Studia Psychologica, Vol. 61, No. 2, 2019, 71-85 doi: 10.21909/sp.2019.02.773

Financial Product Choices: Does Attribute Preference Help Avoid the Attraction Effect?

Viera Bačová, Róbert Hanák
Institute of Experimental Psychology, Centre for Social and Psychological Sciences, Slovak Academy of Sciences

The attraction effect occurs when a third option is added to two seemingly equivalent options but it competes against only one of the original options. This increases the likelihood of the dominating option being chosen. In attraction effect studies, it is assumed that both attributes of the options are of equal importance to the decision maker. We aimed to examine whether attribute preference would affect the occurrence of the attraction effect when choosing financial products. A total of 487 undergraduate students were randomly assigned to groups with the financial product choice of two or three options. We found that when participants had no clear attribute preference the attraction effect occurred more frequently. Those with a clear preference for one attribute succumbed to the effect only when choosing a product unfamiliar to them. The research sheds light on two conditions of the attraction effect: the product familiarity and the attribute preference.

Key words: attraction effect; financial product choices; attribute importance

The Attraction Effect

Making a decision involves having to choose between two or more items or products. As these possess different attributes (for example a higher price or a higher quality), varying in importance to the decision maker, identifying which is most attractive is often difficult. When one item does better on one attribute and the other on another attribute, and if the decision maker considers both attributes equally important, a trade-off is required, which is not easily achieved. Therefore the decision maker looks for ways to simplify the decision and reduce the effort required (Payne, Bettman, & Johnson, 1993). One way is to use the context provided by the set of options (Tversky & Simonson, 1993).

One of the context effects to have gained most attention among decision theorists is the attraction effect (Huber, Payne, & Puto, 1982). It occurs when a third option is added to two seemingly equivalent options but it competes against only one of the original options, considerably increasing the likelihood of the dominating option being chosen. Numerous experimental studies have confirmed the robustness of the attraction effect in various experimental

Acknowledgment
This research was supported by a grant from Slovak Ministry of Education, Science, Research and Sport of the Slovak Republic VEGA 2/0118/17 *Risk assessment in decision making of individuals on the personal and company/business finances and business opportunities*.

Correspondence concerning this article should be addressed to Prof. Viera Bačová, Institute of Experimental Psychology, Centre for Social and Psychological Sciences, Slovak Academy of Sciences, Dúbravská cesta 9, 841 04 Bratislava, Slovak Republic. E-mail: viera.bacova@savba.sk

Received March 14, 2018

designs and in a wide range of domains and products (Dhar & Simonson 2003; Mishra, Umesh, & Stem, 1993; Simonson & Tversky 1992; Sivakumar & Cherian 1995; Wedell & Petibone, 1996). Although the robustness of the attraction effect has recently been disputed (Frederick, Lee, & Baskinet, 2014; Huber, Payne, & Puto, 2014; Simonson, 2014; Yang & Lynn, 2014), new experimental evidence and review studies have emerged emphasizing that "caution needs to be exercised before discounting the attraction effect's practical relevance" (Lichters, Sarstedt, & Vogt, 2015, p. 3).

For the attraction effect to occur, the newly added option (also labelled the "decoy" or "irrelevant option") has to resemble the option we wish to make more favorable (also labelled the "target"). The decoy does less well on one of the attributes but not the other. The option the decoy makes less favorable on a particular aspect is called the "competitor" (Hedgcock & Rao, 2009; Huber et al., 1982; Simonson, 1989; Simonson & Tversky, 1992). The decoy is only minimally preferable, but sets up a relationship of dominance between the two original options (therefore the attraction effect is also known as the asymmetric dominance effect; Ariely & Wallsten, 1995).

Conditions Affecting the Attraction Effect and Attribute Importance

Several research studies have sought to identify the conditions under which the attraction effect occurs and its intensity. The literature contains mixed findings on the role of conditions which attenuate or amplify the attraction effect. Malkoc, Hedgcock, and Hoeffler (2013) have summarized some of the factors that moderate (amplify or diminish) the attraction effect. They suggested that the attraction effect may be enhanced by, for example, the need to justify one's choice, higher product quality (options), product similarity, depletion, and decoy popularity. Attenuating factors may include meaningfulness, product familiarity, category knowledge, involvement, and preference strength. Nonetheless, there are some conditions which are clearly thought to influence the attraction effect but have yet to be investigated in relation to it.

In most attraction effect studies, it is assumed that the two original options are equivalents as they are chosen approximately equally often. At the same time, it is assumed that both attributes are of equal weight (importance) to the decision maker. But what if the decision maker considers one attribute to be more important than the other? Thus far in attraction effect research, the importance of the attributes to the decision maker has attracted little attention. Studies that have addressed this issue have produced contradictory conclusions.

In Malaviya and Sivakumar's experiments (1998), the attraction effect was more evident in participants who considered one of the attributes more important than in those who considered both attributes to be equally important. According to Malaviya and Sivakumar (1998), if one attribute is more important it makes it more difficult for decision makers to choose because they have to consider both the weight and local value of the attributes. Therefore instead of thinking hard, the decision maker is more likely to be influenced by the relationship between the decoy and the target option. The choice is easier for those who consider both attributes equally important, as they consider only the total value of the options. By contrast, Simonson (1989) considers the opposite case to be more valid. The attraction effect is more likely to occur if the attributes are equally important to the decision maker. This is because the choice becomes more complex and so the decision maker is more prone to rely on heuristics and the local dominance of the target option. In their study, Wedell and Pettibone (1996) concluded

that the dominance of the target can be increased when the more important attribute dominates in the decoy, since this then reinforces the perception that the target option has a higher total value. The option that does better on a more important attribute will be preferred and ultimately chosen as decision makers can use the local dominance of the option to make their choice.

Financial Products and the Attraction Effect

Financial products have some specific characteristics that set them apart from the products typically investigated in relation to decision heuristics and biases.

On one hand, there are reasons why the attraction effect can be expected in the financial domain. Firstly, it is not possible to describe financial products without using realistic numerical attributes relevant to their use in the real environment. Expressing stimuli numerically (as opposed to using verbal descriptions and images) has been found to enhance the attraction effect (Frederick et al., 2014; Simonson, 2014). Secondly, the sophistication of financial products makes calculating the best option highly demanding, and math and financial skills and sometimes expertise are required, therefore, we may expect the use of heuristics including the attraction effect.

On the other hand, there are reasons why we should not expect the attraction effect in the financial domain. Firstly, given the importance of financial decisions, we can suppose that the decision maker is highly engaged and more vigilant when choosing financial products. In relation to non-financial products, Mishra et al. (1993) have observed that a high degree of engagement reduces the attraction effect. Another reason for not having the attraction effect is that in financial decisions the choice is often between undesirable options only; that is, all options may require expenditure. "When a choice is made from a set of undesirable options, a more vigilance-oriented mind-set is evoked, leading to the elimination of the otherwise robust attraction effect." (Malkoc et al., 2013, p. 318).

With these specific properties that act for and against the attraction effect, it is very difficult to determine which conditions will lead to the attraction effect occurring in financial product decision making.

Several researchers have investigated the attraction effect in financial product decisions (e.g., Herne, 1999 – monetary gambles; Schwarzkopf, 2003 – investment; Zhumakadyr uulu & König, 2014 – loans). Yet many issues in this domain have yet to be resolved. Perhaps the main question to consider first is whether, when making a financial choice, the decision maker does all the necessary calculations to determine which option is most beneficial. The complexity of some financial products makes choosing difficult, as rationally calculating the optimal option requires the decision maker to have a certain cognitive capacity and willingness to make the effort. At least two scenarios are possible in this situation. In the first scenario, the decision maker is not willing to spend the time and effort on a tiring financial task and this enhances the likelihood of the attraction effect occurring (Pocheptsova, Amir, Dhar, & Baumeister, 2009). In the second scenario, the decision maker may engage the cognitive processes to a greater extent, and this eliminates the attraction effect. We are inclined to believe that because people rarely follow the rules of rational decision making in general, attraction effect heuristics are also used to decide financial issues and this tendency is boosted when the person has to make a decision concerning a less well-known and more complex financial product. In this study our aim is to fill a gap in the literature by examining the role of attribute importance in financial product choices.

The Research

Our aim was to investigate the attraction effect in decisions where participants have to choose between options involving financial products. We looked specifically at how much people care about the attributes of the options and how this affects the attraction effect both when choosing a more familiar financial product and a less familiar one. Our hypotheses inquires whether the subjective weighting of attributes can be used as a clue in making complicated financial choices.

If one of the attributes has greater subjective importance, this gives the decision maker an easy-to-use instruction manual for choosing the option that does best on the preferred attribute. Hence we formulated our first hypothesis as:

H1. Participants who regard one of the attributes of a financial product as more important are less prone to yield to the attraction effect as they choose the option that does best on the more important attribute, regardless of the decoy.

Participants who consider both attributes of a financial product to be equally important have to choose between two relatively equal options. Since making a trade-off is demanding, these participants will be prone to reduce task complexity and use heuristics, in this case the attraction effect. As they do not have a clear preference for some of the options, it will be possible to observe the attraction effect in their choices:

H2. Participants who regard both attributes of a financial product to be equally important are more prone to yield to the attraction effect as the decoy influences their choice.

We specified Hypothesis 3 as follows:

H3. Participants who regard both attributes of a financial product to be equally important yield to the attraction effect more when choosing less familiar financial products compared to choosing more familiar products.

Method

Participants

A total of 487 undergraduate students at Slovak universities completed the experiment without financial incentive. Participants were asked by their lecturers if they would be willing to answer some questions and make a few financial decisions for extra credit points.

Undergraduates were selected for several reasons. Given the characteristics of the financial products, we needed a sample that was cognitively capable of understanding what was expected, and able to compare the options and make the choice with some cognitive effort. The second reason for using university students was that we wanted to test the effects of familiarity and engagement with financial products. Some financial products (such as risky stock investment) remain unknown to students in early adulthood, but they understand the nature of the decision-making task.

The age of the participants ranged from 18 to 26 years old (M = 21.65, SD = 2.49). Women accounted for 72% of the sample. We performed a power calculation before data collection. The a priori G power calculation for Chi-square test, comparing two groups, with expected medium effect size (w = 0.3, α error probability = .05), produced a sample size of 145 respondents.

Of the 487 students, 206 (42.3%) were studying management or economics, 168 (34.5%) social science and 113 (23.2%) natural or technical science.

Design and Materials

In the experiment the participants were randomly assigned to one of three between-subject conditions. They then indicated their choices in the financial decision tasks. The control group (N = 163) had only two options in the

decision tasks (A, B), Experimental Group 1 (N = 188) had three options (A, B, Ca) and Experimental Group 2 (N = 136) had three options (A, B, Cb). The options were not randomized in the groups. Respondents were assigned to the groups randomly by MS Excel random number generator.

Six different products were presented in the financial decision tasks and participants were asked to choose one option in each of the six tasks. The products, which we assumed the participants would be more familiar with, were Mobile Rate, Current Account, and Savings Account; the expected less familiar products were Loan, Car Accident Insurance, and Investment. Each option was described using the two attributes most widely promoted in "real" world financial marketing. The descriptions of all the financial products and the attribute values used in the decision tasks are given in Table 1 in the Appendix. The control group's options differed in that one option (A) did better on Attribute 1, and the second option (B) did better on Attribute 2 (core set of options). The financial values of the attributes were calculated before the stimuli materials were created as we wished to produce two de facto equal options (A and B) comparable in values with none being evidently inferior to the other. For example, the difference in the financial value of Option A and Option B in the Loan task is only a few percentage points (3.4%). Therefore, the stimuli material allowed respondents to make their decisions according to their preferences rather than being "pushed" into choosing the evidently financially better option of the two, especially where the second one had an evidently inferior financial value. This principle that the options should be comparable or equal in total values has been an essential part of the attraction effect stimuli.

In the experimental conditions, we created the third option, the decoys (Ca and Cb), using the same strategy for both attributes: we increased the frequency of the superior attribute of the target, by narrowing the difference in attribute where the target was more advantageous than the competitor. In Experimental Group 1 with the three options A, B, and Ca, the third option was a decoy (Ca), which had one inferior attribute compared to the target option (A). It was expected that this manipulation would encourage participants to choose option A. Similarly, in Experimental Group 2, the participants were presented with the three options A, B, and Cb, where the decoy (Cb) had one inferior attribute compared to target option B. It was expected that this manipulation would encourage participants to choose option B. An example of the decoys created for the Loan financial product for both experimental groups is given in Figures 1a and 1b in the Appendix.

Procedure

All the materials were presented on a computer screen using Google software. The entire task lasted about 15 minutes and participants were allowed to pace themselves. All the participants first answered a few socio-demographic questions and then proceeded to make their financial choices.

After choosing one option for all six financial products, the participants were asked which specific attribute in each option they considered more important. They could choose the first attribute, the second, or both attributes. It is important to note that we obtained the attribute importance ratings after the participants had chosen one of the two or three options, as asking about their preferred attribute in advance could have led them to make a different choice (Ariely & Wallsten, 1995). We asked the following question about attribute importance (this example is for the first financial decision task – the Loan): *When evaluating the attributes of the loan, the attribute I consider more important is a) interest rate; b) processing fee;*

c) interest rate and processing fee are equally important to me.

Afterwards the participants were asked how familiar they were with each of the financial products on a three-point scale (*Not familar; Partly familiar; Very familiar*). We also measured how engaged they were with the financial product, asking them if they were willing to invest time and energy into seeking information when making a decision (*Not willing to devote time and energy; Willing to devote reasonable time and energy; Willing to devote time and energy*).

Measuring the Attraction Effect

The attraction effect was defined as the change in the proportion of participants choosing the two original options after the decoy had been introduced. To measure this change, in our study we compared the proportion of the participants choosing the target option and the competitor in the group without the decoy and in the groups where the decoy had been included in the set of options. Since the proportion choosing the decoy also rises (though minimally), the technique for calculating the effect varies depending on whether the proportion selecting the decoy is included. Based on a detailed description of computational techniques in Malaviya and Sivukumar's study (1998) and the most commonly used techniques (e.g., Lee et al., 2016; Malkoc et al., 2013; Pocheptsova et al., 2009), we decided to compute the difference in target choices when the decoy was absent and when it was present, without normalizing the target's share to account for the decoy share. Therefore, the participants who chose the decoy option were not included in the comparisons. The advantage of this procedure is that there is no need to adjust the proportion who selected the decoy, and the direct absolute measure of the effect can be obtained. Malaviya and Sivukumar (1998) used and compared the various measures of the attraction effect and found that interpretations of the outcomes on the four measures were generally similar. The attraction effects were calculated separately for each product.

Results

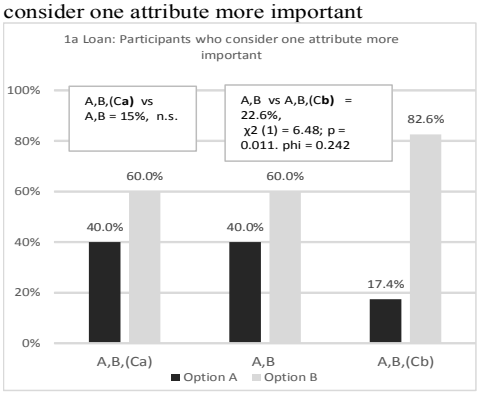
For all the financial products, one third to two thirds of the participants considered the two attributes of the product to be equally important. Those participants who preferred one attribute over the other mostly agreed on which attribute they preferred in five of the six products. The attribute considered more important served as the target-attribute for our participants and was used in the attraction effect calculations. The subjective importance of the attributes (%) is given in Table 2 in the Appendix.

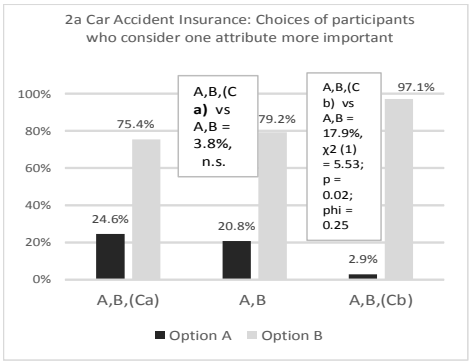
To test our hypotheses, we compared the choices of participants who considered both attributes equally important with those of participants who considered the target-attribute more important. We calculated the manifestation of the attribution effect for the less important attribute as well; however, these results were either not significant or the number of participants was not suitable for analysis.

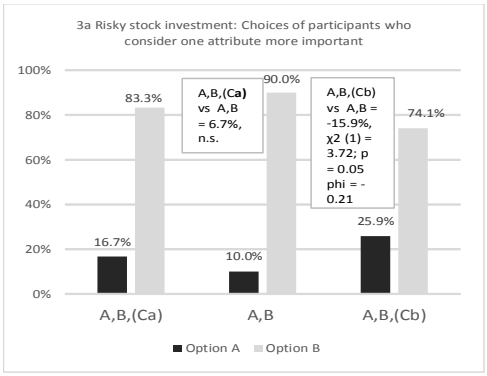
We produced descriptive results and performed a separate pairwise comparison analysis to show whether there was a statistically significant difference between the proportion choosing option A and the proportion choosing option B in the control group and in the two experimental groups. The results concerning whether the attraction effect could be seen in the participant choices according to whether there was preference for one attribute versus no preference are given separately for each product in Graphs 1a, 1b to 3a, 3b for the less familiar and Graphs 4a, 4b to 6a, 6b more familiar products.

As already mentioned, we decided to calculate the attraction effect as the difference in target-option choices in the groups with the decoy and

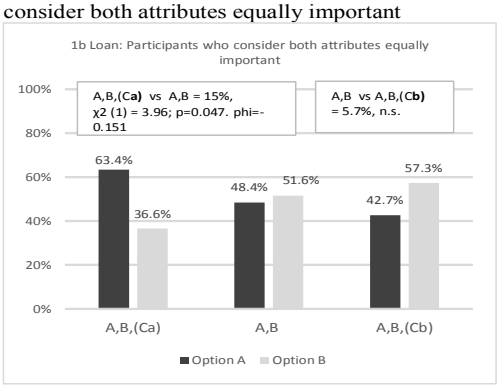
Below: three graphs representing the three less familiar financial products for participants who consider one attribute more important

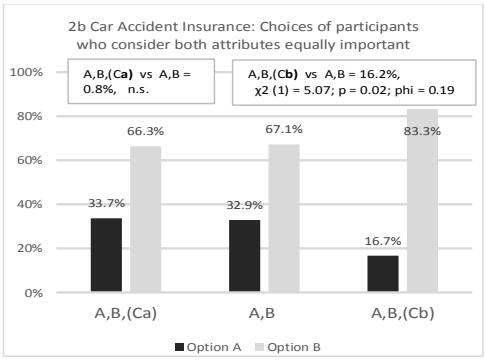


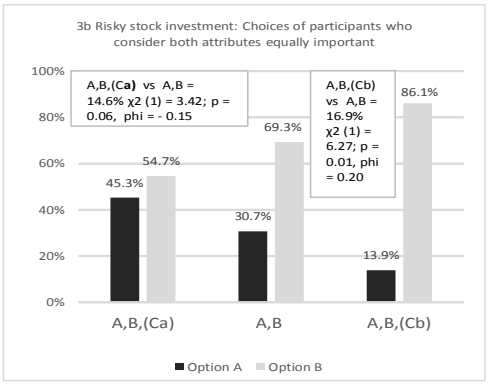




Below: three graphs representing the less familiar financial products for participants who consider both attributes equally important

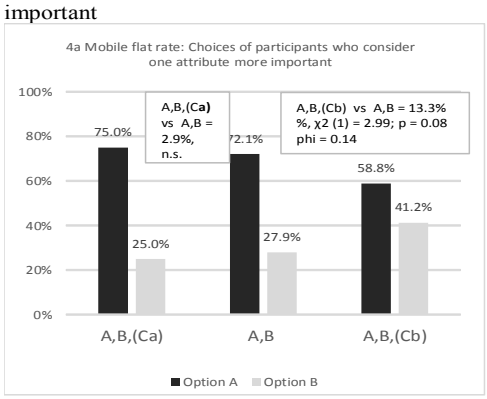


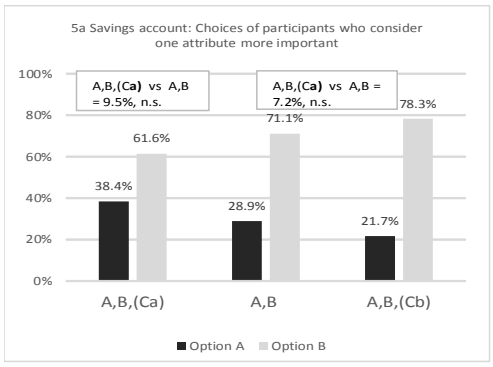


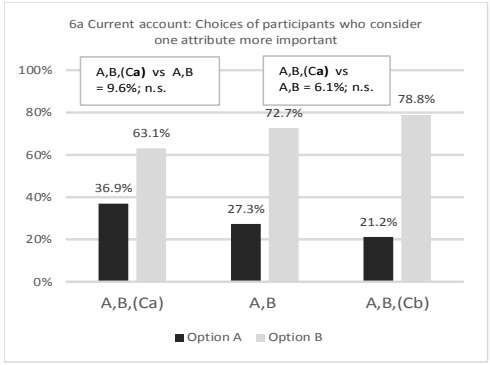


Graphs 1a, 1b, 2a, 2b, 3a, and 3b. The attraction effect in the participant choices according to whether there was preference for one attribute versus both attributes equally important for the less familiar products.

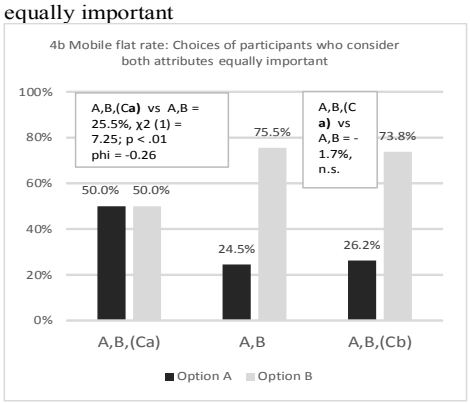
Below: three graphs which represent three more familiar financial products for participants who consider one attribute more important

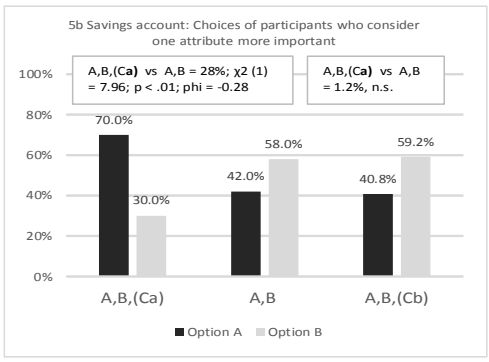


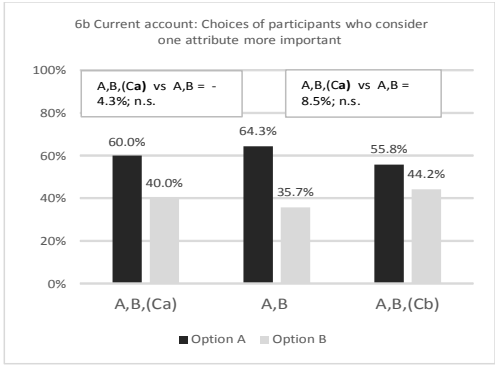




Below: three graphs which represent three more familiar financial products for participants who consider both attributes equally important







Graphs 3a, 3b, 4a, 4b, 5a, and 5b. The attraction effect in the participant choices according to whether there was preference for one attribute versus both attributes equally important for more familiar products.

without the decoy, without normalizing the target share to account for the decoy share. Therefore, the participants who chose the decoy option were not included in the comparisons and are not shown in the results graphs[1].

