?

Study 1

Method

The study was performed in a train. Two female students entered a wagon at a rail station and entered a compartment in a manner that would indicate unfamiliarity with each other (the second entering a short time after the first, each entering at a different end of the wagon). Earlier, a research assistant had randomly selected a compartment that met the study criteria. In the first version, the compartment was to have one traveller; in the second (with the expected bystander effect) there was to be three, but their positioning in the compartment had to clearly indicate they were not acquainted with one another.

After entering the compartment the students removed their coats and sat opposite each other. After around 5 minutes, one of them turned to the other and said "Excuse me, I need to go to the toilet, would you mind keeping an eye on my coat?" The second responded "Yes, of course". A minute after the first student had left, the second went to her coat, removed a wallet from it and then took a PLN 50 banknote (approx. USD 10) from the wallet. She put the banknote in her pocket and replaced the wallet in the coat, then sat back in her seat. After another two minutes or so, the first student returned from the toilet and took her seat inside the compartment. If the other passengers did not react to what they had seen, the women left the compartment upon the train reaching the next station.

The dependent variable in the study was defined as any reaction at all leading to the "victim" learning of the theft. This could have been preventing the theft, loudly drawing the victim's attention upon her return, threatening to call the police or to leave a little note with the information "that woman stole from you". In other words – every circumstance which would lead to the theft being discovered. The study was performed in one day along the railway between two of Poland's major cities. 20 attempts were made, 10 of each scenario (either one or three witnesses to the theft).

Results

From among the 10 people in the first version (one witness), 9 reacted and informed the victim of the theft; from among the 10 situations in the second version (three witnesses), the theft was disclosed in four of them. The dependency reached statistical significance ($\chi^2 = 5.49$; $df = 1$; $p = 0.019$; Cohen's $d = 1.23$).

Discussion of Study 1

The results confirmed the hypotheses. Exactly as in the experiments by Darley and Latane, there was a dependency between the number of witness to the interaction and the desire to provide assistance to the person in need. It should be remembered, however, that the objective of the study was only to confirm empirically that a field (natural) situation would lead to results similar to those generated by earlier experimenters. It was decided in a successive study to examine what would occur if, a survey study substituted for the realistically arranged situation.

Study 2

Method

A survey with the following description of a situation was prepared:

"Imagine the following situation – you are in a train, sitting in an eight-person compartment alone/together with two other people. At the next station, one young woman joins you, followed a few moments later by a second. After a few minutes, one of the women says to the other "Excuse me, I need to go to the toilet, would you mind keeping an eye on my coat?" The second replies "Yes, of course." When the first one leaves, the second one waits around a minute, walks up to the coat of the woman who had just left, takes out a wallet, removes PLN 50 from it and puts the banknote in her pocket. She returns the wallet to the coat, smoothens it to make it appear as though nobody has touched it, and then returns to her seat. **What would you do in this situation?**"

Each group received the same version of the survey (with information how many people were in the compartment).

The study – to enhance environmental accuracy – was conducted on the platforms of railway stations in Wrocław and Jelenia Góra. Participants were randomly selected for participation in the study and randomly assigned to a particular group (travelling alone or with two other people). Under the description of the situation they were given an empty field to write down their most likely reaction. These descriptions were then transcribed, and competent judges were asked to assess these descriptions in terms of their helpfulness. Four competent judges (two women and two men) were asked to read the description of the situation (without information about the number of people in the compartment) and assess the reactions on a scale of 1 to 5 (1 – total absence of reaction, 5 – very helpful reaction to the problem).

Results

The study included 80 participants (40 per group, each group evenly divided by sex). The average score for helpfulness of reaction in conditions of travelling alone was 3.78, while for travelling with two other witnesses it was 4.01. The difference between the groups did not achieve statistical significance ($t = 0.774$, $df = 78$, $p = 0.441$, Cohen's $d = 0.17$). The number of '1' responses was also examined in the groups (meaning a situation which the judges defined as the total absence of reaction). In the group of people "travelling alone" there were 6 such people, while there were 5 in the second group.

General discussion

The results generated in the second study paint an entirely different picture from that of experiment 1. There was no difference between the groups, both in the competent judges' assessments of helpfulness of witness reaction, and in the number of people who did not react at all to the situation. It can be observed that the average helpfulness of reactions described by study participants was quite high (just over 4). Thus, the participants described their likely reaction far more positively (for example, from the perspective of their self-image) than what was observed in the experimental field study. Naturally, it should be emphasized that the described research is not the first to stress the role of drawing attention to the differences recorded by investigators employing various methodological approaches. We may recall the pioneering studies of Stanley Milgram concerning obedience to authority figures (Milgram, 1963). Another exceptionally interesting example of empirical demonstration of the differences between results achieved from analysis of declarations and of real behaviours comes from Hofling and his "hospital experiment" (Hofling et al., 1966).

The results generated are consistent with the theses proposed by Cialdini (2009) and Baumeister et al. (2007). Those researchers pointed to the damage done to psychology as a scientific discipline resulting from the excessive focus on studies relying on self-descriptive methods, survey methods and all other methods which in fact do not examine real behaviours in real-life conditions. Robert Cialdini admits that for some reason he published NOT ONE SINGLE ARTICLE in serious journals from 1996–2009 describing field study experiments (while publishing others performed according to other methods). This led him to stop taking on PhD candidates, as conducting field studies with him reduced their chances of getting published in prestigious journals, and thus significantly threatened their scientific careers. While Cialdini's declaration is not completely airtight (his output includes, for example, a text from 2008 which he co-authored and in which a portion of the studies described were field studies – Nolan et al., 2008), it does demonstrate a clear trend of abandoning studies of real behaviour in favour of other methods.

Baumeister, Vohs and Funder write that attempts to find studies in leading psychological journals examining real behaviours by subjects borders on the search for a needle in a haystack (one example of such a journal they cite is the Journal of Personality and Social Psychology). Essentially, the "behaviour" most often boils down to filling in a survey or in executing various "finger movements" at a computer – pressing the right button or operating a mouse. The number of studies concerning real experiments continues to decline – in 1976, it was 80% while today it is not more than 20%.

The studies described in the article demonstrate – as it would see – that besides the issues described by Cialdini and Baumeister et al., there is yet another, no less serious complication. Studies skipping over an analysis of real behaviours in natural conditions may simply be a source of artefacts. It could therefore be said that the presented results are a good illustration of how false an image we obtain when we forget as researchers that the study of real human behaviour in the natural environment is of key importance in our discipline.

References

Aronson, E. (2010). *Not by Chance Alone: My Life as a Social Psychologist.* New York: Basic Books.

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*(4), 396–403.

Cialdini, R. B. (2009). We have to break up. *Perspectives on Psychological Science*, *4*(1), 5–6.

Darley, J. M. & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology, 8*, 377–383. doi: 10.1037/h0025589

Fischer, P., Krueger, J. I., Greitemeyer, T., Vogrincic, C., Kastenmüller, A., Frey, D., & ... Kainbacher, M. (2011). The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological Bulletin, 137*(4), 517–537. doi: 10.1037/a0023304

Hofling, C. K., Brotzman, E., Dalrymple, S., Graves, N., & Pierce, C. M. (1966). An experimental study in nurse-physician relationships. *The Journal of nervous and mental disease*, *143*(2), 171–180.

Milgram, S. (1963). Behavioral study of obedience. *The Journal of Abnormal and Social Psychology*, *67*(4), 371.

Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., Griskevicius, V. (2008). Normative social influence is underdetected. *Personality and Social Psychology Bulletin, 34*(7), 913–923.