The Graphs 1a, 1b to 6a, 6b show the graphical representation and calculations of the control and experimental group choices and the attraction effect. Where the results were not significant, we have not reported the chi square or effect sizes, just the percentage value. Values close to significance are reported as well. A positive value means that the decoy was effective, but in most cases it was not significant. In all cases, the decoy led to an increase in the target-option being selected except in the case of Risky Stock Investment, where a repulsion effect was observed.

For the participants who considered both product attributes to be important, the attraction effect was observable in five of the six financial products. It was not observed in the Current Account product. For those participants who preferred one attribute, the attraction effect was observed in two of the six financial products. In both cases, the decoy increased the attractiveness of the option that did better on the attribute considered more important by the participants. Both of the products where the attraction effect was observed were less familiar and had long-term consequences (Loan and Car Accident Insurance).

As expected, the participants had a high degree of engagement with almost all the financial products. Above 87% reported a willingness to spend time and energy choosing and purchasing each of the financial products. Good familiarity with the most frequently used financial products (Mobile Rate, Current Account, and Savings Account) was reported by 84–95% of the participants. Participants reported being less familiar with the less frequently used financial products (Loan 76%, Car Accident Insurance 56%, and Investment 39%). The product students were least familiar with was the Risky Stock Investment, with 61% of participants being unfamiliar with it. The exact data on participants' familiarity and engagement with the financial products is shown in Table 3 in the Appendix.

To test Hypothesis 3 we compared the choices of the less familiar and more familiar financial products with respect to whether the participants considered one or both attributes to be important. For the more familiar products (Mobile Rates, Current Accounts and Savings Accounts), the attraction effect was observed only in participants considering both attributes equally important. This applied to two of the three products (Mobile Rates and Savings Accounts). For those participants who considered one attribute to be important, the attraction effect did not appear in any of the three familiar products.

For the less familiar products (Loan, Car Accident Insurance, and Investment), the attraction effect was observed for all three products for participants preferring two attributes. For participants preferring one attribute, the attraction effect was observed in relation to two of the three products (Loan and Car Accident Insurance), for the third product (Investment), the attraction effect occurred in reverse, manifesting as the repulsion effect. Concerning Hypothesis 3, the results showed that participants who considered two attributes to be equally important yielded more to the attraction effect in both familiar and less familiar products compared to those who prefered only one attribute.

[1] The decoy was chosen by 5–11% of the participants; the smallest proportion selected the decoy in the Mobile Flat Rates task (4.4–5.9% of participants) and the largest proportion selected the decoy in the Car Accident Insurance task (10–11% of participants). These figures are similar to those obtained in most of the attribution effect studies (for example, Malaviya & Sivakumar, 1998; Malkoc et al., 2013; Mishra et al., 1993; Simonson, 1989).

Discussion

In our study, we attempted to verify the assumption that the attraction effect occurs when a financial product is being selected under the condition of a preference for both or one of the attributes of the financial product. Unlike other studies in this field, our design included questions about the importance participants assigned to the attributes of each option. This allowed us not only to test whether the attraction effect occurred, but also whether it was present when participants considered the attributes of the options to be more or less or equally important. We also attempted to establish participants' familiarity and engagement with the financial products.

The main result of our study is that when people care more about one attribute than the other the attraction effect plays less of a role in financial product choices. The attraction effect was present more frequently in decision making when participants had no clear attribute preference. Those with a clear preference for one attribute succumbed to the effect only when choosing a product they were less familiar with (and which had serious potential consequences). We can consider our hypotheses to be supported. These results have brought new insights as the two moderators of the attraction effect (familiarity and attribute preference) have not been extensively studied so far.

The attraction effect occurs when people use the dominance relationship as a heuristic to avoid trade-offs between attributes (Hedgcock & Rao, 2009), and our participants tended to do this in their financial choices. Our findings that the attraction effect was less likely to occur when participants had a preference for one attribute and more likely to occur when they considered both attributes important support Ariely and Wallsten's theory (1995) that the conflict inherent in choosing between options that differ in two similarly weighted attributes helps to create the conditions for context-dependent choices to occur. Our results also indicate that the greater subjective importance of one attribute reduces the attraction effect because the preference of one attribute over the other serves as a guide when choosing a financial product.

From one to two thirds of our participants considered both attributes of the financial products to be equally important. A similar proportion of participants considered one attribute more important. This can be labeled as the focal or critical attribute. It was a target-attribute for our participants, and we assume that it served as an important guide to them when thinking about which financial product to choose.

The financial task that was least familiar to our participants was the Risky Stock Investment. In participants' decision on this product, we identified a repulsion effect, which is the corollary to the attraction effect. The latter occurs when a third inferior option is added to a set of two non-dominating options. In this scenario the share of participants choosing the "dominating" option (referred to as "the target") increases. If the addition of the third inferior option increases the share of participants choosing the non-dominating option (referred to as "the competitor"), we refer to this as the repulsion effect (Simonson, 2014).

Where the Investment Stocks were concerned, those participants who considered the risk of making a loss (target-attribute) more important than the return, the addition of the decoy reducing the risk of loss led to an increase in the proportion choosing the option that did better on the return, which was the competitor. The decoy did not therefore have the intended effect, as it elicited the repulsion effect rather than the attraction effect. According to Simonson (2014), it is the nature of the relationship between the two adjacent options that moderates the resulting effect. We could interpret this phenomenon as entering loss aver-

sion into the interaction between the decoy and the gain and risk attributes of the investment. According to the prospect theory (Tversky & Kahneman, 1991), there are differences in sensitivity to gain and loss, and there are also individual differences in risk tolerance, and these were manifested in participants' preferences for the risk or gain attributes or both.

A practical implication of our findings is that expressing the attribute values numerically does not prevent biases in consumer decisions on financial products when the total values of the options are difficult to calculate. Another implication may be that if consumers have a clear preference for an attribute of a financial option, the attraction effect may be less strong than when there is no attribute preference.

Asymmetric dominance is not the only or even the main factor in any choice. There is no doubt that a choice of an option from the set of options is not determined only by set configuration. The driving forces behind product choices can be product type, the importance and nature of the attributes on which the options differ, the attribute values, decision maker characteristics, and various other factors (Simonson, 2014).

Limitations

There are several limitations of this study. First limitation of the research could be that the attraction effect was tested using university students, and they may be more familiar with some but not with all of the financial products. Students were less knowledgeable about some financial products, such as the Loan or Car Accident Insurance, but they were well informed about Mobile Rates and Current Accounts. Another limitation is that university students are better cognitively predisposed than the representative sample of the adult population for solving the tasks we used, therefore, our results could differ from the general adult population. Third possible limitation is that financial products are specific decision tasks, which are closely related to the country of origin and its financial background, financial culture, financial education, level of development, regulation and law and are not easily transferable to other countries without proper adjustment. Purely Slovak sample describes the situation in Slovakia, which could differ from countries.

Conclusion

The impact of attribute importance on the attraction effect remains unclear, especially with regards the financial domain. We consider our study to be a pilot study on attraction effect in the financial domain, which can provide a basis for further research. Our findings indicate that under certain conditions the attraction effect and repulsion effect can be observed in the financial domain as well. However, more specific and detailed research exploring the effect under combinations of conditions of familiarity versus engagement versus attribute importance is needed to determine how the attraction effect works. Our study may also provide inspiration to create a more elaborate model of the conditions in which one can expect the attraction effect to occur in the financial domain.

References

Ariely, D., & Wallstein, T. S. (1995). Seeking subjective dominance in multidimensional space: An explanation of the asymmetric dominance effect. *Organizational Behavior And Human Decision Processes, 63*(3), 223–232. doi: 10.1006/obhd.1995.1075

Dhar, R., & Simonson, I. (2003). The effect of forced choice on choice. *Journal of Marketing Research, 40*(2), 146–160. doi: 10.1509/jmkr.40.2.146.19229

Frederick, S., Lee, L., & Baskin, E. (2014). The limits of attraction. *Journal of Marketing Research, 51*(4), 487–507. doi: 10.1509/jmr.12.0061

Hedgcock, W., & Rao, A. R. (2009). Trade-off aversion as an explanation for the attraction effect: A functional magnetic resonance imaging study. *Jour-*

nal of Marketing Research, 46, 1–13. doi: 10.1509/jmkr.46.1.1

Herne, K. (1999). The effects of decoy gambles on individual choice. *Experimental Economics*, *2,* 31–40. doi: 10.1023/A:1009925731240

Huber, J., Payne, J. W., & Puto, C. P. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *The Journal of Consumer Research, 9*(1), 90–98. doi: 10.1086/208899

Huber, J., Payne, J. W., & Puto, C. P. (2014). Let's be honest about the attraction effect. *Journal of Marketing Research, 51*(4), 520–525. doi: 10.1509/jmr.14.0208

Lee, C. F., Chuang S. C., Chiu, C. K., & Lan, K. H. (2016). The influence of task difficulty on context effect – compromise and attraction effects. *Current Psychology*, 1–18. Retrieved from: http://link.springer.com/article/10.1007%2Fs12144-016-9428-0. doi: 10.1007/s12144-016-9428-0

Lichters, M., Sarstedt, M., & Vogt, B. (2015). On the practical relevance of the attraction effect: A cautionary note and guidelines for context effect experiments. *Academy of Marketing Science Review*, *5*(1), 1–19. doi: 10.1007/s13162-015-0066-8

Malaviya, P., & Sivakumar, K. (1998). The moderating effect of product category knowledge and attribute importance on the attraction effect. *Marketing Letters*, *9*(1), 93–106. doi: 10.1023/A:1007976305757

Malkoc, S. A., Hedgcock, W., & Hoeffler, S. (2013). Between a rock and a hard place: The failure of the attraction effect among unattractive alternatives. *Journal of Consumer Psychology, 23*(3), 317–329. doi: 10.1016/j.jcps.2012.10.008

Mishra, S., Umesh, U. N., & Stem, D. E. (1993). Antecedents of the attraction effect: An information-processing approach. *Journal of Marketing Research, 30*, 331–349. doi: 10.1177/002224379303000305

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge University Press. doi: 10.1017/CBO9781139173933

Pocheptsova, A., Amir, O., Dhar, R., & Baumeister, R. F. (2009). Deciding without resources: Resource depletion and choice in context. *Journal of Marketing Research*, *46*(3), 344–355. doi: 10.1509/jmkr.46.3.344

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, *16*, 158–74. doi: 10.1086/209205

Schwarzkopf, D. L. (2003). The effects of attraction on investment decisions. *The Journal of Behavioral Finance, 4*(2), 96–108. doi: 10.1207/S15427579JPFM0402_06

Simonson, I. (2014). Vices and virtues of misguided replications: The case of asymmetric dominance. *Journal of Marketing Research, 51*(4), 514–519. doi: 10.1509/jmr.14.0093

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29*(3), 281–295. doi: 10.1177/002224379202900301

Sivakumar, K., & Cherian, J. (1995). Role of product entry and exit on the attraction effect. *Marketing Letters*, *6*(1), 45–51. doi: 10.1007/BF00994039

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4), 1039–1061. doi: 10.2307/2937956

Tversky, A., & Simonson, I. (1993). Context-dependent preferences. *Management Science, 39*(10), 1179–1189. doi: 10.1287/mnsc.39.10.1179

Wedell, D. H., & Pettibone, J. C. (1996). Using judgments to understand decoy effects in choice. *Organizational Behavior and Human Decision Processes, 67*(3), 326–344. doi: 10.1006/obhd.1996.0083

Yang, S., & Lynn, M. (2014). More evidence challenging the robustness and usefulness of the attraction effect. *Journal of Marketing Research, 51*(4), 508–513. doi: 10.1509/jmr.14.0020

Zhumakadyr uulu, K., & Koenig, N. (2014). The role of context effect in the choice of financial products: An experimental investigation in the Kyrgyz Republic. *Central Asia Business Journal, 6*(1), 5–20.

Appendix

Figures

Figure 1a Loan – Experimental Group 1

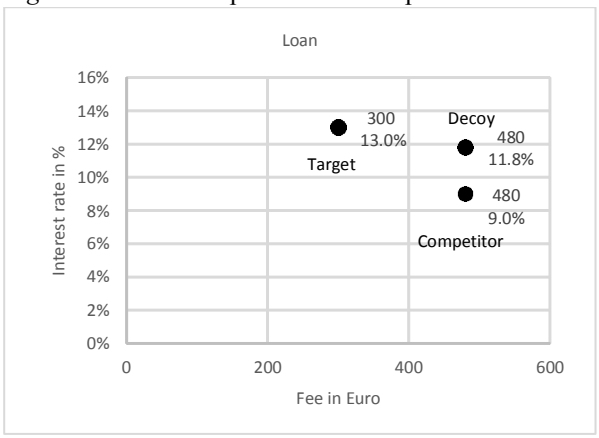


Figure 1b Loan – Experimental Group 2

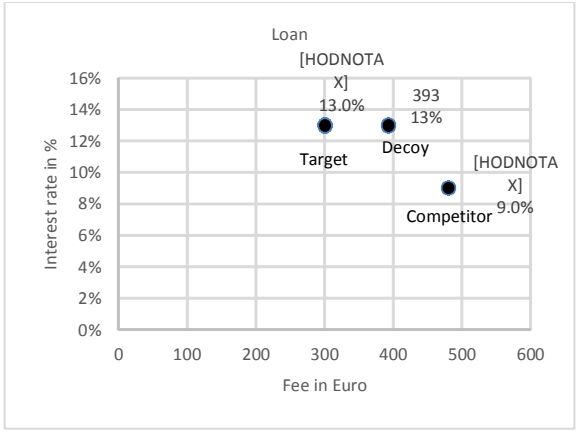


Figure 1a, 1b. Example decoys created for the Loan financial product for both experimental groups

Tables

Table 1 *Descriptions of financial products and attribute values used in the decision tasks*

PRODUCT	Attribute 1	Attribute 2	PRODUCT	Attribute 1	Attribute 2
Loan	Interest rate (%)	Processing fee (€)	**Car accident insurance**	Yearly insurance payment (€)	Accident repair fee (€)
Option A	13%	€300	Option A	€250	€180
Option B	9%	€480	Option B	€130	€300
Decoy Ca	13%	€393	Decoy Ca	€250	€235
Decoy Cb	11.80%	€480	Decoy Cb	€170	€300
Risky stock investment	Estimated yield (%)	Probability of losing all money invested (%)	**Mobile flat rate**	Monthly fee (€)	Discount on new mobile phone (€)
Option A	37%	10%	Option A	15	€20
Option B	30%	5%	Option B	19	€110
Decoy Ca	34%	10%	Decoy Ca	17	€20
Decoy Cb	30%	5.40%	Decoy Cb	19	€96
Savings account	Annual interest rate (%)	Early withdrawal fee (%)	**Current account**	Account maintenance fee (€)	Reward for each card payment (%)
Option A	0.90%	4%	Option A	€1.90	0.60%
Option B	1.25%	8%	Option B	€1	0.30%
Decoy Ca	0.90%	4.60%	Decoy Ca	€1.90	0.42%
Decoy Cb	1.07%	8%	Decoy Cb	€1.30	0.30%

Table 2 *Subjective importance of the attributes (%)*

Financial product	Participants preferring Attribute 1	Participants preferring Attribute 2	Participants preferring both attributes equally
Loan	**40.2%** Interest rate	4.9 % Processing fee	54.8%
Car accident insurance	**32.2%** Yearly insurance payment	16.8 % Accident repair fee	50.9%
Risky stock investment	19.3% Estimated yield	**31.4%** Probability of losing all money invested	49.3%
Mobile flat rate	**52.8%** Monthly fee	16% Discount on new mobile phone	31.2%
Savings account	**55%** Annual interest rate	11.3% Early withdrawal fee	33.7%
Current account	**39.8%** Account maintenance fee	19.1% Reward for each card payment	41.1%

Table 3 *Participants' familiarity and engagement with the financial products*

Financial product	Familiarity (%)			Engagement (%)		
	Not familiar	Partly familiar	Very familiar	Not willing to devote time and energy	Willing to devote reasonable time and energy	Willing to devote time and energy
Loan	24.4	65.5	10.1	6.4	32.6	61.0
Car accident insurance	44.4	45.6	10.0	8.6	36.8	54.6
Risky stock investment	61.0	31.2	7.8	12.9	37.8	49.3
Mobile flat rate	4.5	43.1	52.4	5.8	39.6	54.6
Savings account	15.6	57.7	26.7	5.5	38.4	56.1
Current account	5.3	49.5	45.2	3.5	38.4	58.1

Studia Psychologica, Vol. 61, No. 2, 2019, 86-98 doi: 10.21909/sp.2019.02.774

The Relationship Between an Alternative Form of Cognitive Reflection Test and Intertemporal Choice

Jiuqing Cheng, Cassidy Janssen
Department of Psychology, University of Northern Iowa, United States

The cognitive reflection test (CRT) has been popular because it has demonstrated a good predictive validity of a variety of biases in judgment and decision making. Thomson and Oppenheimer (2016) further developed a second version of the cognitive reflection test, CRT-2. Although CRT-2 has been found to be associated with several biases in judgment and decision making, its relationship with intertemporal choice remains unclear. Previous studies have shown that intertemporal choice characterizes the competition between intuition and reflection, and can be predicted by the original CRT. To further validate CRT-2, the present study tests the relationship between CRT-2 and intertemporal choice. The study finds that better performance on CRT-2 is significantly associated with fewer impulsive intertemporal choices in both gain and payment conditions. Moreover, impulsive choices are related to intuitive errors but not non-intuitive errors generated from CRT-2. The study suggests that CRT-2 provides some more items for researchers to select to characterize individual differences in thinking style and judgment and decision making.

Key words: cognitive reflection, CRT-2, intertemporal choice, dual-process theory

The Cognitive Reflection Test (CRT) is a popular test that is used to measure rational thinking and normative choice preference (Frederick, 2005). CRT contains three items, and an iconic item is the famous bat and ball problem: "A bat and a ball cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost?" As one can imagine, a "10 cents" answer appears to be intuitive but nevertheless incorrect. To find the correct answer, the respondent needs to override the intuitive impulse, and perform reasoning deliberately (Frederick, 2005; Kahneman, 2011).

Researchers believe that the CRT responses characterize the interaction between two competing mental processes as defined by the dual-process theory (Frederick, 2005; Kahneman, 2011; Sinayev & Peters, 2016). According to this theory, two processes (systems) exist in our mind: whereas System 1 is fast, intuitive and impulsive; System 2 is slow, deliberative and controlled (Sloman, 1996; Evans, 2008; Kahneman, 2011). To deliver a correct answer on a CRT item, System 2 needs to check, inhibit, and outperform System 1.

The dual-process theory has long been used to address biased judgment and decision-making, and a variety of such biases are linked to System 1's impulse and intuition (Evans, 2008; Kahneman, 2011). Consistently, a series of studies have revealed an association between CRT and biased judgment and decision-making. For example, in the intertemporal choice task, participants with lower CRT scores displayed a stronger preference for the immediate smaller rewards than for the later larger rewards and hence, were more impulsive in their choices

Correspondence concerning this article should be addressed to Dr. Jiuqing Cheng, Bartlett 2068, Department of Psychology, University of Northern Iowa, Cedar Falls, Iowa, United States, 50613. E-mail: jiuqing.cheng@uni.edu

Received September 28, 2018

(Bialek & Sawicki, 2018; Frederick, 2005; Sinayev & Peters, 2016). In the gamble choice task, participants with lower CRT scores exhibited excessive risk-averse, hence they were not able to maximize the potential earning (Frederick, 2005). Additionally, fewer correct answers on CRT were associated with greater conjunction fallacy and base-rate neglect (Hoppe & Kusterer, 2011; Oechssle et al., 2009). Not surprisingly, performance on CRT also correlated with scholastic assessment test (SAT, a popular test used for college admission in the United States) scores and grade point average (GPA, a classical measure to index overall academic performance), both of which require logical reasoning and deliberation (Frederick, 2005; Thomson & Oppenheimer, 2016).

Thus far, the development of CRT has advanced our understanding of judgment and decision making; nonetheless, some concerns have also been raised. For example, Primi, Morsanyi, Chiesi, Donati, and Hamilton (2016) argued that CRT might be too difficult and hence lead to a floor effect particularly in relatively poorly educated populations.

A more significant concern deals with CRT's overexposure. As CRT gains its popularity in research and media report, participants may learn the items and the answers before taking the test. For instance, in Thomson and Oppenheimer (2016, study 1), more than sixty percent of the participants had been exposed to at least one item before the study. The knowledge of the test can artificially inflate the score. In line with this, in Haigh (2016), those who had seen at least one item scored significantly higher than those without any prior knowledge of CRT. Similarly, Białek and Pennycook (2017) analyzed six previously published studies and found that in four studies participants with prior knowledge of CRT obtained a higher score than those who did not have such knowledge.

However, it is worth noting that although prior knowledge of CRT may increase test scores, CRT's predictive ability (its core ability) remains robust. For example, in Białek and Pennycook (2017), even though participants with prior knowledge of CRT scored better, there was no significant difference in CRT's predictive ability (correlations between CRT and other tasks) between experienced and inexperienced participants. Meyer, Zhou, and Frederick (2018) tracked mTurk workers who took CRT repeatedly and found that on average, scores improved by merely 0.024 items per exposure. More importantly, CRT's predictions did not significantly vary with repeated exposure. In Stagnaro, Pennycook, and Rand (2018), CRT was correlated with religious belief measures, and such correlations were stable across years.

Importantly, one recent study provided new insights into the impact of CRT's exposure on its predictive power. Šrol (2018) found that this impact was moderated by the need for cognition. In this study, CRT's predictive ability of performance on heuristics and bias tasks was improved by its exposure only in those with a high level of need for cognition. However, in that sample, only 16% of participants were categorized into the group with a high level of need for cognition. Thus, when combining all participants together, there was no overall difference in CRT's predictions between exposed and unexposed participants. Nonetheless, Šrol (2018) indicated that participants' metacognitive characteristics might moderate how exposure affected CRT's prediction.

Another concern pertains to the confounding effect of numeracy. Sinayev & Peters (2016) proposed and empirically demonstrated that both cognitive reflection and numeracy were needed to generate correct answers for CRT. Numeracy refers to the ability to comprehend and utilize numerical information (Peters & Bjalkebring, 2015; Sinayev & Peters, 2016). According to Sinayev and Peters (2016), to generate a correct answer, participants went through two steps. In the first step, participants needed

to inhibit the intuitive impulse (i.e., cognitive reflection). In the second step, participants engaged in math calculation (i.e., numeracy involvement). Consistent with their hypothesis, Sinayev and Peters (2016) found that the numeracy component, teased apart from the CRT response, could significantly predict judgment and decision-making biases as described above. Thus, the relationship between CRT and judgment and decision-making biases was confounded with numeracy.

Given the concerns, some researchers have introduced modified CRT measures (Baron, Scott, Fincher, & Metz, 2015; Primi et al., 2016; Sirota & Juanchich, 2018; Thomson & Oppenheimer, 2016; Toplak, West, & Stanovich, 2014). For example, to mitigate the potential floor effect, Primi et al. (2016) added three new items and found only a very small proportion of participants answered all items incorrectly. The new version performed well in younger and less educated populations. To address the overexposure problem, Toplak et al. (2014) added four more items to CRT. Sirota and Juanchich (2018) further tested this seven-item version with three formats: open-ended questions, two-option multiple choices, and four-option multiple choices. Both studies found that the extended CRT retained its predictive power, regardless of the question format.

Exploring a Second Version of CRT: CRT-2

Among the modified CRT measures, the present study specifically focuses on CRT-2, which was developed by Thomson and Oppenheimer (2016). We have two reasons. First, compared to the measures that contained both original CRT and new items (Baron et al., 2015; Primi et al., 2016; Sirota & Juanchich, 2018; Toplak et al., 2014), CRT-2 adopts a completely new set of items (specific items are found in the Methods section). Our main goal is to further validate these items by testing the relationship between CRT-2 and intertemporal choice. More broadly speaking, the study aims to further investigate whether CRT-type trick questions can predict biased judgment and decision making. CRT-2 has the potential to provide more items for researchers to select to characterize individual differences in cognition.

Another reason to focus on CRT-2 is that CRT-2 might rely less on (though not exclude) numeracy. First, CRT-2 adopts items that aim to reduce such an effect. As can be seen in the Methods section, among the four items, the first and the third items do not appear to need any computation. Second, in Thomson and Oppenheimer (2016), the correlation between CRT-2 and numeracy was significantly weaker than the correlation between the original CRT and numeracy. Third, as demonstrated in Primi et al. (2016), numeracy was a significant covariate that mediated the gender effect on CRT. That is, the fact that males had better performance on CRT was in part because males performed better on numeracy. In Thomson and Oppenheimer (2016), males scored higher on both CRT and numeracy than did females. However, there was no difference in performance between females and males on CRT-2. Taken together, it is reasonable to believe that CRT-2 might rely less on numeracy than does the original CRT.

In Thomson and Oppenheimer (2016), CRT-2 was correlated with need for cognition, base rate neglect, college GPA, and SAT scores, indicating it could replicate some of the important findings generated by the original CRT. Nevertheless, the study did not find a significant relationship between CRT-2 and intertemporal choice. As described in that article, one reason might be that the intertemporal choice task was not reliable in the study. The low reliability might be because there were only a few items. Moreover, only one relationship reached a statistical significance level when testing the correlation between CRT-2 and each of the

intertemporal choice items separately. We note that with the limited number of items, the task might not be able to capture a stable choice preference.

In the present study, we are interested in clarifying the relationship between CRT-2 and intertemporal choice for two reasons. First, intertemporal choice is related to a series of important life activities and consequences. For example, research has found that more impulsive intertemporal choices are associated with lower income, lower credit score, lower college GPA, and a greater chance of having obesity and abusing substances (de Wit, 2008; Kirby, Winston, & Santiesteban 2005; Meier & Sprenger, 2011; Reimers, Maylor, Stewart, & Chater, 2009; Schiff et al., 2016). Thus, it is of interest to examine a test that can characterize individual differences in intertemporal choice.

Second and more importantly, researchers have demonstrated that making intertemporal choices reflects the competition between System 1 and System 2 as defined by the dual-process theory. For example, McClure, Laibson, Loewenstein, and Cohen (2004) identified two competing brain regions (part of the limbic system vs. dorsal lateral prefrontal cortex) when participants were making different selections in an intertemporal choice task. These two brain regions resembled the characteristics of System 1 and System 2 (e.g., intuition vs. calculation). Additionally, with modeling, Price, Higgs, Maw and Lee (2016) found that intertemporal choice could be well explained by a two-parameter model that depicted the dual-process theory. Moreover, recent studies with mouse-tracking demonstrated that the trajectories were less direct when making less impulsive intertemporal choices, and concluded that participants had to inhibit the temptation of choosing the sooner smaller rewards in order to maximize their benefit in the long run (Cheng & González-Vallejo, 2017; Dshemuchadse, Scherbaum, & Goschke, 2013; Stillman, Medvedev, & Ferguson, 2017). Therefore, testing the relationship between intertemporal choice and CRT-2 helps to illustrate whether CRT-2 captures cognitive reflection (System 1 vs. System 2), as does the original CRT.

Overview of the Present Study

CRT-2 appears to provide some new items that pertain to cognitive reflection and judgment and decision making. Some recent studies combined CRT and CRT-2 and had used the new composite to address honesty, analytical thinking style, and attitude toward fake news (Capraro & Peltola, 2018; Pennycook & Rand, 2017; Yilmaz & Saribay, 2017). However, we believe the validity of CRT-2 needs to be addressed before its extensive application.

The present study aims to test the validity of CRT-2 by examining its correlation with intertemporal choice. To address the reliability issue, we employed an intertemporal choice task that was recently employed in other studies (Cheng & González-Vallejo, 2016; Dai & Busemeyer, 2014; Scholten, Read, & Sanborn, 2014). In this task, participants make repeated choices between a sooner, smaller reinforcer and a later, larger reinforcer. With a series of choice pairs, we hope to increase the reliability of the task and to obtain a stable choice preference from participants.

Furthermore, for CRT scoring, most studies so far have used the number of correct responses. Such a scoring method measures cognitive reflection and has demonstrated good predictive ability (Pennycook, Cheyne, Koehler, & Fugelsang, 2015). However, as implied in Pennycook et al. (2015), while greater cognitive reflection may predict more long-term oriented choices, the pattern is different from the concept that intuition can predict more impulsive choices. In other words, for CRT-2, even its correct response could predict intertemporal choice preference, the extent to which CRT-2 measures

intuition in intertemporal choice remains unclear. From the perspective of face validity, if CRT-2 taps into intuitive thinking style, two patterns should be revealed. First, among the errors, there should be at least a portion of intuitive errors. Too few intuitive errors among all errors would indicate that CRT-2 is unable to capture the intuitive thinking style. Second, the intuitive error should be able to predict intertemporal choice preference in the opposite direction predicted by the correct response. Following Pennycook et al. (2015) and Sinayev and Peters (2016), we employ the scoring method with the correct response, intuitive error and other error. For CRT-2, the intuitive and other types of errors can be found in the Methods section. The study aims to further examine whether the performance of CRT-2 is consistent with its face validity regarding both reflective and intuitive thinking styles.

One issue of CRT-2 is its relatively low reliability (Cronbach's α). In Thomson and Oppenheimer (2016), with the same group of participants, CRT-2's reliability was .51, lower than CRT's reliability (.62). In Primi et al. (2016), CRT's reliability was .65. Białek and Pennycook (2017) reviewed six past studies on CRT and found that the reliability ranged from .53 to .76. In Šrol (2018), CRT's reliability was as high as .78. Thus, it appears that for the original CRT, its reliability varies across samples. For CRT-2, it is not clear whether its reliability also varies between studies. More importantly, consistently low reliability would reduce the merit of CRT-2. Thus, the present study tests CRT-2's reliability with a different sample.

It is worth noting that in the majority of studies with intertemporal choice, only the gain condition is adopted. That is, participants make selections between two rewards. In such a condition, excessive preference for the immediate/sooner, smaller rewards over the later, larger rewards is considered being impulsive, and lower CRT scores are supposed to be associated with greater impulsive choices. To obtain a reliable relationship between CRT-2 and intertemporal choice, the present study also employs a payment condition where participants make selections between a sooner, smaller payment and a later, larger payment. In this condition, excessive preference for the later, larger payment over the sooner, smaller payment is regarded as the impulsive choice pattern, because participants have to pay more money in the long run (Cheng, Lu, Han, & González-Vallejo, 2012; Perry & Carroll, 2008). We hypothesize that lower CRT-2 scores and more intuitive errors are correlated with more impulsive choices in both gain and payment conditions.

Methods

Participants

Prior to data collection, this study was approved by the Institutional Review Board (IRB) to ensure it met the ethical guidelines. In the present study, all participants were recruited from the participant pool at the authors' institution. The participant pool was comprised of freshmen and sophomore students who were taking Elementary Psychology. Data collection stopped at the end of the semester when the participant pool was closed. As a result, one-hundred and forty-five college students participated in this study via Qualtrics to receive course credit. Three participants completed fewer than half of the items. Another three completed zero or only one item on CRT. Hence these six participants were removed from the study. In the remaining 139 participants, there were 68 females, 67 males and four did not reveal their gender. We note that this sample size was comparable to the one tested in Thomson and Oppenheimer (2016).

Sensitivity analysis was performed with G*Power 3.1.9 to estimate the effect sizes with the current sample size. α was set at .05 and

statistical power was set at .80. As a result, the study had sufficient power to detect a correlation coefficient of .23 (two-tailed), and differences between two independent means of d = 0.49 (two-tailed, one group had 68 females and the other group had 67 males).

Materials and Procedures

All participants completed CRT-2 and two conditions of intertemporal choice tasks (gains vs. payments), as described below.

CRT-2 scale. Four items of CRT-2 were adopted from Thomson and Oppenheimer (2016, p. 101). To clarify the impact of intuitive error on decision preference, we adopted two kinds of scoring criteria (Sinayev & Peters, 2015; Thomson & Oppenheimer, 2016). The first one simply differentiated the incorrect and correct answers. The second kind not only identified the correct and incorrect answers, but it also teased apart the errors into two categories: intuitive errors and other errors. The items and the scoring keys are listed below. For each item, any answer that is different from the correct or intuitive answer is considered as a non-intuitive incorrect answer.

1. If you're running a race and you pass the person in second place, what place are you in? (intuitive answer: first; correct answer: second)

2. A farmer had 15 sheep and all but 8 died. How many are left? (intuitive answer: 7; correct answer: 8)

3. Emily's father has three daughters. The first two are named April and May. What is the third daughter's name? (intuitive answer: June; correct answer: Emily)

4. How many cubic feet of dirt are there in a hole that is 3' deep x 3' wide x 3' long? (intuitive answer: 27; correct answer: none)

Intertemporal choice tasks. The intertemporal choice task employed in the present study was similar to those reported in some previous studies (Cheng & González-Vallejo, 2016; Scholten et al., 2014). The current study employed two conditions of intertemporal choice tasks with hypothetical gains and payments. In the gain condition, participants were asked to make forty choices between a sooner gain and a more delayed gain. All attributes, including magnitude and delay, varied across all choice pairs. To mimic the earning and payment (for the payment condition) in everyday life where whole numbers rarely occur, in all choice pairs, the magnitude contained two decimal places. As an example, participants were asked to make a choice between \$137.55 in 67 days vs. \$90.29 in 34 days, and then moved to another choice pair: \$205.05 in 55 days vs. \$149.85 in 32 days. Across all choices, the averages of the sooner and later delays were 28.68 and 54.43 days, respectively. The averages of the smaller and larger gains were \$195.97 and \$345.75, respectively.

The delays and magnitudes used in the payment condition were exactly the same as those used in the gain condition. There were two differences between the conditions. First, in the payment condition, participants were asked to make choices between a sooner smaller payment and a more delayed larger payment (as opposed to selecting between gains in the gain condition). Second, the sequences of the choice pairs were different between the two conditions. Doing so aimed to reduce the memory effect so that memory of choices in one condition would not affect choices in the other. In an earlier experiment performed by the authors, upon completing the task, participants were asked whether they noticed that the attributes were the same between the two conditions. None reported affirmatively.

Following previous studies (Cheng et al., 2012; Scholten et al., 2014), the present study employed the proportion of choosing the long-term advantageous options (later larger gain in the gain condition, and sooner smaller payment in the payment condition) to index the choice

preference. A higher proportion in both conditions indicates a less short-sighted (impulsive) choice preference.

Results

Reliability of the Measures

In the current study, when only differentiating correct and incorrect answers, CRT-2's Cronbach's α was .60, with a 95% confidence interval between .48 and .70. When differentiating correct answers, intuitive errors and other errors, CRT-2's Cronbach's α slightly increased to .61, with a 95% of confidence interval between .50 and .71. Given the confidence intervals, such reliability was comparable to the findings in other studies of CRT-2 (Thomson & Oppenheimer, 2016; Yilmaz & Saribay, 2017).

For the gain and the payment conditions of the intertemporal choice task, the Cronbach's α were .93 (95% CI between .91 and .95) and .92 (95% CI between .89 and .93), respectively. Thus, choice preference in the current study was reliable and could be used for further analyses.

Performance of CRT-2

On average, participants answered 2.39 items correctly (59.8% correct rate), with an *SD* of 1.17. As seen in Figure 1, the percentages of participants who gave zero to four correct answers were: 9.4, 12.2, 24.5, 38.1 and 15.8, respectively. Thus, based on the current sample, the distribution of CRT-2 scores was not severely skewed. Moreover, CRT-2 did not meet a floor or ceiling effect. Table 1 further presents the results regarding CRT-2 performance when differentiating intuitive and non-intuitive errors. As can be seen, when participants made errors, the majority errors (73.6%) were intuitive ones.

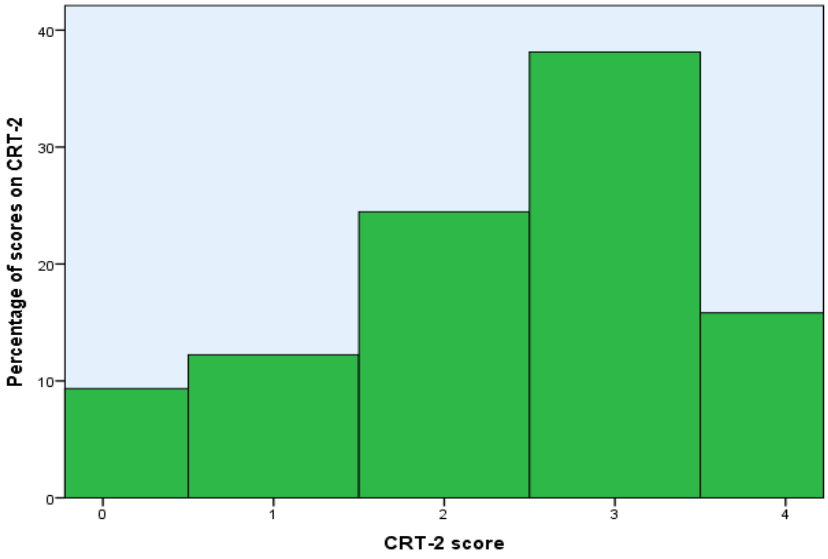


Figure 1 Percentage of different scores on CRT-2.

Table 1 *CRT-2 performance when differentiating intuitive and other errors*

	Item 1	Item 2	Item 3	Item 4
Non-intuitive error (%)	2.9	2.2	3.6	33.8
Intuitive error (%)	30.2	20.1	22.3	46.0
Correct answer (%)	66.9	77.7	74.1	20.1

Table 2 *Pearson correlations between CRT-2 items*

	Item 1	Item 2	Item 3	Item 4
Item 1	--	.28**	.35***	.24**
Item 2		--	.35***	.10
Item 3			--	.30***

Note. ** $p < .01$; *** $p < .001$.

As displayed in Table 1, the last item was more difficult than the other three. Given the different levels of difficulty, one might ask whether including the last item decreased the reliability of CRT-2. This was not the case in the present study, as removing the last item resulted in a Cronbach's α of .60 (95% CI between .46 and .70). Moreover, as shown in Table 2, items displayed significant inter-correlations, with the only exception between Item 2 and Item 4.[1] Thus, all four items should be included in CRT-2.

CRT-2 and Intertemporal Choice

In the gain condition, the mean proportion of choosing the later larger gain over the sooner smaller gain was 0.64 ($SD = 0.25$). In the payment condition, the mean proportion of choosing the sooner smaller payment over later larger payment was 0.67 ($SD = 0.22$). Similar to other studies (Cheng et al., 2012; Estle et al., 2006), there was a trend that participants selected more long-term advantageous options in the payment condition than in the gain condition, $t(138) = 1.64$, $p = .10$, $d = 0.14$, although not statistically significant.

Table 3 shows Pearson correlations between CRT-2 responses and preference of intertemporal choice. As shown, overall CRT-2 performance and intuitive error were significantly related to choice preference in both of the gain and payment conditions. Following Lee and Preacher (2013), Fisher's z test was applied to examine whether the correlation strength was significantly different between when using CRT-2 total score and when using intuitive error to predict choice preference. In the gain condition, there was no significant difference between the two correlations, Fisher's $z = 1.25$, p(two-tailed) = .212. A similar non-significant pattern was also found in the payment condition, Fisher's $z = 1.05$, p(two-tailed) = .295. Thus, CRT-2 total score and intuitive error had a similar predictive ability on choice preference in both gain and payment conditions.

Contrary to CRT-2 total score and intuitive error, error due to non-intuitive reasons was not associated with choice preference in either condition. The non-intuitive error was not related

[1] For all correlations in the present study (Tables 2 and 3), there was little difference in correlation coefficients between when using Pearson correlation and Spearman correlation. The significance of the correlations remained the same when using either type of the correlation.

to the intuitive error, either. We did not apply Fisher's *z* test to compare the predictive ability between intuitive error and other error because the latter one simply could not predict choice preference.

Gender Effect

Table 4 exhibits the comparisons on CRT-2 and choice preference between female and male participants (those who did not report gender were excluded in this section). Similar to Thomson and Oppenheimer (2016), there was no difference in any of the CRT responses between females and males. Additionally, there was no gender effect on intertemporal choice preference.

Discussion

The present study examined the relationship between CRT-2 and intertemporal choice. The overall performance on CRT-2 (e.g., average total score and inter-correlations between items) was comparable between the present study and Thomson and Oppenheimer (2016). Primi et al.

Table 3 *Pearson Correlations between CRT-2 responses and preference of intertemporal choice*

	CRT correct rate	Intuitive error	Other error	Proportion of LL	Proportion of SS
CRT correct rate	--	-.83***	-.42***	.29**	.31***
Intuitive error		--	-.16	-.23**	-.26**
Other error			--	-.15	-.12
Proportion of LL				--	.56***

Note. * $p < .05$; ** $p < .01$; *** $p < .001$.
CRT correct rate: number of correct items out of 4.
Intuitive error: proportion of intuitive errors (out of 4).
Other error: proportion of other errors (out of 4).
Proportion of LL: the proportion of choosing the later larger gain in the gain condition.
Proportion of SS: the proportion of choosing the sooner smaller payment in the payment condition.

Table 4 *Gender effect on CRT-2 and choice preference*

Item	Females *Mean* (*SD*)	Males *Mean* (*SD*)	*t*-test (*df* = 133)
CRT-2 Item 1	0.69 (0.47)	0.66 (0.48)	$t = 0.42$, $p = .672$, $d = 0.07$
CRT-2 Item 2	0.74 (0.44)	0.81 (0.40)	$t = -0.97$, $p = .333$, $d = 0.17$
CRT-2 Item 3	0.74 (0.44)	0.75 (0.44)	$t = -0.14$, $p = .885$, $d = 0.02$
CRT-2 Item 4	0.16 (0.37)	0.24 (0.43)	$t = -1.12$, $p = .267$, $d = 0.19$
CRT correct rate	0.58 (0.27)	0.61 (0.32)	$t = -0.62$, $p = .540$, $d = 0.11$
Intuitive error	0.31 (0.26)	0.28 (0.29)	$t = 0.62$, $p = .537$, $d = 0.11$
Other error	0.11 (0.13)	0.10 (0.20)	$t = 0.08$, $p = .941$, $d = 0.01$
Proportion of LL	0.63 (0.24)	0.65 (0.25)	$t = -0.29$, $p = .776$, $d = 0.05$
Proportion of SS	0.70 (0.19)	0.65 (0.24)	$t = 1.36$, $p = .176$, $d = 0.24$

(2016) concerned a potential floor effect for the original CRT. As illustrated in Figure 1, less than 10% of participants answered all items of CRT-2 incorrectly. Meanwhile, 15.8% of participants answered all items of CRT-2 correctly. Hence, the study did not detect any obvious floor or ceiling effect, indicating the CRT-2's difficulty appeared to be appropriate for college students.

Compared to Thomson and Oppenheimer (2016) and Yilmaz and Saribay (2017), the internal consistency of CRT-2 in the present study was similar (when taking 95% confidence internal into account). As stated, at the apparent level, the first and the third item in CRT-2 did not need any computation, whereas the other two items were more related to mathematics. Thus, the inconsistency between the items' relationship with mathematics might decrease CRT-2's internal consistency. While a Cronbach's α of .60 was far from being perfect, it was still close to CRT's Cronbach's α in some studies as cited earlier. Hence, we believe CRT-2's internal consistency should not be a fundamental problem that prevents its future usage.

The present study computed three scores: CRT-2's total score (i.e., the correct answer rate), the percentage of intuitive errors, and the percentage of other errors. Similar to Thomson and Oppenheimer (2016), the majority of errors were intuitive errors. Moreover, there was no significant relationship between intuitive errors and other errors. Thus, intuitive errors and other errors appeared to capture different constructs of thinking style.

Most importantly, the present study employed a reliable intertemporal choice task and found that more CRT-2 corrected responses were significantly related to fewer impulsive intertemporal choices in both gain and payment conditions. Additionally, we also found that intuitive errors but not other errors were significantly positively related to impulsive choice preference. Furthermore, the strength of the correlation between choice preference and CRT-2 correct responses was similar to the strength of the correlation between choice preference and intuitive errors. The similar predictive ability between the correct responses and intuitive errors might be due to the fact that the intuitive errors accounted for 73.6% of total errors.

The findings stated above had a few implications. First, in addition to the correct responses, intuitive errors could also predict impulsive preference in intertemporal choices. By contrast, non-intuitive errors were not able to do so. While we admit that both CRT-2 and intertemporal choice tap into a variety of psychological constructs such as general intelligence and numeracy, we believe the current findings generated by CRT-2 are at least consistent with the notion of cognitive reflection and intuitive thinking style. In other words, the performance of CRT-2 was in line with its face validity. To more clearly demonstrate that CRT-2 can capture cognitive reflection and intuition, in future studies, more CRT-type scales, thinking style scales (for example, the Faith in Intuition scale used in Pennycook et al., 2015), and judgment and decision making tasks are needed for cross-validation. Additionally, the study implied that for CRT and other similar scales, to examine their validity, researchers can go beyond the total score (i.e., the number of correct responses). The percentage of intuitive errors and the relationship between intuitive errors and other behavioral tasks should also be tested.

Combined with previous findings in Thomson and Oppenheimer (2016), the present study implied that CRT-2 could provide some more valid items for researchers to characterize individual differences. In a broader sense, the present study suggested that in addition to the three original CRT items, *CRT-type questions* generally have good predictive power of biased judgment and decision making.

Limitations of the present study should also be addressed. First, we did not directly ask participants whether they had seen any of the CRT-

2 items before. Thus, we could not illustrate to what extent CRT-2 was free of prior experience. Second, Thomson and Oppenheimer (2016) found that compared to CRT, CRT-2's correlation with objective numeracy scales was weaker. While teasing apart numeracy is appealing, the current study did not measure numeracy. Similar to Thomson and Oppenheimer (2016), the present study found that there was no gender effect on CRT-2, inciting that CRT-2 seemed to be more gender neutral than the original CRT. Nonetheless, the gender effect on the original CRT may have resulted from not only objective numeracy (numerical skills) but also math anxiety, self-efficacy, and rational thinking (Primi, Donati, Chiesi, & Morsanyi, 2018; Ring, Neyse, David-Barett, & Schmidt, 2016; Sladek, Bond, & Phillips, 2010; Zhang, Highhouse, & Rada, 2016). Thus, the present study simply replicated the non-significant gender effect on CRT-2. However, we believe such a pattern did not provide sufficient insight into the relationship between CRT-2 and numeracy. Hence, future studies are needed to clarify whether CRT-2 is less affected by objective and/or subjective numeracy. Recently, a new version of CRT (termed verbal CRT) based on non-mathematical problems was developed. This version has a weaker relationship with numeracy and is more gender neutral (Sirota, Kostovičová, Juanchich, Dewberry, & Marshall, 2018). We believe developing such a version is the right step to tease apart cognitive reflection and numeracy.

The third limitation pertains to the study's external validity. The current study employed college students from a participant pool. Although with such a sample, CRT-2 performed well, we realize that further studies are needed to examine whether CRT-2 can also be applied to populations with different ages and education levels.

In sum, the present study reveals that with a reliable intertemporal choice task, CRT-2's correct response and intuitive errors are able to predict choice preference in both gain and payment contexts. The study suggests that CRT-2 provides some more items for researchers to select to characterize individual differences in thinking style and judgment and decision making.

References

Baron, J., Scott, S., Fincher, K. S., & Metz, S. E. (2015). Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition, 4*(3), 265–284. doi: 10.1016/j.jarmac.2014.09.003

Bialek, M., & Pennycook, G. (2017). The Cognitive Reflection Test is robust to multiple exposures. *Behavior Research Methods, 50*(5),1953–1959. doi: 10.3758/s13428-017-0963-x

Bialek, M., & Sawicki, P. (2018). Cognitive reflection effects on time discounting. *Journal of Individual Differences, 39*(2), 99–106. doi: 10.1027/1614-0001/a000254

Capraro, V., & Peltola, N. (2018). Lack of deliberation drives honesty among men but not women. Retrieved from SSRN: https://ssrn.com/abstract=3182830. doi: 10.2139/ssrn.3182830

Cheng, J., Lu, H., Han, X., González -Vallejo, C., & Sui, N. (2012). Temporal discounting in heroin-dependent patients: No sign effect, weaker magnitude effect, and the relationship with inhibitory control. *Experimental and Clinical Psychopharmacology, 20*(5), 400–409. doi: 10.1037/a0029657

Cheng, J., & González-Vallejo, C. (2016). Attribute-wise mechanism vs. alternative-wise mechanism in intertemporal choice: Testing the proportional difference model, trade-off model and hyperbolic model. *Decision*, *3*(3), 190–215. doi: 10.1037/dec0000046

Cheng, J., & González-Vallejo, C. (2017). Action dynamics in intertemporal choice reveal different facets of psychological states. *Journal of Behavioral Decision Making, 30*(1), 107–122. doi: 10.1002/bdm.1923

Dai, J., & Busemeyer, J. R. (2014). A probabilistic, dynamic, and attribute-wise model of intertemporal choice. *Journal of Experimental Psychology: General, 143,* 1489–1514. doi: 10.1037/a0035976

Dai, J., Pleskac, T. J., & Pachur, T. (2018). Dynamic cognitive models of intertemporal choice. *Cognitive Psychology, 104,* 29–56. doi: 10.1016/j.cogpsych.2018.03.001

de Wit, H. (2008). Impulsivity as a determinant and consequence of drug use: A review of underlying pro-

cesses. *Addiction Biology*, *14*, 22–31. doi: 10.1111/j.1369-1600.2008.00129.x

Dshemuchadse, M., Scherbaum, S., & Goschke, T. (2013). How decisions emerge: Action dynamics in intertemporal decision making. *Journal of Experimental Psychology: General, 142,* 93–100. doi: 10.1037/a0028499

Estle, S. J., Green, L., Myerson, J., & Holt, D. D. (2006). Differential effects of amount on temporal and probability discounting of gains and losses. *Memory & Cognition, 34*, 914–928. doi: 10.3758/BF03193437

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment and social cognition. *Annual Review of Psychology, 59*, 255–278. doi: 10.1146/annurev.psych.59.103006.093629

Frederick, S. (2005). Cognitive reection and decision making. *Journal of Economic Perspectives, 19(4)*, 25–42. doi: 10.1257/089533005775196732

Haigh, M. (2016). Has the standard Cognitive Reflection Test become a victim of its own success? *Advances in Cognitive Psychology, 12*, 145–149. doi: 10.5709/acp-0193-5

Hoppe, E. I., & Kusterer, D. J. (2011). Behavioral biases and cognitive reection. *Economics Letters, 110*(12), 97–100. doi: 10.1016/j.econlet.2010.11.015

Kahneman, D. (2011). *Thinking fast and slow.* New York: Farrar, Strauss, Giroux.

Kirby, K. N., Winston, G. C., & Santiesteban, M. (2005). Impatience and grades: Delay-discount rates correlate negatively with college GPA. *Learning and Individual Differences 15,* 213–22. doi: 10.1016/j.lindif.2005.01.003

Lee, I. A., & Preacher, K. J. (2013). Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Retrieved from http://quantpsy.org/corrtest/corrtest2.htm

McClure, S. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, *306*(5695), 503–507. doi: 10.1126/science.1100907

Meier, S., & Sprenger, C. D. (2012). Time discounting predicts creditworthiness. *Psychological Science, 23*(1)*,* 56–58. doi: 10.1177/0956797611425931

Meyer, A., Zhou, E., & Frederick, S. (2018). The non-effects of repeated exposure to the Cognitive Reflection Test. *Judgment and Decision Making, 13*, 246–259. http://journal.sjdm.org/18/18228a/jdm18228a.html

Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization*, *72*(1), 147–152. doi: 10.1016/j.jebo.2009.04.018

Pennycook, G., & Rand, D. G. (2018). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. Retrieved from https://ssrn.com/abstract=3023545. doi: 10.2139/ssrn.3023545

Perry, J. L., & Carroll, M. E. (2008). The role of impulsive behavior in drug abuse. *Psychopharmacology (Berl)*, *200*(1), 1–26. doi: 10.1007/s00213-008-1173-0

Peters, E., & Bjalkebring, P. (2015). Multiple numeric competencies: When a number is not just a number. *Journal of Personality and Social Psychology, 108*(5), 802–22. doi: 10.1037/pspp0000019

Price, M., Higgs, S., Maw, J., & Lee, M. (2016). A dual-process approach to exploring the role of delay discounting in obesity. *Physiology & Behavior, 162*, 46–51. doi: 10.1016/j.physbeh.2016.02.020

Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the Cognitive Reflection Test applying item response theory (IRT). *Journal of Behavioral Decision Making, 29*, 453–469. doi: 10.1002/bdm.1883

Primi, C., Donati, M., Chiesi, F., & Morsanyi, K. (2018). Are there gender differences in cognitive reflection? Invariance and differences related to mathematics. *Thinking & Reasoning, 24*(2), 258–279. doi: 10.1080/13546783.2017.1387606

Reimers, S., Maylor, E. A., Stewart, N., & Chater, N. (2009). Associations between a one-shot delay discounting measure and age, income education and real-world impulsive behavior. *Personality and Individual Differences, 47*, 973–978. doi: 10.1016/j.paid.2009.07.026

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin, 135,* 943–973. doi: 10.1037/a0017327

Ring, P., Neyse, L., David-Barett, T., & Schmidt, U. (2016). Gender differences in performance predictions: Evidence from the Cognitive Reflection Test. *Frontiers in Psychology,* 7:1680. doi: 10.3389/fpsyg.2016.01680

Schiff, S., Amodio, P., Testa, G., Nardi, M., Montagnese, S., Caregaro, L., di Pellegrino, G., & Sellitto, M. (2016). Impulsivity toward food reward is related to BMI: Evidence from intertemporal choice in obese and normal-weight individuals. *Brain and Cognition*, *110*, 112–119. doi: 10.1016/j.bandc.2015.10.001

Scholten, M., Read, D., & Sanborn, A. (2014). Weighing outcomes by time or against time? Evaluation rules in intertemporal choice. *Cognitive Science, 38,* 399–438. doi: 10.1111/cogs.12104

Sinayev, A., & Peters, E. (2015). Cognitive reection vs. calculation in decision making. *Frontiers in Psychology, 6*, 532. doi: 10.1111/cogs.12104

Sirota, M., & Juanchich, M. (2018). Effect of response format on cognitive reflection: Validating a two- and four-option multiple choice question version of the cognitive reflection test. *Behavior Research Methods, 50*(6), 2511–2522. doi: 10.3758/s13428-018-1029-4

Sirota, M., Kostovičová, L., Juanchich, M., Dewberry, C., & Marshall, A. C. (2018). Measuring cognitive reflection without maths: Developing and validating the verbal cognitive reflection test. *PsyArXiv Preprints*. doi: 10.31234/osf.io/pfe79

Sladek, R. M., Bond, M. J., & Phillips, P. A. (2010). Age and gender differences in preferences for rational and experiential thinking. *Personality and Individual Differences, 49*(8), 907–911. doi: 10.1016/j.paid.2010.07.028

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*(1), 3–22. doi: 10.1037/0033-2909.119.1.3

Šrol, J. (2018). These problems sound familiar to me: Previous exposure, Cognitive Reflection Test, and the moderating role of analytic thinking. *Studia Psychologica, 60*(3), 195–208. doi: 10.21909/sp.2018.03.762

Stagnaro, M. N., Pennycook, G., & Rand, D. G. (2018). Performance on the Cognitive Reflection Test is stable across time. *Judgment and Decision Making, 13*(3), 260–267. http://journal.sjdm.org/18/18201/jdm18201.html

Stillman, P. E., Medvedev, D., & Ferguson, M. J. (2017). Resisting temptation: Tracking how self-control conflicts are successfully resolved in real time. *Psychological Science*, *28*(9), 1240–1258. doi: 10.1177/0956797617705386

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the Cognitive Reflection Test. *Judgment and Decision Making, 11*(1), 99–113. http://journal.sjdm.org/15/151029/jdm151029.html

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning, 20*, 147–168. doi: 10.1080/13546783.2013.844729

Yilmaz, O., & Saribay, S. A. (2017). The relationship between cognitive style and political orientation depends on the measures used. *Judgment and Decision Making, 12*(2), 140–147. http://journal.sjdm.org/16/161128a/jdm161128a.html

Zhang, D. C., Highhouse, S., & Rada, T. B. (2016). Explaining sex differences on the Cognitive Reflection Test. *Personality and Individual Differences, 101*, 425–427. doi: 10.1016/j.paid.2016.06.034

Studia Psychologica, Vol. 61, No. 2, 2019, 99-109 doi: 10.21909/sp.2019.02.775

Mental Simulation as a Remedy for Biased Reasoning

Daša Strachanová, Lenka Valuš
Institute of Experimental Psychology, Center of Social and Psychological Sciences
Slovak Academy of Sciences

Prompting mental simulation with a counterfactual scenario has been found to enhance rationality in individuals and groups. Building upon previous findings and the dual-process accounts of reasoning, we hypothesized that debiasing power of mental simulation lies in inhibiting System 1 and facilitating System 2 responses. Therefore, we examined whether counterfactual priming mitigates biased reasoning via changes in cognitive reflection. Each participant of our between-subject experiment (N = 462) solved two out of three tasks on biased reasoning: one before and one after being exposed to the counterfactual scenario. The tasks were designed to elicit selectively seeking hypothesis-confirming evidence, ignoring alternative explanations, and unwillingness to reconsider the default option. In addition, the participants completed two sets of cognitive reflection problems at the beginning and at the end of the experiment. Mental simulation reduced people's tendencies to ignore alternative explanations and hypothesis-disconfirming evidence, and the latter effect was mediated by intuition inhibition.

Key words: debiasing; counterfactual priming; mental simulation; cognitive reflection

Introduction

Decades of research have focused on documenting all the deviations from optimal judgments and choices people fall prey to. Lists of these cognitive biases seem endless (see Benson, 2016, for a nice visualization), and we moved from the concept of so-called "homo economicus" to a much more realistic image of human reasoning. As a result, finding ways of improving decision-making, especially when high stakes are at play, is being widely discussed. This, however, proved to be a much bigger challenge than we might have expected. Early debiasing attempts based on warnings about the possibility of bias occurrence, explaining nature of the bias, and providing a dose of feedback failed to succeed (Fischhoff, 1982). How far have we got, a quarter century later?

In the current debiasing efforts, two approaches have been applied: we either try to change the environment or the decision-maker. Typical examples of the first category are nudge techniques (Thaler & Sunstein, 2008), helpful external representations (e.g., Sirota, Kostovičová, & Juanchich, 2014) or inducing the sense of accountability (Lerner & Tetlock, 1999). Soll, Milkman, and Payne (2016) list various other possibilities for modifying the environment in favor of better decisions. The other way, based on changing people, is much more demanding. We can cultivate rationality with trainings focused on a particular single competence, such as statistical reasoning (Sirota, Kostovičová, & Vallée-

Acknowledgment
This study was funded by the Slovak Research and Development Agency and is part of research project APVV-16-0153 "Cognitive failures – individual predictors and intervention possibilities".

Correspondence concerning this paper should be addressed to Lenka Valuš, Institute of Experimental Psychology, Center of Social and Psychological Sciences, Slovak Academy of Sciences, Dúbravská cesta 9, Bratislava, 84104, Slovak Republic. E-mail: valus.lenka@gmail.com

Received November 13, 2018

Tourangeau, 2015) and considering alternatives (Hirt & Markman, 1995), or we can use a complex intervention. Morewedge et al. (2015), for instance, reduced crucial cognitive biases via teaching a whole set of competences, such as hypothesis testing, considering countervailing evidence, alternative explanations and perspectives, base rates and anchors, and encouraging people to reconsider their initial answers. There are, however, some indirect methods as well, mindfulness meditation being one of them (Hafenbrack, Kinias, & Barsade, 2014). Another indirect debiasing intervention is based on mental simulation.

During the process of mental simulation people think of various relevant but potentially converse alternatives (Galinsky & Moskowitz, 2000). In the naturalistic decision-making paradigm, mental simulation is a conscious, deliberate, and analytic strategy used to evaluate different courses of action (Klein, 2008). It can be trained or primed. In a single-shot intervention, mental simulation can be induced with a counterfactual priming (Galinsky & Moskowitz, 2000). The method is based on a scenario, which encourages people to produce counterfactuals – i.e. thoughts of what might have been, e.g., "If only I had tried more …".

They represent alternative realities for a past event (e.g., "... I would have been better off."). Taken together, counterfactual scenarios are supposed to activate a mental simulation mindset in which various alternatives are considered (Galinsky & Moskowitz, 2000). As a result, mental simulation might affect cognitive performance in situations when taking into account different views and explanations is crucial. Indeed, it has been found to mitigate confirmation bias in both individuals (Galinsky & Moskowitz, 2000) and groups (Kray & Galinsky, 2003), to enhance performance in the hidden profile task (Liljenquist, Galinsky, & Kray, 2004), and to reduce prejudice (Miller, Markman, Wagner, & Hunt, 2013).

Here we decided to test the effect of counterfactual priming on the three components of biased reasoning: selectively seeking hypothesis-confirming evidence, neglecting alternative explanations and unwillingness to reconsider the default option in response to new information. *Biased search for information* in line with prior hypothesis or belief is a key component of the confirmation bias, one of the most robust cognitive deviations (Nickerson, 1998). *Neglecting alternative explanations* is closely tied to many cognitive biases. One of many examples is the fundamental attribution error, manifested as overestimation of internal factors and underestimation of external factors when evaluating other people's behavior (Ross, 1977). The last one, *sticking with the default*, is at the heart of status-quo bias – the tendency to prefer the current state of affairs despite the existence of better alternatives (Samuelson & Zeckhauser, 1988). These three tendencies are in sharp contrast to open-minded thinking (Baron, 1993) and, therefore, represent barriers to rational judgment and decision-making (e.g., Svedholm-Häkkinen & Lindeman, 2018).

Besides examining the effect of mental simulation on biased reasoning, we were also interested in its explanation, which has yet to be empirically tested. In the context of the dual-processes account of human reasoning (e.g., Wason & Evans, 1975), mental simulation is believed to correspond to the System 2 processes (Klein, 2008). In line with this assumption, counterfactual thinking enhanced performance on analytic problems (Kray, Galinsky, & Wong, 2006; Markman, Lindberg, Kray, & Galinsky, 2007). We, therefore, hypothesized that the effect of mental simulation on tasks which require both overriding automatic responses and considering alternatives is linked to activation and inhibition of the two Systems. Specifically, we expected that mental simulation induced with a counterfactual scenario would support the analytic System 2 processes and

inhibit the intuitive System 1 processes, and this would result in less biased reasoning. As a proxy for intuitive versus analytic thinking we chose cognitive reflection. A high score on the cognitive reflection tests (Frederick, 2005; Sirota, Kostovičová, Juanchich, Dewberry, & Marshall, 2018; Toplak, West, & Stanovich, 2014) means that the person is able to exert cognitive effort, suppress their automatic incorrect responses, and engage in deliberative thinking. It is a very strong predictor of susceptibility to cognitive biases (e.g., Toplak, West, & Stanovich, 2011, 2014), rational choices and real-life decision outcomes (e.g., Campitelli & Labollita, 2010; Juanchich, Dewberry, Sirota, & Narendran, 2016).

Taken together, we aimed at examining whether counterfactual priming would mitigate biased reasoning via changes in cognitive reflection. We anticipated that thanks to mental simulation, activated with a counterfactual scenario, people would inhibit their intuition, engage in analytic thinking, and their responses would be less biased as a result. Specifically, they would seek both hypothesis-confirming and disconfirming evidence, they would consider alternative explanations, and would be willing to reconsider the default option when exposed to novel information.

Method

Participants

We assumed a counterfactual priming effect similar to the one in Galinsky and Moskowitz's (2000) study, $d = 0.55$. Based on a power calculation (two-sided tests, $\alpha = .05$, $1 - \beta = .95$), we planned to recruit at least 314 participants in total (adjusted for the assumption that one fifth would fail to pass an attention check item). Four hundred eighty-six people in a convenience sample filled out our online questionnaire. After excluding those who failed to pass the attention check item, our final sample consisted of 462 participants (60% female, 18 to 59 years, $M = 24.3$, $SD = 6.8$). Two thirds were university students. Concerning the study programs, the participants were students or graduates in natural and technical sciences (48%), social sciences and humanities (40%) or other fields (art, sport, etc.; 12%).

Design and Procedure

The participants first answered basic socio-demographic items. Afterwards, they solved a first series of cognitive reflection items (CRT1). Then, they were randomly redirected to one of six groups, as depicted in Figure 1. These groups represent three control groups (CG): CG for Task 1, CG for Task 2, and CG for Task 3. After solving one of the three tasks, all the participants proceeded to our experimental manipulation – the counterfactual priming. Next, we asked them to solve one of the remaining two tasks. For instance, roughly half of those who solved Task 3 prior to the intervention, were administered Task 1 after the priming. The other half solved Task 2. Taken together, there were six groups: Task 1 - Priming - Task 2, Task 1 - Priming - Task 3, Task 2 - Priming - Task 1, etc. As a result, there was a different composition of the control and the experimental group for each of the three tasks. Finally, the participants answered a second series of cognitive reflection problems (CRT2). The two CRT tests were administered in a randomized order. Priming exposure acted as an independent variable, performance in the three tasks as dependent variables, and correct and intuitive CRT scores (control group: CRT1 / experimental group: CRT2) as mediators.

Materials and Measures

We conducted two pilot studies in the first phase of the research to empirically evaluate

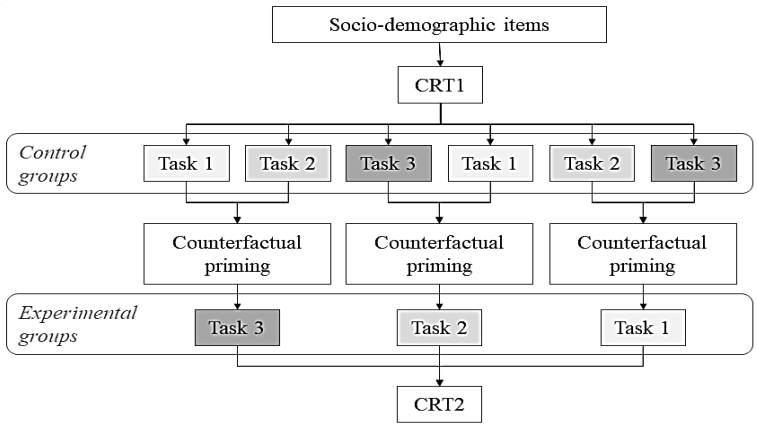


Figure 1 The design and procedure used in the experiment

stimuli and material. First, we adopted a *counterfactual scenario* from Galinsky and Moskowitz (2000) and pretested it with a sample of 39 participants. The story concerned Jane, a woman at a concert of her favorite band. She switched her seat to have a better view but, as a result, she did not win a trip to Hawaii since the winner was sitting exactly at the place she had left. The other, neutral, scenario did not contain this outcome – the winning seat had nothing to do with Jane's choice. The participants were asked to generate thoughts that might run through Jane's head after the concert. Answers of the two groups (neutral vs. counterfactual scenario) were coded by two independent raters. Given a high reliability, $r = .82$, $p < .001$, the pairs of ratings were averaged. The neutral scenario produced significantly less counterfactual thoughts than the counterfactual scenario ($M = 0.3$, $SD = 0.4$ vs. $M = 0.8$, $SD = 0.7$, respectively), $t(37) = -2.73$, $p = .010$, $d = 0.88$. With respect to the results and in order to increase the effectiveness of the manipulation, we added a sentence at the end of the scenario: "On her way home, Jane is thinking of what she would have gained and what she would have lost if she had not switched her seat".

The objective of the other pilot study ($N = 26$) was to prepare two short *cognitive reflection tests* (CRT) of roughly the same difficulty. We tested the original CRT tasks (Frederick, 2005), their alternatives with a modified wording (Kostovičová, Dudeková, & Konečný, 2013), new items from the extended version (Toplak et al., 2014), and verbal CRT problems (Sirota et al., 2018). We checked the items for their difficulty, discrimination power, content clarity and familiarity. Based on the results, we designed two versions, consisting of 2 verbal and 2 numerical tasks, for instance: "How many of each animal did Moses take on the ark?" (Sirota et al., 2018) or "Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class?" (Toplak et al., 2014, p. 151). Performance on the two tests was almost identical (CRTa: $M = 1.6$,

$SD = 1.0$; CRTb: $M = 1.6$, $SD = 0.9$), $t(25) < 0.01$, $p = 1.000$, $d < 0.01$. We coded the answers as correct, intuitive or other. The scores of correct as well as intuitive responses ranged from 0 to 4.

The indicators of *biased reasoning* focused on selectively seeking hypothesis-confirming evidence (Task 1), neglecting alternative explanations (Task 2), and unwillingness to reconsider the default option in response to new information (Task 3). The first task was inspired by Galinsky and Moskowitz (2000). The story was about Tereza who attends speed-dating and before the event starts, she spots a man whose behavior suggests he is an introvert. Since she knows she will have only one minute to talk to him, she wants to prepare some questions to test whether her impression is correct. The participants were asked to choose 5 out of 10 questions. Half of the questions captured typical behavior of an introvert. These five "introverted" questions were supposed to elicit hypothesis-confirming answers (e.g., "Do you like spending evenings at home?"). The content of the other five questions concerned typical behavior of extroverts. These five "extroverted" questions were meant to elicit hypothesis-disconfirming answers (e.g., "Do you like meeting new people?"). The more introverted questions and the fewer extroverted questions a participant chooses, the more s/he seeks information that would support the initial hypothesis of the man being an introvert. Thus, the difference between the number of introverted and extroverted questions (range: -5 to 5; -5: extroverted questions only / +5: introverted questions only) captures the level, to which a person *selectively seeks hypothesis-confirming evidence*.

In the second task, the participants were asked to imagine that they have a new colleague, Adam. He seems hardworking, diligent, and punctual, yet his mood and behavior are strange and unpredictable. After the scenario, which provided both reasons for situation- and personality-based explanations of Adam's actions, the participants answered three questions on 7-point Likert scales. The first two items were only distractors and concerned their overall impression and their view on the colleague's potential. The third, crucial one focused on the explanation of the colleague's unstable behavior (-3: situational factors only; 3: colleague's personality only). We were interested in people's preference for one or the other category (situation vs. personality), but our main focus was on the absolute score (range 0 to 3). Thus, given the doubts on the robustness of the fundamental attribution error (e.g., Malle, 2006), and since human behavior is generally influenced by both internal and external factors, and the scenario did not encourage preference for any of them, the participants should have considered both explanations. Thus, the higher the score the more a person *neglects an alternative explanation*.

The last task was about vaccination against HPV (human-papillomavirus, the main cause of cervical cancer). The participants were informed that this vaccine is not mandatory in Slovakia. After being provided with arguments on possible benefits of the vaccination, and on the risks of side-effects as well (Kostovičová, Bašnáková, & Bačová, 2017), they answered two questions. The first one concerned whether vaccination is one of the greatest discoveries of medicine or rather one of its most controversial topics. In the other – critical one, the participants expressed whether they would be *willing to reconsider the default option* (non-mandatory vaccination) *in response to new information* provided (no/yes).

Results

First of all, scores of the correct CRT answers in the two measurements were moderately correlated, $r = .41$, $p < .001$, and so were the scores

of the intuitive answers, $r = .36$, $p < .001$. The numbers of correct CRT answers, $r = -.19$, $p = .001$, and intuitive CRT responses, $r = .16$, $p = .006$, in the second measurement correlated with selectively seeking hypothesis-confirming evidence. We found no significant correlation between the three indicators of biased reasoning.

There was no substantial effect of experimental manipulation on correct CRT responses, regardless of the specific order. Overall, the CRT scores before and after intervention were very similar ($M = 2.0$, $SD = 1.2$ vs. $M = 1.9$, $SD = 1.2$), $t(461) = 1.54$, $p = .324$, $d = 0.10$. However, counterfactual priming *reduced intuitive CRT responses* ($M = 1.3$, $SD = 1.2$ vs. $M = 1.0$, $SD = 1.0$), $t(461) = 6.04$, $p < .001$, $d = 0.40$. It also enhanced the number of other, incorrect but non-intuitive, answers ($M = 0.7$, $SD = 0.9$ vs. $M = 1.2$, $SD = 1.2$), $t(461) = -7.24$, $p < .001$, $d = 0.48$. These patterns were the same when we separately analyzed changes in performance in verbal and numerical CRT problems. Taken together, mental simulation prompted with a counterfactual scenario helped override automatic intuitive responses but failed to sufficiently help the participants to come up with correct solutions.

At the same time, responses in the third task remained unaffected. Specifically, two thirds of the control (66%, $n = 107$) and the experimental group (68%, $n = 101$) expressed that vaccination is one of the greatest discoveries of medicine rather than one of its most controversial topics, $\chi^2(1) = 0.11$, $p = .745$, $\varphi = .02$. And, more importantly, willingness to reconsider the default option was also almost identical in the control (44%, $n = 90$) and the experimental group (43%, $n = 85$), $\chi^2(1) = 0.07$, $p = .791$, $\varphi = .02$. Correlation between the two responses was significant, yet weak, $r_S = .22$, $p < .001$.

As for the assessment of the colleague's behavior, responses of the experimental group were more in line with the situational explanation (MRank = 146.0) compared to the control group (MRank = 159.3), yet the difference was non-significant, $M\text{-}WU = 10538.5$, $p = .175$, $r_m = 0.08$. When comparing the level to which one neglects an alternative explanation, we found that the score was significantly higher in the control group ($M = 1.8$, $SD = 1.0$ vs. $M = 1.4$, $SD = 1.0$), $t(302) = 3.55$, $p < .001$, $d = 0.41$. Thus, counterfactual priming *enhanced considering alternative explanations*.

Comparisons of the numbers of introverted and extroverted questions (range 0 to 5), and the overall score (range -5 to 5) for the "speed-dating" task are presented in Figure 2. The number of hypothesis-confirming (introverted) questions was slightly higher in the control group, but without a significant difference, $t(307) = 1.76$, $p = .080$, $d = 0.20$. The number of hypothesis-disconfirming (extroverted) ques-

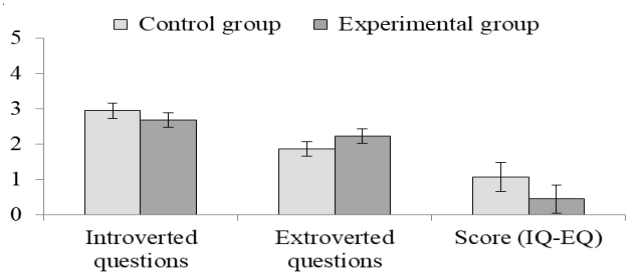


Figure 2 Comparison of hypothesis dis/confirming choices

tions was significantly higher in the experimental group, $t(307) = -2.43$, $p = .016$, $d = 0.28$. And finally, the overall score was significantly higher in the control group, $t(307) = 2.17$, $p = .031$, $d = 0.25$, which means that *counterfactual priming reduced selectively seeking hypothesis-confirming evidence*.

Since the counterfactual priming did not affect correct CRT responses and willingness to reconsider the default option, we performed two mediation analyses (Hayes, 2013) with priming as an independent variable and intuitive CRT responses as a mediator. In the first tested model with considering alternative explanations as a dependent variable, only the effects of intervention were significant. The other tested model with seeking hypothesis-confirming evidence is depicted in Figure 3. We found that *the number of intuitive CRT answers mediated the effect of counterfactual priming on selectively seeking hypothesis-confirming evidence*. Yet, the indirect effect (-0.083) and the *R*-squared mediation effect size (.004) were quite small. Order of the two CRT versions was entered as a covariate.

Discussion

Cognitive errors are costly and are getting costlier; therefore, the main challenge for the current judgment and decision-making research is development of improvement strategies (Milkman, Chugh, & Bazerman, 2009). Debiasing efforts can be focused either on modification of the environment (such as helpful external representations) or the decision makers themselves (such as cultivating rationality via training). Both approaches have their pros and cons but – except for some nudge techniques – their implementation in the real-world settings poses many pragmatic difficulties and requires a lot of time and effort. Previous research showed that mental simulation prompted with a counterfactual scenario represents a "fast and frugal" way of reducing biases for which considering alternatives is crucial (Galinsky & Moskowitz, 2000; Kray & Galinsky, 2003; Liljenquist et al., 2004; Miller et al., 2013). Given the link between higher performance in analytic problems and counterfactual thinking (Kray et al., 2006; Markman et al., 2007), we hypothesized that the effect of mental simulation on biased reasoning would be mediated by changes in cognitive reflection. In other words, we anticipated that exposure to counterfactual scenario would result in lower levels of seeking hypothesis-confirming evidence, ignoring alternative explanations, and unwillingness to reconsider the default option thanks to inhibiting intuitive System 1 processes and activating deliberative System 2 thinking.

We found that counterfactual priming failed to enhance the cognitive reflection scores, yet

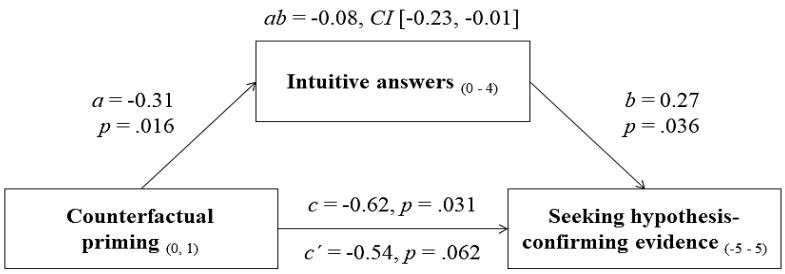


Figure 3 Intuition inhibition as a mediator of the effect of counterfactual priming on selectively seeking hypothesis-confirming evidence

it reduced the level of intuitive responses. This means that mental simulation mindset helped in suppressing appealing automatic answers but was not efficient in generating the correct ones instead. As Stanovich (2018) points out, some cognitive errors cannot be unambiguously credited to miserly processing. Thus, sometimes people miss some important mindware component, such as knowledge of scientific and probabilistic reasoning, or math skills. Given the strong association of the CRT with numeracy (e.g., Liberali et al., 2012), cognitive reflection performance might represent a quite stable individual disposition, which is rather difficult to change (e.g., Stagnaro, Pennycook, & Rand, 2018). Yet, a positive effect of mental simulation on intuition inhibition is a promising finding, which deserves further examination.

The other case, which contradicted our assumptions, is finding no effect of the experimental manipulation on the willingness to reconsider the default option in response to new information. We suppose the explanation is twofold. First, we chose a quite controversial and emotionally charged topic, which polarizes society into two opposite groups, one of which is significantly smaller in number (Masaryk & Hatoková, 2016; Tvardzík, 2015). Thus, people's preferences might be so strong that they are resistant to such a simple debiasing method. This is in line with our previous findings on domain specificity of the effects of counterfactual thinking (Strachanová & Grežo, 2018). In addition, maybe the participants refused to reconsider the default option not because they are against vaccination as such but because we used the term "mandatory". Next time we would replace the forced binary choice with a Likert scale as well. We also plan to further investigate the moderating role of the domain in debiasing via mental simulation.

On the other hand, counterfactual priming substantially facilitated search for hypothesis-disconfirming evidence and consideration of alternative explanations. These patterns suggest that mental simulation might help in overcoming cognitive failures such as belief bias, confirmation bias, fundamental attribution error, myside bias, self-serving bias, survivorship bias, and many more. It is even more promising when we take into account that counterfactual thinking can be systematically trained (Hendrickson, 2008), so it becomes a natural component of one's cognitive strategy repertoire. Mental simulation toward future, for instance, is a typical supporting strategy in naturalistic decision-making (Klein, 2008; Klein et al., 2003). Counterfactual thinking might therefore become a part of training for managers, teachers and other professionals, and can be utilized in educational campaigns. In order to determine whether it is appropriate to use counterfactuals in direct or indirect debiasing methods, we need to explore the nature of the bias (Kahn, Luce, & Nowlis, 2006). Most importantly, we have to know whether the decision maker is aware of the underlying processes leading to the bias. Since we believe that the three components of biased reasoning we have focused on in our research are largely subconscious, indirect debiasing method is a proper strategy for their reduction. Although the effects of counterfactual priming are short-term only, indirect debiasing strategies eliminate some of the barriers associated with direct interventions, such as bias blind spot or perceived relevance (Lilienfeld, Ammirati, & Landfield, 2009).

The current study has several limitations, out of which two were already mentioned: CRT as a problematic measure of analytic (or intuitive) reasoning, and the task with mandatory vaccination and a binary question as a problematic measure of willingness to reconsider the default option. Instead of a Likert scale, an open question for participants to state their own reasons for a colleague's behavior might produce more ecologically valid answers in the task on

neglecting alternative explanation. In the last task on selectively seeking hypothesis-confirming evidence, a few neutral questions (neither introverted nor extroverted) might help as distractors. Overall, multiple tasks to measure the same aspect of biased reasoning and composite scores would ensure results of a higher reliability. We also admit that our sample was not representative, especially in terms of age and education. And, most importantly, we cannot consider our search for explanation of debiasing power of mental simulation successful. Although intuition inhibition mediated the effect of counterfactual priming on seeking hypothesis-confirming evidence, the indirect effect was rather small and the finding is inconclusive.

Thus, future research should address the missing parts of the "counterfactual priming" puzzle. First, the mediators of the effect: either analytic thinking measured in a different way, or some other constructs that are linked to biased reasoning and cognitive performance, such as open-mindedness (e.g., Svedholm-Häkkinen & Lindeman, 2018) or metacognitive awareness (Schraw, 1998). Future research should focus on moderating variables as well. These might be some personal characteristics which make people less or more resistant to the intervention, and task features as well. According to the recent neuroscience findings (De Brigard, Spreng, Mitchell, & Schacter, 2015), for instance, effectiveness of the counterfactual priming might depend on whether the protagonist in the scenario is an unknown person or the participant. Another moderator worthy of investigation is mental simulation toward past versus future. Follow-up studies could also focus on particular cognitive biases, which have not been tested under the mental simulation condition. Last but not least, the effect of mental simulation as both indirect and direct debiasing strategy should be tested in the real-world settings.

References

Baron, J. (1993). Why teach thinking? An essay. *Applied Psychology: An International Review*, *42*(3), 191–214. doi: 10.1111/j.1464-0597.1993.tb00731.x

Benson, B. (2016). *Cognitive bias cheat sheet*. Retrieved from https://betterhumans.coach.me/cognitive-bias-cheat-sheet-55a472476b18

Campitelli, G., & Labollita, M. (2010). Correlations of cognitive reflection with judgments and choices. *Judgment and Decision Making*, *5*(3), 182–191.

De Brigard, F., Spreng, N. R., Mitchell, J., & Schacter, D. (2015). Neural activity associated with self, other, and object-based counterfactual thinking. *NeuroImage*, *109*, 12–26. doi: 10.1016/j.neuroimage.2014.12.075

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422–444). Cambridge: Cambridge University Press.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*(4), 25–42. doi: 10.1257/089533005775196732

Hafenbrack, A. C., Kinias, Z., & Barsade, S. G. (2014). Debiasing the mind through meditation: Mindfulness and the sunk-cost bias. *Psychological Science*, *25*(2) 369–376. doi: 10.1177/0956797613503853

Hayes, A. F. (2013). *An introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.

Hirt, E., & Markman, K. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology*, *69*(6), 1069–1086. doi: 10.1037/0022-3514.69.6.1069

Galinsky, A. D., & Moskowitz, G. B. (2000). Counterfactuals as behavioral primes: Priming the simulation heuristic and consideration of alternatives. *Journal of Experimental Social Psychology*, *36*(4), 384–409. doi: 10.1006/jesp.1999.1409

Hendrickson, N. (2008). *Counterfactual reasoning. A basic guide for analysts, strategists, and decision makers.* Washington, D.C.: The National Intelligence University.

Juanchich, M., Dewberry, C., Sirota, M., & Narendran, S. (2015). Cognitive reflection predicts reallife decision outcomes, but not over and above personality and decision making styles. *Journal of Behavioral Decision Making*, *29*(1), 52–59. doi: 10.1002/bdm.1875

Kahn, B. E., Luce, M. F., & Nowlis, S. M. (2006). Debiasing insights from process tests. *Journal of*

Consumer Research, *33*(1), 131–138. doi:10.1086/500492

Klein, G. (2008). Naturalistic decision making. *Human Factors*, *50*(3), 456–460. doi: 10.1518/001872008X288385

Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., & Hollnagel, E. (2003). Macrocognition. *IEEE Intelligent Systems*, *18*(3), 81–85. doi: 10.1109/MIS.2003.1200735

Kostovičová, L., Bašnáková, J., & Bačová, V. (2017). Predicting perception of risks and benefits within novel domains. *Studia Psychologica*, *59*(3), 176–192.

Kostovičová, L., Konečný, M., & Dudeková, K. (2013). Čo to vlastne meriame? Reflexia Testu kognitívnej reflexie [What do we actually measure? Reflection of the Cognitive Reflection Test]. In I. Sarmány-Schuller (Ed.), *Otázniky rozhodovania: teória, empíria, život, 31. Psychologické dni: zborník príspevkov* [31st Psychological days: Conference proceedings] (pp. 166–171). Bratislava: STIMUL.

Kray, L. J., & Galinsky, A. D. (2003). The debiasing effect of counterfactual mind-sets: Increasing the search for disconfirmatory information in group decisions. *Organizational Behavior and Human Decision Processes*, *91*(1), 69–81. doi: 10.1016/S0749-5978(02)00534-4

Kray, L. J., Galinsky, A. D., & Wong, E. (2006). Thinking within the box: The relational processing style elicited by counterfactual mind-sets. *Journal of Personality and Social Psychology*, *91*(1), 33–48. doi: 10.1037/0022-3514.91.1.33

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, *125*(2), 255–275. doi:10.1037/0033-2909.125.2.255

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making*, *25*(4), 361–381. doi: 10.1002/bdm.752

Lilienfeld, S. O., Ammirati, R., & Landfield, K. (2009). Giving debiasing away: Can psychological research on correcting cognitive errors promote human welfare? *Perspectives on Psychological Science*, *4*(4), 390–398. doi: 10.1111/j.1745-6924.2009.01144.x

Liljenquist, K. A., Galinsky, A. D., & Kray, L. J. (2004). Exploring the rabbit hole of possibilities by myself or with my group: The benefits and liabilities of activating counterfactual mindsets for information sharing and group coordination. *Journal of Behavioral Decision Making*, *17*(4), 263–279. doi: 10.1002/bdm.479

Malle, B. F. (2006). The actor-observer asymmetry in attribution: A (surprising) meta-analysis. *Psychological Bulletin*, *132*(6), 895–919. doi:10.1037/0033-2909.132.6.895

Markman, K. D., Lindberg, M. J., Kray, L. J., & Galinsky, A. D. (2007). Implications of counterfactual structure for creative generation and analytical problem solving. *Personality and Social Psychology Bulletin*, *33*(3), 312–324. doi: 10.1177/0146167206296106

Masaryk, R., & Hatoková, M. (2016). Qualitative inquiry into reasons why vaccination messages fail. *Journal of Health Psychology*, *22*(14), 1880–1888. doi: 10.1177/1359105316656770

Milkman, K. L., Chugh, D., & Bazerman, M. H. (2009). How can decision making be improved? *Perspectives on Psychological Science*, *4*(4), 379–383. doi: 10.1111/j.1745-6924.2009.01142.x

Miller, A. K., Markman, K. D., Wagner, M. M., & Hunt, A. N. (2013). Mental simulation and sexual prejudice reduction: The debiasing role of counterfactual thinking. *Journal of Applied Social Psychology, 43*(1), 190–194. doi: 10.1111/j.1559-1816.2012.00992.x

Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing decisions: Improved decision making with a single training intervention. *Policy Insights from Behavioral and Brain Sciences*, *2*(1), 129–140. doi: 0.1177/2372732215600886

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220. doi: 10.1037/1089-2680.2.2.175

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology*, *10*, 173–220. doi: 10.1016/S0065-2601(08)60357-3

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, *1*(1), 7–59. doi: 10.1007/BF00055564

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, *26*(1), 113–125. doi: 10.1023/A:1003044231033

Sirota, M., Kostovičová, L., & Juanchich, M. (2014). The effect of iconicity of visual displays on statistical reasoning: evidence in favor of the null hypothesis. *Psychonomic Bulletin & Review*, *21*(4), 961–968. doi: 10.3758/s13423-013-0555-4

Sirota, M., Kostovičová, L., Juanchich, M., Dewberry, C., & Marshall, A. C. (2018). *Measuring cognitive reflection without math: Developing and validating verbal cognitive reflection test.* Manuscript submitted for publication.

Sirota, M., Kostovičová, L., & Vallée-Tourangeau, F. (2015). How to train your Bayesian: A problem-representation transfer rather than a format-representation shift explains training effects. *The Quarterly Journal of Experimental Psychology*, *68*(1), 1–9. doi: 10.1080/17470218.2014.972420

Soll, J. B., Milkman, K. L., & Payne, J. W. (2016). A user's guide to debiasing. In G. Keren & G. Wu (Eds.), *The Wiley Blackwell handbook of judgment and decision making II* (pp. 924–951). Malden, MA: Wiley-Blackwell.

Stanovich, K. E. (2018). Miserliness in human cognition: The interaction of detection, override and mindware. *Thinking & Reasoning*, *24*(4), 423–444. doi: 10.1080/13546783.2018.1459314

Stagnaro, M. T., Pennycook, G., & Rand, D. G. (2018). Performance on the Cognitive Reflection Test is stable across time. *Judgment and Decision Making*, *13*(3). doi: /10.2139/ssrn.3115809

Strachanová, D., & Grežo, M. (2018). Keby som sa rozhodol inak: Kontrafaktové myslenie a efekt utopených nákladov v rozhodovaní [If I had decided differently: Counterfactual thinking and the sunk-cost effect in decision-making]. In E. Maierová, L. Viktorová, J. Suchá, & M. Dolejš (Eds.), *PhD existence 2018 - Nekonečno v psychologii: Sborník příspěvků* [PhD existence 2018 - Infinity in psychology: Conference proceedings] (pp. 97–103). Olomouc: Univerzita Palackého.

Svedholm-Häkkinen, A. M., & Lindeman, M. (2018). Actively open-minded thinking: Development of a shortened scale and disentangling attitudes towards knowledge and people. *Thinking & Reasoning*, *24*(1), 21–40. doi: 10.1080/13546783.2017.1378723

Thaler, R. H., & Sunstein, C (2008). *Nudge*. Yale University Press.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics and biases tasks. *Memory & Cognition*, *39*(7), 1275–1289. doi: 10.3758/s13421-011-0104-1

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, *20*(2), 147–168. doi: 10.1080/13546783.2013.844729

Tvardzík, J. (2015, May 20). Očkovanie: Kolektívna imunita je v ohrození. Epidémia je reálna [Vaccination: Collective immunity is at risk. The epidemic is real]. *Trend*. Retrieved from https://www.etrend.sk/ekonomika/ockovanie-kolektivna-imunita-je-v-ohrozeni-epidemia-je-realna-varuje-expertka.html?split=1

Wason, P. C., & Evans, J. S. B. T. (1975). Dual processes in reasoning? *Cognition*, *3*(2), 141–154. doi: 10.1016/0010-0277(74)90017-1

Studia Psychologica, Vol. 61, No. 2, 2019, 110-119 doi: 10.21909/sp.2019.02.776

Interpersonal Style of Coaching, Motivational Profiles and the Intention to be Physically Active in Young Athletes

Alfonso Valero-Valenzuela, David Manzano-Sánchez
Department of Physical Activity and Sport, Faculty of Sports Science, University of Murcia, Spain

Juan Antonio Moreno-Murcia
Department of Health Psychology, Faculty of Sociosanitary Sciences, Miguel Hernández University, Elche, Spain

Diego Andrés Heredia León
Department of Educational Sciences, Faculty of Teacher Training and Education, Universidad Católica de Cuenca, Ecuador

The aim of this study was to assess the relationships among the motivational profile, the coach's interpersonal style and the intention to be physically active in young athletes. A sample of 254 athletes (mean age of 12.81 years), who used to participate in official competitions, was used. The measurements taken were of the young athletes' perception of the interpersonal style of the coach, satisfaction of basic psychological needs, motivation towards sports practice and intention to be physically active. Bivariate correlation, cluster and multivariate analyses were carried out. The cluster analysis revealed two profiles: one with high self-determined motivation and greater values of autonomy support, competence, relatedness and intention to be physically active, and another less self-determined cluster that showed low levels of competence, relatedness and intention to be physically active. It is suggested that adding some strategies based on autonomy support to training with athletes may ensure adherence to sports practice.

Key words: self-determination, autonomy support, motivation, athletics

Introduction

Among the aspects that most influence the adherence to sports is the teaching methodology used by teachers and coaches (Martín-Albo, Nuñez, & Navarro, 2003). Unfortunately, there are still many cases where methodology is applied with an excessively rigid lesson structure and motivation is encouraged that is too focused on external aspects (Valero, 2004). This creates athletes with low motivation, which can interfere with the adherence to sports practice and the intention to be physically active (Almagro & Paramio-Pérez, 2017; Almagro, Sáenz, & Moreno, 2010; Liao, Chou, Huh, Levenhal, & Dunton, 2017; Ulrich-French & Smith, 2009). This is the reason why it is necessary to study the factors which can influence such motivation in sports.

In this regard, the "Self-Determination Theory" (SDT) (Deci & Ryan, 1985, 2000) appears as a macro-theory of human motivation related to personality development and functioning in social contexts. It focuses on the degree to which human behavior is self-determined

Correspondence concerning this article should be addressed to Dr. David Manzano-Sánchez, Department of Physical Activity and Sport, Faculty of Sports Science, University of Murcia, Calle Argentina, 19, 30720 San Javier, Murcia, Spain. E-mail: David.manzano@um.es

Received November 17, 2018

or volitional. The SDT not only concerns the specific nature of positive development trends, but also examines the social environments that support these trends.

In addition, the origin of motivation and its consequences at a cognitive, behavioral and affective level on the individual are analyzed (Vallerand, 1997). Within the sport context, one of the social factors that acquire a fundamental role in athletes' motivation is the interpersonal style that the coach uses when giving instructions (Haerens et al., 2017). This may range from frequent extrinsic incentives (controller style) to a prominent role of the athlete (autonomy support), participating in decision making and acquiring greater responsibility (Reeve et al., 2014). As indicated by Vallerand's (1997) hierarchical model of intrinsic and extrinsic motivation, the impact of social factors (interpersonal style) is mediated by basic psychological needs, these factors being key for the autonomy, competence and relatedness satisfaction needs, which lead to greater motivation and, in turn, greater adherence to physical practice. Therefore, autonomy support from the authority role is a determining factor that correlates positively with the satisfaction of basic psychological needs and with more autonomous motivation (Frielink, Schuengel, & Embregts, 2018), as well as with some consequences such as athletes' well-being (Balaguer, Castillo, & Duda, 2008), physical activity practice (Almagro et al., 2010; Almagro, Sáenz, & Moreno, 2012; Gillet, Vallerand, Amoura, & Baldés, 2010; Naisseh, Martinent, Ferrand, & Hautier, 2015) and adherence to sport practice (Haerens, Hirk, Cardon, De Bourdeaudhuij, & Vansteenkiste, 2010), among others.

Within the context of athletics, it has been found that greater intrinsic motivation corresponds to greater satisfaction with athletics practice and less boredom (Manzano & Valero, 2014), which may lead to an increase of the intention to be physically active in the future, although so far no studies have been found that confirm this hypothesis. In contrast to a traditional methodology characterized by a controlling style, there are alternative approaches, such as the Ludotechnic Model (Valero & Conde, 2003) or the Competitive Games proposed by Patón, Ferreiro, and Nemiña (2018). They put into practice an interpersonal style based on autonomy support that avoids repetition and monotony (which lead to loss of motivation), achieving an improvement in intrinsic motivation, competence and relatedness needs (Patón et al., 2018), as well as satisfaction with athletic practice and technical execution improvement (Morales, Valero, Manzano, & Jiménez, 2016).

Regarding the participants' motivational profiles and their future involvement in physical activity practice, authors such as Haerens et al. (2010) showed a relationship between motivation and intention to be physically active, obtaining higher values of intention to practice sports in the most self-determined profiles. Similarly, Friederichs, Bolman, Oenema, and Lechner (2015) obtained that a more self-determined profile was related to a more active lifestyle. However, none of these studies mentioned are really based on a sample of athletes. On the other hand, Haerens et al. (2017) found relationships between autonomy support, intrinsic motivation and satisfaction of basic psychological needs in a group of elite athletes. Despite this, the study does not include a variable that has at least one consequence derived from the complete sequence of Vallerand's (1997, 2007) hierarchical model, such as the intention to practice sports.

Taking into consideration all of the above, the main objective of our study was to determine the athletes' profiles according to the level of self-determined motivation and to analyze the differences in these profiles with respect to the interpersonal style, the basic psychological needs and the intention to be physically

active. We expect the study to confirm that there are relationships between the different variables included in the study and also different motivational profiles: one group with higher level of self-determined motivation, which will correspond to greater autonomy support, basic psychological needs satisfaction and intention to be physically active, compared to another group with lower level of self-determined motivation, which corresponds to low level of basic psychological needs satisfaction and intention to be physically active.

Method

Participants

The sample of this study initially comprised of 313 athletes who belonged to 15 clubs from the Spanish Athletics Federation. They had similar low and middle-level socio-demographic profiles, and they were selected based on accessibility and convenience. Inclusion criteria for participation in the study were: a) regular attendance in the Athletics school (≥ 70%) and b) completion of all the questionnaires.

After discarding the questionnaires that had not been carried out in their entirety, applying the statistical procedures for the detection of inconsistencies, as well as homogenizing the age of the participants, the final sample was composed of 254 athletes (97 males and 157 females) who participated in cross country official competition (from 2 to 4 kilometers), with ages ranging from 10 to 16 years old (M = 12.81, SD = 1.89) and mean training frequency of 3–4 times per week.

Material and Instruments

Autonomy Support. The Scale of Autonomy Support by Moreno, Huéscar, Andrés, and Sánchez (in press) was used. The questionnaire consists of eleven items that the participants have to answer about the coach's style in the sessions (e.g., "With his/her explanations, he/she helps us understand why the activities we do are useful"). The previous item used was: "In my athletic training, my coach ...". It consists of a 5-point Likert-type scale, from 1 (Surely not) to 5 (Surely), with internal consistency values of $\alpha = .65$, $\Omega = .77$.

Basic Psychological Needs. The Spanish version (Moreno, González-Cutre, Chillón, & Parra, 2008) of the Basic Psychological Needs in Exercise Scale (BPNES) was used. Participants answered on a Likert-type scale from 1 (Totally disagree) to 5 (Totally agree) including a total of 12 items. The questionnaire started with the following item: "In my physical education classes…". The internal consistency was $\alpha = .85$, $\Omega = .89$ for competence, $\alpha = .63$, $\Omega = .76$ for autonomy and $\alpha = .69$, $\Omega = .79$ for relatedness.

Motivation. The Spanish version by Moreno-Murcia, Marzo, Martínez, and Conte (2011) of the Behavioral Regulation in Sport Questionnaire (BRSQ) by Lonsdale, Hodge, and Rose (2008) was used. Participants answered on a Likert-type scale from 1 (Nothing is true) to 7 (Very true). The questionnaire started with the following sentence: "I participate in this sport ...". It is composed of 36 items that measure 9 categories with 4 items each: general intrinsic motivation (IM) (e.g., "Because I enjoy it"), IM knowledge (e.g., "For the pleasure that I get from knowing more about this sport"), IM stimulation (e.g., "For the enthusiasm I feel when I am involved in the activity") and IM execution (e.g., "Because I enjoy when I try to achieve long-term goals"); integrated regulation (e.g., "Because it is part of what I am"), identified regulation (e.g., "Because the benefits of sport are important to me"), introjected regulation (e.g., "Because I would feel ashamed if I abandon it") and external regulation (e.g., "Because if I do not do it, others would not be happy with me"), corresponding the four to extrinsic moti-

vation (EM); and amotivation (e.g., "However, I do not know why I do it). The internal consistency values were α = .74, Ω = .85 for general IM, α = .84, Ω = .89, for IM knowledge, α = .83, Ω = .89 for IM stimulation, α = .79, Ω = .87 for IM achievement, α = .84, Ω = .89 for integrated regulation, α = .73, Ω = .83 for identified regulation, α = .77, Ω = .85, for introjected regulation, α = .66, Ω = .79 for external regulation and, finally, α = .75, Ω = .85 for amotivation.

Intention to be physically active. We used the questionnaire called "Intention to be physically active" (IPA) by Hein, Müür, and Koka (2004) and validated for Spanish primary school children by Arias, Castejón, and Yuste, (2013) and for Spanish children over 12 years by Moreno, Moreno, and Cervelló (2007). This questionnaire is composed of 5 items. The introductory sentence used was: "Regarding your intention to practice some physical/sports activity ...". The answers were provided on a Likert-type scale from 1 (Totally disagree) to 5 (Totally agree). The reliability values were Cronbach's α = .70, Ω = .81.

Design and Procedure

A descriptive methodology with transversal design was used (Montero & León, 2007). The design was approved by the Ethics Committee of the University of Murcia (1414/2016). The necessary authorizations were obtained from the presidents of the Athletics Federation of Murcia Region and the clubs. Informed written consent was also obtained from athletes and parents when they were under-age. Once the consent was obtained, the questionnaires were administered. The researcher was present to give a brief explanation of the objective of the study, how to complete the instruments and to solve all the doubts that may arise during the process. After the information was provided, the participation of the athletes in the study was voluntary and anonymous. The time required to fill in the scales was approximately 15 minutes, varying slightly according to the age of the athletes. All questionnaires were answered on the athletics track prior to training and at the beginning of the season.

Statistical Analysis

First, descriptive statistics of all variables (means and standard deviations) was obtained, correlation analysis was conducted and the internal consistency of each factor was calculated using Cronbach's alpha and Omega coefficient. Most of the reliability Cronbach coefficients and all the Omega showed values above .70, a criterion considered acceptable for psychological domain scales (Nunnally, 1978). Concerning alpha coefficients only a few fell in the range between .60 and .70, considered acceptable by authors such as Sturmey, Newton, Cowley, Bouras, and Holt (2005). Furthermore, Ventura-León and Caycho-Rodríguez (2017) suggest the Omega coefficient has a more feasible value for social science studies, with non-continuous variables. It is not affected by the sample error or the number of the items, among other issues. Next, efforts were made to try to identify different motivational profiles in the study sample. To do so, a hierarchical cluster analysis was carried out with Ward method, using the variables intrinsic motivation towards knowledge, intrinsic motivation towards stimulation, intrinsic motivation towards execution, integrated, identified, introjected and external regulation, and amotivation.

Subsequently, a multivariate analysis (MANOVA) was performed to verify the possible differences between the motivational profiles and the coach's interpersonal style, the satisfaction of basic psychological needs and the intention to be physically active (dependent variables). Statistical analysis was carried out using the statistical software IBM SPSS Statistics 23.0.

Table 1 *Descriptive statistics (mean and standard deviation), Cronbach's alpha, Omega and correlations between variables*

		M	*SD*	*α*	*Ω*	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Autonomy S.	4.12	.48	.67	.77		.34**	.09	.41**	.37**	.44**	.47**	.37**	.31**	.40**	-.22**	-.28**	-.37**	.29**
2	Competence	5.12	.87	.85	.89			.01	.45**	.44**	.52**	.54**	.56**	.50**	.47**	-.16*	-.19**	-.28**	.35**
3	Autonomy	3.48	.96	.63	.76				.15*	-.01	.08	.07	.04	.12	.14*	.04	.13*	.04	.00
4	Relatedness	5.02	.72	.69	.79					.28**	.40**	.41**	.42**	.39**	.46**	-.13*	-.14*	-.20**	.33**
5	IM General	6.56	.73	.74	.85						.79**	.75**	.68**	.60**	.53**	-.25**	-.36**	-.40**	.47**
6	IM Knowledge	6.32	.90	.84	.89							.83**	.76**	.65**	.69**	-.18**	-.29**	-.35**	.53**
7	IM Stimulation	6.13	1.01	.83	.89								.86**	.79**	.71**	-.16*	-.27**	-.38**	.49**
8	IM Achievement	6.40	.81	.79	.86									.68**	.65**	-.15*	-.28**	-.36**	.44**
9	Integrated R.	5.87	1.21	.84	.89										.68**	.00	-.14*	-.24**	.49**
10	Identified R.	6.15	.93	.73	.83											**-.12***	-.20**	-.25**	.48**
11	Introjected R.	2.34	1.37	.77	.85												.65**	.50**	-.07
12	External R.	1.66	.93	.66	.79													.71**	-.17**
13	Amotivation	1.66	1.04	.75	.85														-.21**
14	IPA	4.66	.45	.70	.81														

Note. *M* = Mean; *SD* = Standard Deviation; *α* = Cronbach's alpha; *Ω* = Omega; *S* = Support; *IM* = Intrinsic Motivation; *R* = Regulation; *IPA* = Intention to be physically active.
**$p < .01$; *$p < .05$

Results

Descriptive Analysis and Correlations

The basic psychological needs obtained a mean score from 3.48 (autonomy) to 5.12 (competence). The most valued motivation was general IM and the least valued were amotivation and external regulation. The analysis of bivariate correlations reflected a significant and positive relationship between competence and relatedness ($p < .01$), but not with autonomy. The general IM had a significant and positive relationship with autonomy support style, competence and relatedness, IM dimensions, integrated regulation, identified regulation and intention to be physically active ($p < .01$). General IM had a negative and significant relationship with introjected regulation, external regulation and amotivation. Furthermore, the analysis gave a significant and positive relationship between the autonomy support style and the needs for competence and relatedness, as well as with IM dimensions towards knowledge, stimulation and achievement, integrated and identified regulation, and with the intention to be physically active ($p < .01$). In addition, the autonomy support style presented a significant and negative relationship with introjected and external regulation and amotivation in $p < .01$ (Table 1).

Cluster Analysis

The phases proposed by Hair, Anderson, Tatham, and Black (1998) were followed in order to carry out cluster analysis. First, participants who did not answer some test items, did not complete the test or did it incorrectly were excluded. In the next step, the univariate distribution of all the grouped variables was examined for normality. Hierarchical cluster analysis using the Ward method was performed to determine the groups that existed in the initial sample, and the dendrogram obtained suggested the existence of two groups (Table 2).

Following Norusis (1992), the small coefficients indicate great homogeneity among the members that make up the cluster, unlike the case of large samples. We conclude that there exist two motivational profiles: a self-determined or highly motivated profile, with high scores in IM (knowledge, stimulation and execution) and the most self-determined types of EM (integrated regulation and identified regulation; clus-

Table 2 M*ean, standard deviation and Z-score in clusters 1 and 2*

	Cluster 1 (n = 193) 76.0%			Cluster 2 (n = 61) 24.0%			
Variables	*M*	*SD*	*Z*	*M*	*SD*	*Z*	*F*
IM General	6.81	.34	.84	5.77	1.02	.03	169.693**
IM Knowledge	6.68	.43	.89	5.21	1.07	.16	310.688**
IM Stimulation	6.54	.52	.91	4.81	1.06	.20	353.604**
IM Achievement	6.73	.38	.86	5.40	0.96	.18	323.767**
Integrated R.	6.29	.84	.78	4.52	1.22	.35	190.963**
Identified R.	6.50	.55	.78	5.06	1.03	.25	242.984**
Introjected R.	2.20	1.29	-.34	2.89	1.49	.77	9.761**
External R.	1.46	.71	-.49	2.37	1.25	.77	44.549**
Amotivation	1.41	.71	-.56	2.69	1.49	.64	62.985**

Note. *M* = Mean; *SD* = Standard Deviation; *IM* = Intrinsic Motivation; *R* = Regulation.
**$p < .01$; *$p < .05$

ter 1) and a less self-determined or less motivated profile, with higher scores in amotivation, external regulation and introjected regulation (cluster 2). Differences of .50 in the Z-scores were used as a criterion to describe whether one group scored relatively "high" or "low" compared to the other (Wang & Biddle, 2001).

Differential Analysis

Multivariate analysis of variance (MANOVA) was performed to examine the characteristics of each motivational profile according to the coach's interpersonal style as perceived by the athletes, the satisfaction of their basic psychological needs and the intention to be physically active. To do so, the clusters were used as independent variables, and autonomy support, basic psychological needs (competence, autonomy and relationship with others) and intention to be physically active as dependent variables (Table 3).

The analysis of the results shows significant differences among clusters (Wilk's $\Lambda = .600$, $F(5,931) = 27.42$, $p < .01$) in autonomy support ($F = 48.32$, $p < .01$, $\eta^2 = .16$), competence ($F = 80.65$, $p < .01$, $\eta^2 = .24$), relationship with others ($F = 40.76$, $p < .01$, $\eta^2 = .13$) and intention to be physically active ($F = 76.59$, $p < .01$, $\eta^2 = .23$), yielding the highest scores in the self-determined profile.

Discussion

The objective of the study was to determine the existing profiles among the athletes according to the level of self-determined motivation and to analyze the differences in these profiles with respect to autonomy support, basic psychological needs and the intention to be physically active. The study's hypothesis was fulfilled by finding two profiles, one with higher levels of self-determination versus another one with low levels of self-determination. The more self-determined profile showed higher results in the autonomy support, two of the three satisfaction of basic psychological needs and the intention to be physically active.

Autonomy support was positively related to the psychological needs of competence and relatedness and the more self-determined motivation. These results coincide with those obtained in different studies, where relationships among autonomy support, satisfaction of basic psychological needs and motivation in athletes of various individual and collective sports

Table 3 *Multivariate analysis of interpersonal style, basic psychological needs and intention to be physically active according to the motivational profile*

	Cluster 1 (n = 193) 76.0%		Cluster 2 (n = 61) 24.0%			
Variables	*M*	*SD*	*M*	*SD*	*F*	η^2
Autonomy Support	4.23	.42	3.78	.47	48.32**	.16
Competence	5.36	.67	4.36	1.00	80.65**	.24
Autonomy	3.47	1.00	3.49	.83	.01	.02
Relatedness	5.17	.64	4.54	.75	40.76**	.13
IPA	4.78	.28	4.27	.64	76.59**	.23
Wilk's Λ						.600**
Multivariate F						27.42**

Note. *M* = Mean; *SD* = Standard Deviation; η^2 = Cohen's value (size effect); IPA = Intention to be physically active.
** $p < .01$

were reported (Almagro et al., 2012; Almagro et al., 2010; Balaguer et al., 2008; Gillet et al., 2010; Naisseh et al., 2015). It is not unusual in this kind of study to assess one of the three basic psychological needs (Almagro et al., 2010), or that one or more of the three basic psychological needs presents no correlation to the self-determination motivation, as has happened in this research (Balaguer et al., 2008), so it is recommended to pay attention to this in future studies.

The cluster analysis revealed the existence of two motivational profiles in the sample studied: a self-determined profile with high IM values and the most self-determined types of EM (integrated and identified regulation), and a less self-determined profile with higher values of introjected regulation (EM) and amotivation. These results coincide to a large extent with the study by Haerens et al. (2017) in a sample of elite athletes and with Almagro et al. (2012) in Spanish adolescent athletes. Although they are studies with different athletes and contexts, similar profiles were obtained, clearly distinguishing between participants with higher rates of internal motivation and others with greater external motivation or amotivation.

The analysis of variance showed that the self-determined profile was positively related to the interpersonal style of autonomy support and two of the three basic psychological needs, while the non-self-determined profile was negatively related to the autonomy support and the basic psychological needs of competence and relatedness. These results are in line with those obtained by Haerens et al. (2017) and Almagro et al. (2012). Therefore, regardless of the context, a more self-determined motivational profile is linked to athletes with greater satisfaction with some or even every one of their basic psychological needs and greater autonomy support provided by their mentors.

We found higher intention to be physically active in the self-determined profile, so the results obtained are in line with the profile studies by Haerens et al. (2010) and Friederichs et al. (2015) in school children and adults with low levels of physical activity. There are other studies that involved athletes, but did not carry out a profile analysis. Gillet et al. (2010) determined that a coach using an autonomy support style predicted a greater self-determined motivation and better sports performance in judokas of different categories and ages, while Almagro et al. (2010) found that the climate of autonomy created by the coach predicted the perception of autonomy of their athletes, their intrinsic motivation and adherence to sports practice in athletes of 12-17 years involved in different individual and collective sports.

As limitations of this study we must indicate that it is a cross-sectional and descriptive work, where causality relationships cannot be established. Furthermore, it should be taken into consideration that the intention to be physically active is not exactly the same as real physical activity and there could be some differences between them. On the other hand, future investigations may contemplate the possibility of carrying out a prediction analysis, to verify whether the relationships of this study can follow Vallerand's (1997) hierarchical model, considering the coach's interpersonal style as a trigger variable for motivation and basic psychological needs, and if, in turn, this leads to behaviors such as the intention to be physically active, life satisfaction or well-being. The aim of this research has been to assess the relationships among the motivational profile, the coach's interpersonal style and the intention to be physically active. Future studies with more accurate models (e.g., hierarchical analysis) should be developed to support SDT as significant determinants of athlete's intentions to be physically active and relationships among micro-theories of personality motivation. In addition, it would be interesting to analyze the existing differences based

on sociodemographic variables such as gender or age, something that was not contemplated in the present study. However, in the absence of more experimental evidence, as indicated by the pedagogical proposal of Koh, Camiré, Bloom, and Wang (2017), it seems essential that sports technicians be trained in interpersonal styles of autonomy support to achieve better behavioral, affective and cognitive effects in their athletes.

In conclusion, two motivational profiles have been found among the athletes: a profile with high levels of motivation (self-determined profile) that is related to the interpersonal style of autonomy support, the satisfaction of the needs of competence and relationships and greater intention to be physically active, and another profile with low motivation (non self-determined).

References

Almagro, B. J., & Paramio-Pérez, G. (2017). Motivación y adherencia a la práctica de baloncesto en adolescentes. *Cuadernos de Psicología del Deporte*, *1*(3), 189–198.

Almagro, B. J., Sáenz, P., & Moreno, J. A. (2010). Prediction of sport adherence through the influence of autonomy-supportive coaching among Spanish adolescent athletes. *Journal of Sports Science and Medicine, 9,* 8–14.

Almagro, B. J., Sáenz, P., & Moreno-Murcia, J. A. (2012). Perfiles motivacionales de deportistas adolescentes españoles. *Revista de Psicología del Deporte*, *21*(2), 223–231.

Arias, J. L., Castejón, F. J., & Yuste, J. L. (2013). Psychometric properties of the intention to be physically active scale in primary education. *Revista de Educación, 362,* 485–505. doi: 10.4438/1988-592X-RE-2013-362-23.

Balaguer, I., Castillo, I., & Duda, J. L. (2008). Apoyo a la autonomía, satisfacción de las necesidades, motivación y bienestar en deportistas de competición: Un análisis de la teoría de la autodeterminación. *Revista de Psicología del Deporte, 17*(1), 123–139.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, *19*(2), 109–134. doi: 10.1016/0092-6566(85)90023-6.

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behaviour. *Psychological Inquiry, 11*, 227–268. doi: 10.1207/S15327965PLI1104_01.

Friederichs, S. A., Bolman, C., Oenema, A., & Lechner, L. (2015). Profiling physical activity motivation based on self-determination theory: A cluster analysis approach. *BMC Psychology, 3*(1), 1–12. doi: 10.1186/s40359-015-0059-2.

Frielink, N., Schuengel, C., & Embregts, P. J. (2018). Autonomy support, need satisfaction, and motivation for support among adults with intellectual disability: Testing a self-determination theory model. *American Journal on Intellectual and Developmental Disabilities, 123*(1), 33–49. doi: 10.1352/1944-7558-123.1.33.

Gillet, N., Vallerand, R. J., Amoura, S., & Baldes, B. (2010). Influence of coaches' autonomy support on athletes' motivation and sport performance: A test of the hierarchical model of intrinsic and extrinsic motivation. *Psychology of Sport and Exercise*, *11*(2), 155–161. doi: 10.1016/j.psychsport.2009.10.004.

Haerens, L., Hirk, D., Cardon, G., De Bourdeaudhuij, I., & Vansteenkiste, M. (2010). Motivational profiles for secondary school physical education and its relationship to the adoption of a physically active lifestyle among university students. *European Physical Education Review, 16*, 117–139. doi: 10.1177/1356336X10381304.

Haerens, L., Vansteenkiste, M., De Meester, A., Delrue, J., Tallir, I., Vande Broek, G., Goris, W., & Aelterman, N. (2017). Different combinations of perceived autonomy support and control: Identifying the most optimal motivating style. *Physical Education and Sport Pedagogy*, *23*(1), 16–36. doi: 10.1080/17408989.2017.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (1998). *Multivariate data analysis*. Upper Saddle River, NJ: Prentice hall.

Hein, V., Müür, M., & Koka, A. (2004). Intention to be physically active after school graduation and its relationship to three types of intrinsic motivation. *European Physical Education Review*, *10*(1), 5–19. doi: 10.1177/1356336X04040618.

Koh, K. T., Camiré, M., Bloom, G. A., & Wang, C. K. J. (2017). Creation, implementation, and evaluation of a values-based training program for sport coaches and physical education teachers in Singapore. *International Journal of Sports Science and Coaching*, *12*(6), 795–806. doi: 10.1080/1750984X.2016.1199727.

Liao, Y., Chou, C. P., Huh, J., Leventhal, A., & Dunton, G. (2017). Associations of affective responses during free-living physical activity and future physical activity levels: An ecological momentary assessment

study. *International Journal of Behavioral Medicine*, *24*(4), 513–519.

Lonsdale, C., Hodge, K., & Rose, E. A. (2008). The Behavioral Regulation in Sport Questionnaire (BRSQ): Instrument development and initial validity evidence. *Journal of Sport and Exercise Psychology*, *30*(3), 323–355. doi: 10.1123/jsep.30.3.323.

Manzano, D., & Valero, A., (2014). Análisis del perfil motivacional de diversos grupos de atletas y su repercusión en el nivel de satisfacción deportiva. *Sport-Tk. Revista EuroAmericana de Ciencias del Deporte*, *2*(2), 9–19.

Martín-Albo, J., Núñez, J. L., & Navarro, J. G. (2003). La evolución motivacional como criterio discriminante en los deportes. *Revista Latinoamericana de Psicología, 35*, 139–150.

Montero, I., & León, O. (2007). A guide for naming research studies in psychology. *International Journal of Clinical and Health Psychology, 7*(3), 847–862.

Morales, M., Valero, A., Manzano, D., & Jiménez, J. (2016). Efectos de una unidad didáctica ludotécnica en el aprendizaje del salto de altura de alumnos de bachillerato. *Agora*, *18*(2), 199–213.

Moreno, J. A., González-Cutre, D., Chillón, M., & Parra, N. (2008). Adaptación a la educación física de la escala de las necesidades psicológicas básicas en el ejercicio. *Revista Mexicana de Psicología, 25*(2), 295–303.

Moreno, J. A., Marzo, J. C., Martínez, C., & Conte, L. (2011). Validation of Psychological Need Satisfaction in Exercise Scale and the Behavioural Regulation in Sport Questionnaire to the Spanish context. *Revista Internacional de Ciencias del Deporte, 7*(26), 355–369. doi: 10.5232/ricyde.

Moreno, J. A., Moreno, R., & Cervelló, E. (2007). El autoconcepto físico como predictor de la intención de ser físicamente activo. *Psicología y Salud, 17*(2), 261–267.

Moreno, J. A., Huéscar, E., Andrés, A., & Sánchez, J. (in press). El estilo interpersonal del profesor en las clases de educación física, del control al apoyo a la autonomía. *International Journal of Performance Analysis in Sport*.

Naisseh, M., Martinent, G., Ferrand, C., & Hautier, C. (2015). Relationship between parents' motivation for physical activity and their beliefs, and support of their children's physical activity: A cluster analysis. *Psychological Reports*, *117*(1), 230–243. doi: 10.2466/06.21.PR0.117c17z0.

Norusis, M. J. (1992). *SPSS/PC+ Professional statistics, Version 5.0*. Chicago, IL: SPSS.

Nunnally, J. C. (1978). *Psychometric Theory*. Nueva York: McGraw-Hill.

Patón, R. N., Ferreiro, M. C., & Nemińa, R. E. (2018). Effect of a didactic unit based on competitive games in the motivation, basic psychological needs and enjoyment in Primary Education students. *Sportis. Revista Técnico-Científica del Deporte Escolar, Educación Física y Psicomotricidad, 4*(1), 111–125. doi: 10.17979/sportis.2018.4.1.2900.

Reeve, J., Vanteenkiste, M., Assor, A., Ahmad, I., Cheon, S., Jang, H., & Wang, J. (2014). The beliefs that underlie autonomy-supportive and controlling teaching: A multinational investigation. *Motivation and Emotion*, *38*(1), 93–110. doi: 10.1007/s11031-013-9367-0.

Sturmey, P., Newton, J. T., Cowley, A., Bouras, N., & Holt, G. (2005). The PAS-ADD checklist: Independent replication of its psychometric properties in a community sample. *British Journal of Psychiatry, 186,* 319–323. doi: 0.1192/bjp.186.4.319

Ulrich-French, S., & Smith, A. L. (2009). Social and motivational predictors of continued youth sport participation. *Psychology of Sport and Exercise*, *10*(1), 87–95. doi: 10.1016/j.psychsport.2008.06.007.

Valero, A. (2004). Situación actual de la metodología en la iniciación deportiva al atletismo. *Revista de Educación Física, 94,* 13–20.

Valero, A., & Conde, J. L. (2003). *La iniciación al atletismo a través de los juegos*. Málaga: Aljibe.

Vallerand, R. J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. *Advances in Experimental and Social Psychology*, *29*, 271–360. doi: 10.1016/S0065-2601(08)60019-2.

Vallerand, R. J. (2007). Intrinsic and extrinsic motivation in sport and physical activity. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology, 3ª ed.,* (pp. 59–83). New York: Wiley.

Ventura-León, J. L., & Caycho-Rodríguez, T. (2017). El coeficiente Omega: Un método alternativo para la estimación de la confiabilidad. *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud, 15*(1), 625–627.

Wang, C. J., & Biddle, S. J. (2001). Young people's motivational profiles in physical activity: A cluster analysis. *Journal of Sport and Exercise Psychology*, *23*(1), 1–22. doi: 10.1123/jsep.23.1.1.

Studia Psychologica, Vol. 61, No. 2, 2019, 120-131 doi: 10.21909/sp.2019.02.777

Cyberbystanders, Affective Empathy and Social Norms

Magdalena Kozubal, Anna Szuster, Julia Barlińska
The Faculty of Psychology, University of Warsaw, Poland

The main aim of the study was to examine the influence of affective empathy and social norms on preventing behavior of cyberbystanders reinforcing cyberbullying. 219 students took part in an experiment conducted in junior high and high schools from three Polish school districts. The goal of the experiment was to check whether the students would forward or delete a humiliating picture. The results indicate a strong impact of previous experiences as a cyberperpetrator on cyberbystanders' reinforcing behavior and a relevant effect of affective empathy activation, which decreased the frequency of cyberbullying enhancing behavior. No significant effect of gender or norm activation was found. Bystanders' negative cyberbullying behavior was effectively reduced through norm priming only in the case of those individuals who were able to appropriately verbalize the contents of violated norms. It indicates that the regulatory role of social norms is subject to cognitive understanding of their contents.

Key words: cyberbullying, cyberbystanders, empathy, affective empathy, social norms

Introduction

In today's world, dominated by technology, the advantages of social media are indisputable. However, the same media are also responsible for generating negative social behaviors like cyberbullying, which violate universally shared norms and cause harm to others (Sproull, Conley, & Moon, 2005).

Cyberbullying is defined as a kind of behavior exhibited by individuals or groups through electronic or digital media, which is marked by repeatedly communicated, hostile or aggressive messages intended to inflict harm or discomfort on others (Tokunaga, 2010, p. 278). It is a phenomenon of peer-to-peer electronic bullying, which involves the presence of a victim, a bully and bystanders (Barlińska, Szuster, & Winiewski, 2013, 2015; Smith, 2011). Bystanders' reaction to bullying acts may be that of supporting the victim, acting as outsiders or assisting and reinforcing the bullying act (Salmivalli, 2010).

The data concerning bystanders suggest that although young people frequently witness cyberbullying, only a limited number of them actually resolve to intervene (Gini, Albiero, Benelli, & Altoe, 2008; Barlińska et al., 2013, 2015). According to the previous research, such factors as empathy, prosocial norms, self-efficacy, high extraversion, severity of a bullying incident and a close relationship with the victim are conducive to the bystander's intervention (Freis & Gurung, 2012; Thornberg et al., 2012; Macháčková et al., 2013; Barlińska et al.,

Acknowledgement
The project is co-financed by the National Research and Development Centre – "Social Innovations" program.

Correspondence concerning this article should be addressed to Magdalena Kozubal, Faculty of Psychology, University of Warsaw, Stawki 5/7, 00-183 Warsaw, Poland.
E-mail: magdalena.kozubal@psych.uw.edu.pl

Received December 2, 2018

2013). In our research we focused on two factors that appear most significant for impacting bystander's online behavior: affective empathy and social norms.

Empathy Activation as a Solution to the Problem of Cyberbullying

A long history of philosophical reflection and almost a hundred years of psychological studies on empathy have established its status as a unique phenomenon in social relations (cf. Hume, 1739/1968; Smith, 1759/1976; Spencer, 1870; Titchener, 1909; Eisenberg & Strayer, 1987; Davis, 1996; Hoffman, 2000; Batson, 2010; Cuff et al., 2016). A majority of authors pointed out that feelings experienced by the subject witnessing states of others are the essence of empathy. Nowadays, in part due to the discoveries of neuroscience (Preston & de Wall, 2002; de Vignemont & Singer, 2006; Decety, 2007), empathy typically refers to two separate categories of the phenomena, with distinction being made between its affective and cognitive variant, i.e. between empathizing based on feeling vs. understanding other people's states. Cognitive empathy is the ability to understand another person's feelings related closely to the theory of mind (Blair, 2005). Affective empathy is concerned with the experience of emotion, elicited by an emotional stimulus. The two types of empathy have basic common attributes: 1) focus on others, 2) a shortened psychological distance, and 3) feeling of closeness.

Empathy plays a fundamental social role as it allows individuals to share experiences, needs and common goals. Its most frequently mentioned aspect is social significance and the benefits associated with morality, altruism, inhibition of aggression, prosocial and helping behavior (Batson & Shaw, 1991; Hoffman, 2006).

The exploration of cyberbullying among adolescents focuses on the developmental aspects of empathy and its basic affective mechanisms as they are believed to be innate. Affective empathy is the process of analogous emotional reacting to the incoming stimulus (Eisenberg, 2000). A necessary condition for affective empathy to arise is the presence of another person. A direct contact with such universal human attributes as mimical expressions and eye contact (both optimal and suboptimal), posture or physical distance (Agryle, 1994) activates affective empathy mechanisms. As these elements of contact are substantially less accessible in online conditions (Kiesler, Siegel, & McGuire, 1984), natural mechanisms controlling aggression tend to significantly decrease in cyberspace (Baron-Cohen, 2011). Therefore, the search for effective mechanisms activating empathy and preventing aggressive online behaviors becomes a priority.

Triggering empathy results in fewer aggressive responses to provocation (Richardson, Green, & Lago, 1998). It evokes a sense of guilt, which, in turn, can reduce antisocial behavior (Hoffman, 2001). Research findings clearly point to a variety of connections between cyberbullying and empathy. Affective empathy was found to be lower in case of adolescents who formerly acted as cyberperpetrators than in case of those who did not engage in bullying acts (Renati, Berrone, & Zanetti, 2012; Berne et al., 2013). Also, the degree of empathy was found to be lower in case of cybervictims, while individuals with higher dispositional empathy, when confronted with cyberbullying acts, are more likely to intervene in a prosocial manner (Freis & Gurung, 2012; Macháčková et al., 2013). Both the affective and the cognitive aspect of empathy reduce cyberaggression (Barlińska et al., 2013, 2015). Nevertheless, more evidence for a limiting influence of empathy on bullying and cyberbullying behavior was gathered in the case of affective empathy (Jolliffe & Farrington, 2004; Renati, Berrone, & Zanetti, 2012). Empa-

thy has been successfully introduced as a part of anti-bullying intervention (Twemlow, Fonagy, Sacco, Gies, & Hess, 2001) and prevention programs at schools (Palladino, Nocentini, & Menesini, 2012; Wölfer et al., 2014; Williford et al., 2013).

An empathic response may prompt a bystander to react in a supportive, non-aggressive manner (Macháčková et al., 2013). Another study confirmed that adolescents in conditions of activated empathy were less likely to forward a cyberbullying message as a bullying reinforcing response (Barlińska, Szuster, & Winiewski, 2013, 2015).

Digital tools, with their superficiality and multitasking character (Carr et al., 2003), justify the search for such an empathy activation strategy that will be compatible with cyberspace attributes. The chances of concentrating one's attention on a clear, accessible and universally recognized stimulus are significantly higher in comparison to a more complex content.

The research using Magnetic Resonance Imaging shows that mere exposure to facial expressions of various emotions results in increased arousal in those parts of the observers' brains that are involved in producing such expressions (Carr et al., 2003). According to the perceptional model of empathy (Preston & De Wall, 2002), both observing and imagining what another person feels automatically triggers the neural pathways responsible for representing the affective states of the observed person. With these mental representations it is possible to recognize other people's emotions and express them. The existence of common affective neural pathways may explain how we can experience other people's emotions as our own. Other studies on neural mechanisms underlying empathy showed that exposure to a face expressing sadness or pain is enough to activate mirror neurons, which are also responsible for arousing empathy in more complex situations involving another person's needs (De Vignemont & Singer, 2006).

The findings point out that affective empathy activated by priming of a human face can be an effective strategy in limiting bystander's negative behavior.

Norms Activation as a Solution to the Problem of Cyberbullying

In common understanding (popular also among psychologists) norms are perceived as a set of social "dos" and "don'ts", as shared expectations of how we should act, reinforced by a threat of a group sanction or a promised reward (Schwartz & Howard, 1981). In every society, selfless acts, the ones filled with compassion, are rewarded, whereas violence towards others is punished. In this context social norms are a set of rules that regulate displays of aggression and social life.

Social norms are instilled through the education process whereby expectations towards an individual are verbalized or showed by social objects (parents, teachers, peers) by means of non-verbal factors. During cognitive development the internalization of norms takes place. The process runs from applying norms only in the externally controlled conditions to including them in the Self-structure. Then, when a norm is violated, the internal sanctions, such as shame, sense of guilt, and lower self-esteem, emerge (Schwartz & Howard, 1981).

The standards of online behavior mostly remain non-codified and fairly relativized. Also, parents and teachers are not as present in cyberspace as in reality. Therefore, teenagers in their online peer groups may have difficulties with the creation and compliance with their own norms and standards.

The regulatory role of norms was evidenced by a variety of research. The mere reading of the stories describing violation of social norms

activates those brain areas, which are responsible for recognizing mental states of others and responding to the aversive emotional expressions, particularly anger (Berthoz, Armony, Blair, & Dolan, 2002). Breaking a norm (punishing a partner with an electric shock for making a mistake or lying) was found to increase helpful behavior in the next task performed in a different context (Berkowitz, 1972). Once a rule is breached, a desire to redress appears. In Macaulay's (1970) research, during a fundraising campaign for hungry children, the behavior of passers-by was affected by the demonstrative behavior of the experimenter's assistant, who either refused or offered a donation. In both cases the proportion of people offering donations increased compared to the group, which was not confronted with any behavior priming pattern. According to the researcher's interpretation, observation of other person's behavior not only provides a behavioral pattern, but also plays a role as the already acquired norm activation factor. This kind of activation may also effectively reduce negative behaviors on the Internet.

The regulatory role of social norms in bystander's reactions to cyberbullying has not been that widely studied. However, some pieces of evidence suggest that it also affects the regulation of teen behavior in cyberspace. Positive peer injunctive norms concerning the social life at school are connected with a smaller number of cyberbullying experiences. In contrast, negative peer injunctive norms regarding the approval of risky behaviors result in a greater number of such experiences (Pyżalski, 2013). Research shows that adolescent bystanders overestimate reinforcement of bullying among their peers, which is positively correlated with active assisting in bullying acts (Sandstrom, Makover, & Bartini 2012). A bystander's belief that bullying is wrong is a factor motivating to intervene. However, if a bystander believes that such an intervention is not his/her moral responsibility, this, in turn, can lead to a passive or a negative behavior (Macháčková, 2013; Thornberg et al., 2012).

The results of one of the few longitudinal studies on cyberbullying showed that social norm violating behaviors are conducive to involvement in cyberbullying (Ybarra & Mitchell, 2004; Sticca, Ruggieri, Alsaker, & Perren, 2013, 2015). Thus, the content of one's norms may affect the choice of reactions towards cyberbullying acts, i.e. the decision to remain passive or to reinforce bullying.

The data presented show that cyberbullying reinforcing behaviors may be limited through exposure to norm violating behaviors.

The Role of Cyberperpetration Experience and Gender

The previous data confirmed that exposure to aggression increases the probability of accepting different bullying forms and tolerating violence both in reality and in cyberspace (Ybarra & Mitchell, 2004; Völlink, Bolman, Dehue, & Jacobs, 2013). Being a perpetrator of cyberbullying is an important predictor of cyberbystander's bullying reinforcing behavior (Barlińska et al., 2013, 2015; Macháčková et al., 2013). Therefore, it was expected that in the study discussed above the said dependency would be replicated.

The empirical findings concerning the relationship between gender and cyberbystander's behavior are equivocal. In some studies females were found to be offering greater support and assistance than males when witnessing cyberaggression, whereas in others no sex differences with regard to bystander's both positive and negative reactions to cyberbullying were reported (Barlińska et al., 2013, 2015; Macháčková et al., 2013). Therefore, in the present study the role of gender was also monitored.

Aim of the Study

In the current study of an experimental character, the focus was on the factors preventing cyberviolence among bystanders. We concentrated on priming empathy and social norms as well as on the influence of cyberbullying perpetrators' gender and experiences.

The effectiveness of empathy activation in preventing cyberbullying reinforcing behavior via watching a testimony of a cyberbullying victim has been proved in our previous studies (Barlińska et al., 2013; Barlińska et al., 2015). In the current study, we focused only on affective empathy priming methods activated through the exposure of a human face. Sufficient data points to the regulatory potential of social norms in bystander's reactions to cyberbullying. We expected that cyberbullying reinforcing behavior may be inhibited through exposure to norm violating behaviors and their verbalization. Cyberperpetration experience and gender are two other factors included in the study, which were proved to modify adolescents' reactions to cyberbullying.

The main hypotheses were as follows:

Priming of both affective empathy and social norms would decrease the frequency of cyberbullying reinforcing behaviors.

The experience of cyberbullying as its perpetrator would increase the frequency of cyberbullying reinforcing behaviors.

Method

Participants

The sample consisted of $N = 219$ pupils from junior high schools and high schools (118 boys and 101 girls), aged 12–18 (Mage = 14.77, SDage = 1.43 years). In terms of activating affective empathy, the number of participants in the group was 63 ($N = 63$), in the conditions of activating the normative system 59 ($N = 59$) and in a control conditions 97 ($N = 97$).

Instruments

Faces

In order to activate affective empathy, 14 pictures, taken from the Ekman and Friesen's ''Pictures of Facial Affect'' collection (POFA; 1976), had been uploaded on a special online platform. The pictures presented 6 male and 8 female faces expressing sadness. The adolescents were asked to mark whether they liked a given face or not (a masking task). After choosing an answer another picture appeared on the screen.

The Pictorial Test of Social Incompatibility

In order to activate social norms, the Pictorial Test of Social Incompatibility (PTSI) was used. This questionnaire consisted of 11 black and white schematic pictures. Each of them presented a norm violating behavior. For instance, a young man with a baseball bat demolishing a bus stop. The task was to answer the question: "What's wrong in this picture?" At the beginning there was an exemplary picture with the previously provided answer. The effectiveness of the "norm priming" procedure can be confirmed by the results of the study referred to in the introduction (see Macaulay, 1970; Berkowitz, 1972).

The measure of the correctness of recognizing violated social norms was the average of the answers identifying the content of the norms (for example, "this is stealing," or "you cannot steal"). One point was given for a correct answer and 0 points for a wrong answer. Half a point was in turn given for those answers which identified, for example, legal rules, such as traffic regulations (a driver should not speed), or unwritten safety rules (a woman's handbag should be closed).

The ratio of correctness was determined by the average of the answers, calculated by means of the following formula: Correctness of Recognizing Social Norms = (n1 + n2 + ... + n11) / 11, where n1, n2, ..., n11 signify points given for answers to each of the 11 questions asked. The higher the ratio, the more substantial the correctness of recognizing the displayed standards[1].

Message from a Peer[2]

"Message from a friend" simulated a social networking website by means of which the pupils had a short chat with a virtual friend, who at the end encouraged them to send a message insulting a different pupil (a photomontage presenting a dog with a boy's head). The participants could choose between sending forward the insulting message (cyberbystander's reinforcing behavior) or deleting it (neutral behavior).

Cyberbullying Questionnaire

A scale of cyberperpetration experience from the Questionnaire of Cyberbullying experience (Barlińska & Wojtasik, 2008) was employed. Responses were given on a 4-point scale (1 = never; 4 = several times). The incidence of cyberperpetrator experience was the averaged score for each question (e.g., "Have you ever blackmailed someone on the Internet?"). It proved to be internally consistent, $\alpha = 0.83$ (*M*perpetrator = 1.03, *SD* = 2.01). The scores were used in further analyses.

Procedure

This experimental study was conducted by using a web application that simulated a social networking site and a messaging service[3]. It took place in computer labs, in small groups. The participants were randomly assigned to the experimental or control groups. Anonymity was guaranteed by the use of unique one–time codes which allowed access to an experimental web application. In the first group (empathy activation condition) the participants were shown sad faces on the monitor, whereas in the second group (norm activation condition) the participants were shown schematic drawings illustrating behaviors that break social norms (The Pictorial Test of Social Incompatibility). In the third group – control condition there was no manipulation. Next, all participants took part in an online task called "Message from a peer".[4] They were asked to make a choice between sending or deleting offensive material. Finally, students completed a Cyberperpetrator experience questionnaire.

Plan of Analysis

The independent variables were both of a situational (activation of a normative system, activation of empathy) and dispositional (ability to recognize norms correctly) character. The controlled factors were gender and previous cyberperpetration experience. The dependent variable was online behavior (reinforcing vs.

[1] The schematic pictures were evaluated by three psychologists. It was assessed to what extent the material presented by these pictures was identified as the violation of social norms. Those pictures which the psychologists assessed in a fully compatible way were selected and included in the study. The responses of its participants were assessed according to the key prepared by the psychologists.

[2] Social desirability was controlled; no correlations were found (Barlińska, Szuster, & Winiewski, 2013).

[3] The study was approved by the ethics committee of the Faculty of Psychology of Warsaw University

[4] A detailed description of the method is presented in Barlińska, Szuster, and Winiewski (2013), Cyberbullying among adolescent bystanders: Role of the communication medium, form of violence and empathy, *Journal of Community and Applied Social Psychology*, *23*, 37-51. doi: 10.1002/casp.2137

neutral) in the situation of being a bystander of a cyberbullying act.

Plan of Analysis I

Logistic regression analyses were conducted in order to evaluate whether the activation of affective empathy and social norms reduced the likelihood of cyberbystander's reinforcing behavior. Additionally, we tested whether cyberperpetrator experience increased the frequency of choosing the reinforcing versus neutral reaction, and whether gender had any relationship to the violence enhancing actions. The logistic regression model was chosen due to having a dichotomous dependent measure and several continuous and binary predictors. It is reported following Peng, Lee, and Ingersoll's (2002) guidelines.

Plan of Analysis II

The Pictorial Test of Social Incompatibility used for priming of norms measures also the correctness of norm recognition. In the second analysis this variable was taken into account. As in the first analysis, logistic regression was conducted to evaluate whether the correctness of recognizing violated social norms would decrease the likelihood of cyberbullying reinforcing behavior of a bystander.

Results

A little over 20.5% ($n = 52$) of the sample selected the behavior that reinforced cyberbullying. The logistic regression analysis was conducted, with the selected behavior (0 = neutral, 1 = cyberbullying reinforcing behavior) as the dependent variable. Table 1 presents the figures of the odds ratio coefficients with 95% confidence intervals, Wald statistics with the level of significance for each variable at each stage of the analysis, overall model match statistics, and the selected detailed parameters. The model suggested a good match (Likelihood ratio test = 62.13, $df = 1$, $p < .001$; Omnibus test = 20.72, $df = 4$, $p < .001$) and reasonably good predictive abilities (Cox & Snell $R^2 = 0.09$; Nagelkerke $R^2 = 0.14$).

Table 1 *The results of the logistic regression analysis for the experimental conditions (activation of affective empathy and activation of social norms), experience as a cyberperpetrator, the influence of the gender category and cyberbystander behavior*

Predictor	*B*	*SE β*	Wald's χ^2	*OR (95% CI)*
Experimental conditions			8.94*	
Act. of empathy (Faces)	-1.80	0.63	8.21**	0.16 (0.05-0.57)
Act. of social norms (Pictures)	0.02	0.38	0.01	1.02 (0.48-2.20)
Exp. cyberperpetrator	0.38	0.11	11.75***	1.46 (1.18-1.81)
Gender[a]	-0.33	0.35	0.89	0.77 (0.3-1.3)
Overall model			χ^2	
Likelihood ratio test		62.13***		
Score test		20.72***		
Hosmer & Lameshow test		12.64		

Note. [a] Gender was coded females = 1, males = 2
Cox & Snell $R^2 = .09$; Nagelkerke $R^2 = .14$
***p > .001; **p > .01; *p > .05.

On the basis of the obtained relevant odds ratio coefficients, it can be concluded that the probability of cyberbystanders' reinforcing behavior in the activation of affective empathy condition (face images) is considerably lower than in the control condition (the odds ratio coefficient, $OR = 0.16$). Furthermore, the probability of cyberbullying reinforcing behavior of a bystander increases with the intensity of the cyberperpetration experience ($OR = 1.50$). The activation of social norms and the influence of gender did not have significant effects.

The results are consistent with previous studies (Barlińska et al., 2013, 2015). They show a strong impact of experiences as a cyberperpetrator. The experience of being a perpetrator increased the frequency of behavior that reinforced cyberbullying. These students more often sent messages ridiculing their peer. In turn, empathy priming was proved analogically effective in decreasing the frequency of cyberbullying enhancing behavior of a cyberbystander. No effect of norms activation was obtained as well.

Table 2 presents the figures of the odds ratio coefficients with 95% confidence intervals, Wald statistics with the level of significance for each variable at each stage of the analysis, overall model match statistics, and the selected detailed parameters. The model suggested a good match (Likelihood ratio test = 12.96, $df = 1$, $p < .001$; Omnibus test chi-square = 17.40, $df = 1$, $p < .001$) and good predictive abilities (Cox & Snell $R^2 = 0.26$; Nagelkerke $R^2 = 0.38$).

On the basis of the obtained relevant odds ratio coefficients, it can be concluded that the probability of cyberbystander's reinforcing bullying behavior is considerably lower when the correctness of recognizing social norms is higher (odds ratio coefficient $OR = 0.00$).

Discussion

The study aimed at exploring the regulatory role of factors modifying adolescent cyberbystander pro-bullying behavior. The results clearly confirm the influence of cyberperpetration experience on perpetrators' activities in cyberspace. Bystanders tend to support other, often unknown web users in performing harm inflicting acts. It is coherent with the already obtained findings of the research on bystander's behavior (Barlińska et al., 2013, 2015; Macháčková et al., 2013). Although no significant effect of gender or norm activation was found, further analysis showed that correct social norm recognition diminishes the likelihood of bystander's negative behavior, which stays in line with some previous findings (Thornberg et al., 2012). Affective empathy priming led to the decreased frequency of bystander's negative behavior. The results concerning the modifying role of empathy a rereplicative (Barlińska et al., 2013, 2015;

Table 2 *The results of the logistic regression analysis for the recognition of social norms and cyberbystander behavior*

Predictor	*B*	*SE β*	Wald's χ2	*OR (95% CI)*
Recognition of social norms	-6.37	1.98	10.34***	0.00- (0.00-0.0873)
Overall model			χ2	
Likelihood ratio test		12.96***		
Omnibus test		17.40***		
Hosmer & Lameshow test		5.18		

Note. Cox & Snell $R^2 = .26$; Nagelkerke $R^2 = .38$
***p > .001; **p > .01; *p > .05.

Macháčková et al., 2013). As opposed to a majority of other studies (in which questionnaire empathy measures were used), in our experimental research the priming material contained a simple stimulus of perceptively expressive character, which induced empathic arousal. It appears that exposing cyberbullying bystanders to a human face expressing sadness can modify their behavior and neutralize actions that reinforce cyberbullying. Proving the effectiveness of this stimulus in cyberspace dominated by indirect contact can yield profound practical implications. The main objective of future research should be to find better solutions for adapting empathy arousing methods to the "language" of the Internet by designing websites and programs devoted to the consequent activation of empathy. The advantage of the employed method was that the stimulus (human face) was in no way connected with the adolescents' behavior recorded in the study. Therefore, the effectiveness of such an activation strategy appears particularly worth considering. In order to activate affective empathy mechanisms, standardized photos of a human face (from the Ekman's catalogue), the effectiveness of which was also confirmed by the research with the use of FMR (Carr et al., 2003), were used with a basic emotion, i.e. sadness, exposed. Nevertheless, the question regarding the effectiveness of neutral face exposure arises. Further research is expected to differentiate which factor is the key to the effectiveness of this manipulation: a face itself or the exposure of a basic emotion.

The simple effect of priming through the exposure of a social norm violation proved to be ineffective in decreasing cyberbullying reinforcement. However, further analysis revealed an additional condition: the priming effect of norm exposure had an impact on negative online behavior when the content of norms was correctly recognized. Contrary to the innate and primal character of affective empathy mechanisms, the regulatory role of norms appears as a result of complex cognitive mechanisms that require social learning (Bandura, 1973). As norms are of an abstract nature, their adoption is a long-term, individually motivated process dependent upon cognitive and social development. This is also a challenge for educators and teachers to help children verbalize and adopt social norms by naming them, discussing them, showing their purpose and significance. Outside of the family home, school is the most important socialization environment. One's ability to identify and verbalize a norm is not only indicative of one's knowledge of the norm existence. It is a proof that such a norm has been understood and processed. This, in turn, profoundly affects the regulatory character of norms and opens up new possibilities for an individual to view a given phenomenon from various standpoints. The ability to verbalize is a manifestation of reflective processing (Kahnemann, 2011) that increases control upon frequently automatic online reactions (like the clicking of the "forward" button). This process enables the rules observed by an individual in a specific situation to be further generalized and applied to a different social context, also to cyberbullying. In cyberspace, young people make choices without adult supervision, which also applies to cyberbullying and its witnessing. The achieved results indicate that correct norms internalization leads to their implementation in the virtual world as well. This is, however, a long-term process which largely depends upon a socialization process and individual differences (Hoffman, 2000). Thus, it is necessary to focus on social norms in the Internet environment and devise preventive methods based on clearly determined specific cyberbullying situations, such as teaching pupils how to intervene effectively as responsible and active bystanders both online and offline.

The current investigation has its strengths and limitations. The main strength is the general design using an innovative experimental approach. Among the main limitations is the narrowed measurement of the dependent variable, which has been operationalized by only one of various cyberbullying acts and giving the participants a limited range of behavioral options. Following this, a conclusion about the effectiveness of both affective empathy and norm activation is limited only to this specific form of cyberbullying. Another limitation, resulting from the ethical nature of experimental methods, is the fact that the cyberbullying situation was only simulated. Future experimental studies on effectiveness of inductions in diminishing the scale of cyberbullying reinforcement from bystanders should use more complex stimuli, which will raise the studies' ecological validity.

The sample included adolescents aged 12 to 18. Since affective empathy is based mostly on basic neurological mechanisms, it can be significant when it comes to social norms, the internalization of which depends also on the stage of development. In further studies on cyberbullying, the preventing role of normative systems should also be investigated.

Empathy activation and correct norm recognition were found to be the two major factors effectively reducing bystander's reactions reinforcing a potential cyberbullying act. The findings support previous results indicating the importance of sensitizing bystanders to a potential harm caused by cyberbullying and of encouraging bystander's positive behavior through the enhancement of empathy skills and the development of anti-bullying norms (Barlińska et al., 2013; DeSmet et al., 2014; Thornberg et al., 2012).

References

Ang, R. P., & Goh, D. H. (2010). Cyberbullying among adolescents: The role of affective and cognitive empathy, and gender. *Child Psychiatry and Human Development*, *41*(4), 387–397. doi: 10.1007/s10578-010-0176-3

Argyle, M. (1994). *The psychology of interpersonal behaviour*. Penguin UK.

Bandura, A. (1973). *Aggression: A Social Learning Analysis*. Prentice-Hall, Englewood Cliffs, N. J.

Barlińska, J., & Wojtasik, Ł. (2008). Peer violence and electronic media – research and social campaign. In M. Barbovschi & M. Diaconescu (Eds.), *Teenagers' actions and interactions online in Central and Eastern Europe. Potentials and empowerment, risks and victimization* (pp. 281–299). Cluj-Napoca, Romania: Cluj University Press Babes- Bolyai University.

Barlińska, J., Szuster, A., & Kozubal, M. (2016). In search of a simple method: Is a human face an effective, automatic filter inhibiting cyberbullying? In M. F. Wright (Ed.), *A social-ecological approach to cyberbullying*. Hauppauge: Nova Publishing. ISBN: 978-1-63483-755-2

Barlińska, J., Szuster, A., & Winiewski, M. (2015). The role of short-and long-term cognitive empathy activation in preventing cyberbystander reinforcing cyberbullying behavior. *Cyberpsychology, Behavior and Social Networking*, *18*(4), 241–244. doi: 10.1089/cyber.2014.0412

Barlińska, J., Szuster, H., & Winiewski, M., (2013). Cyberbullying among adolescent bystanders: Role of the communication medium, form of violence and empathy,. *Journal of Community and Applied Social Psychology, 23,* 37–51. doi: 10.1002/casp.2137.

Baron-Cohen, S. (2011). *The science of evil*. London: Basic Book ISBN-10: 0465031420

Bastiaensens, S., Vandebosch, H., Poels, K., Van Cleemput, K., DeSmet, A., & De Bourdeaudhuij, I. (2014). Cyberbullying on social network sites. An experimental study into bystanders' behavioral intentions to help the victim or reinforce the bully. *Computers in Human Behavior*, 31, 259–271. doi: 10.1016/j.chb.2013.10.036

Berkowitz, L. (1972). Social norms feelings, and other factors affecting helping and altruism, In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, *6* (pp. 63–108). New York: Academic Press.

Berne, S., Frisén, A., Schultze-Krumbholz, A., Scheithauer, H., Naruskov, K., Luik, P., Katzer, C., Erentaite, R., & Zukauskiene, R. (2013). Cyberbullying assessment instruments: A systematic review. *Aggression and Violent Behavior, 18*(2), 320–334. doi: 10.1016/j.avb.2012.11.022

Berthoz, S., Armony, J. L., Blair, R. J. R., & Dolan, R. J. (2002). An fMRI study of intentional and unin-

tentional (embarrassing) violations of social norms. *Brain*, *125*(8), 1696–1708. doi: 10.1093/brain/awf190

Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition*, *14*(4), 698–718. doi: 10.1016/j.concog.2005.06.004

Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences*, *100*(9), 5497–5502. doi: 10.1073/pnas.0935845100

Chalmers, J. B., & Townsend, R. (1990). The effect of training in social perspective taking on socially maladjusted girls. *Child Development, 61*, 178–190. doi: 10.2307/1131057

Cuff, B. M., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review*, *8*(2), 144–153. doi: 10.1177/1754073914558466

Davis, M. H. (1996). *Empathy: A social-psychological approach*. Boulder, CO: Westview Press.

Davis, M. H., Conklin, L., Smith, A., & Luce, C. (1996). Effect of perspective taking on cognition representation of persons: A merging of self and other. *Journal of Personal and Social Psychology, 70*(4), 713–726. doi: 10.1037/0022-3514.70.4.713

De Vignemont, F., & Singer, T. (2006). The empathic brain: How, when and why? *Trends in Cognitive Sciences, 10*(10), 435–441. doi: 10.1016/j.tics.2006.08.008

Decety, J. (2007). A social cognitive neuroscience model of human empathy. In E. Harmon-Jones & P. Winkielman (Eds.), *Social neuroscience: Integrating biological and psychological explanations of social behavior* (pp. 246–270), New York: Guilford Press

Decety, J., & Jackson, P. L., (2004). Functional architecture of human empathy. *Behavioural and Cognitive Neuroscience Reviews, 3*(2), 71–100. doi: 10.1177/1534582304267187

DeSmet, A., Veldeman, C., Poels, K., Bastiaensens, S., Van Cleemput, K., Vandebosch, H., & De Bourdeaudhuij, I. (2014). Determinants of self-reported bystander behavior in cyberbullying incidents amongst adolescents. *Cyberpsychology, Behavior, and Social Networking*, *17*(4), 207–215. doi: 10.1089/cyber.2013.0027

Eisenberg, N. (2000). Emotion, regulation and moral development. *Annual Review Psychology, 51*, 665–697. doi: 10.1146/annurev.psych.51.1.665

Eisenberg, N., & Strayer, J. (1987). Critical issues in the study of empathy. In N. Eisenberg & J. Strayer (Eds.), *Cambridge studies in social and emotional development. Empathy and its development* (pp. 3–13). New York, NY, US: Cambridge University Press.

Freis, S. D., & Gurung, R. A. (2012). A Facebook analysis of helping behavior in online bullying. *Psychology of Popular Media Culture*, 1–9. doi: 10.1037/a0030239

Gini, G., Albiero, P., Benelli, B., & Altoe, G. (2008). Determinants of adolescents' active defending and passive bystanding behavior in bullying. *Journal of Adolescence*, *31*(1), 93–105. doi: 10.1016/j.adolescence.2007.05.002

Hoffman, M. L. (2000). *Empathy and moral development. Implication for caring and justice.* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511805851

Hoffman, M. L. (2001). *Empathy and moral development: Implications for caring and justice*. Cambridge University Press.

Hume, D. (1739/1968). *A treatise of human nature.* In H. A. Selby-Bigge (Ed.). Oxford: Clarendon Press.

Jolliffe, D., & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and Violent Behavior, 9*(5), 441–476. doi:10.1016/j.avb.2003.03.001

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kiesler, S., Siegel, J., McGuire, T. W. (1984). Social psychological aspects of computer mediated communications. *American Psychology, 39,* 1123–1134. doi: 10.1037/0003-066X.39.10.1123

Kiesler, S., Siegel, J., & McGuire, T. W. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist, 39*(10), 1123–1134. doi:10.1037/0003-066X.39.10.1123

Macaulay, J. R. (1970). A'shill for charity. In L. Berkowitz & J. R. Macaulay, *Altruism and helping behavior* (pp.36–48) New York: Academic Press.

Macháčková, H., Dedková, L., Ševčíková, A., & Černá, A. (2013). Bystanders' support of cyberbullied schoolmates. *Journal of Community & Applied Social Psychology, 23*(1), 25–36. doi: 10.1002/casp.2135

Menesini, E., Codecasa, E., & Benelli, B. (2003). Enhancing children's responsibility to take action against bullying: Evaluation of a befriending intervention in Italian middle schools. *Aggressive Behavior*, *29*, 10–14. doi: 10.1002/ab.80012

Palladino, B. E., Nocentini, A., & Menesini, E. (2012). Online and offline peer led models against bullying and cyberbullying. *Psicthema, 24*(4), 634–639.

Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research*, *96*(1), 3–14. doi: 10.1080/00220670209598786

Preston, S. D., & De Waal, F. (2002). Empathy: Its ultimate and proximate bases. *Behavioral & Brain Sciences, 25*, 1–72.

Pyżalski, J. (2013). Beyond peer cyberbullying – involvement of Polish adolescents in different kinds of electronic aggression, *Studia Edukacyjne, 28*, 147–168. ISBN 978-83-232-2731-1.

Renati, R., Berrone, C., & Zanetti, M. A. (2012). Morally disengaged and unempathic: Do cyberbullies fit these definitions? An exploratory study. *Cyberpsychology, Behavior, and Social Networking*, *15*(8), 391–398. doi: 10.1089/cyber.2012.0046

Richardson, D. R., Green, L. R., & Lago, T. (1998). The relationship between perspective-taking and nonaggressive responding in the face of an attack. *Journal of Personality, 66*(2), 235–256. doi: 10.1111/1467-6494.00011

Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior, 15,*112–120. doi: 10.1016/j.avb.2009.08.007

Sandstrom, M., Makover, H., & Bartini, M. (2012). Social context of bullying: Do misperceptions of group norms influence children's responses to witnessed episodes? *Social Influence, 8*, 196–215. doi: 10.1080/15534510.2011.651302.

Schwartz, S. H. (1977). Normative influence on altruism. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 10, pp. 221–279). New York: Academic Press. doi: 10.1016/s0065-2601(08)60358-5.

Schwartz, S. H., & Howard, J. A. (1981). A normative decision-making model of altruism. In J. P. Rushton & R. M. Sorrentino (Eds.), *Altruism and Helping Behavior: Social, Personality, and Developmental Perspectives* (pp. 189–211). Lawrence Erlbaum, Hillsdale.

Smith, A. (1759/1976). *The theory of moral sentiments.* Oxford: Clarendon Press.

Smith, P. K. (2011). Bullying in schools: Thirty years of research. In C. P. Monks & I. Coyne (Eds.), *Bullying in different contexts* (pp. 36–60). Cambridge: Cambridge University press. doi: 10.1017/CBO9780511921018.003

Spencer, H. (1870). 1872. *The principles of psychology, 2*. London: Williams.

Sproull, L., Conley, C., & Moon, J. M. (2005). The kindness of strangers: Prosocial behavior on the Internet. In Y. Amichai-Hamburger (Ed.), *The Social Net: Human behavior in cyberspace* (pp. 139–161). Oxford Press University. doi: 10.1093/acprof:oso/9780199639540.003.0008

Sticca, F., Ruggieri, S., Alsaker, F., & Perren, S. (2013). Longitudinal risk factors for cyberbulyling in adolescence. *Journal of Community & Applied Social Psychology, 23*(1), 52–67. doi: 10.1002/casp.2136

Thornberg, R., Tenenbaum, L., Varjas, J., Meyers, K., Jungert, T., & Vanegas, G. (2012). Bystander motivation in bullying incidents: To intervene or not to intervene. *Western Journal of Emergency Medicine*, 13, 247–252. doi: 10.5811/westjem.2012.3.11792

Titchener, E. B. (1909). *A text-book of psychology, Part 1*. New York, NY, US: MacMillan Co.

Tokunaga, R. S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in Human Behavior*, *26*, 277–287. doi: 10.1016/j.chb.2009.11.014

Twemlow, S. W., Fonagy, P., Sacco, F. C., Gies, M. L., & Hess, D. (2001). Improving the social and intellectual climate in elementary schools by addressing bully-victim-bystander power struggles. In J. Cohen (Ed.), *Caring classrooms/intelligent schools: The social emotional education of young children* (pp. 162-182). Teachers' College Press: New York.

Völlink, T., Bolman, C. A. W., Dehue, F., & Jacobs, N. (2013). Coping with cyberbullying: Differences between victims, bully-victims and children not involved in bullying. *Journal of Community & Applied Social Psychology, 23*, 7–24. doi: 10.1002/casp.2142

Williford, A., Elledge, L. Ch., Boulton, A. J., DePaolis, K. J., Little, T. D., & Salmivalli, Ch. (2013) Effects of the KiVa Antibullying Program on cyberbullying and cybervictimization frequency among Finnish youth, *Journal of Clinical Child & Adolescent Psychology*, *42*(6), 820–833. doi: 10.1080/15374416.2013.787623

Wölfer, R., Schultze-Krumbholz, A., Zagorscak, P., Jäkel, A., Göbel, K., & Scheithauer, H. (2014). Prevention 2.0: Targeting cyberbullying @ school. *Prevention Science*, *15*(6), 879-887.

Ybarra, M. L., & Mitchell, K. J. (2004). Online aggressors/targets, aggressors, and targets: A comparison of associated Youth characteristics. *Journal of Child Psychology and Psychiatry, 45*(7), 1308–1316. doi: 10.1111/j.1469-7610.2004.00328.x