

## Social stereotype activation and memory performance: a replication study

ALEXANDRA EMILIA GÎRBĂ<sup>1\*</sup>, ANDREI CORNELIU HOLMAN<sup>1</sup>

Initial receipt: 11.08.2018 Final revision received: 12.12.2018 Accepted: 15.12.2018

**Abstract:** The present study is an exact replication of Dijksterhuis, Aarts, Bargh and van Knippenberg (2000). On the relation between associative strength and automatic behavior. The present replication includes, additionally, a phase of elaboration and verification of the prime words on Romanian population. The pre-test of the prime words was necessary because we believe that concepts like *bingo* or *Florida*, used in the original study, are not associated with the elderly in the Romanian culture. The selected words from this phase were used as prime words in the two experiments. The first experiment examined the relation between past contact and stereotype activation effects on memory performance. In the second experiment we examined the relation between past contact, stereotype associative strength and stereotype activation effects on memory performance. The present replication supports only partially the results from the original study. The first study does not confirm the influence of semantic activation by subliminal methods of the stereotype of “elderly” on the number of recalled words. Contrarily, the influence of the subliminal activation of this stereotype on the number of recalled words emerged as significant in the second study. It also failed to replicate the mediation effect of the strength of association between the concept of elderly and forgetfulness in the relationship between the activation of the stereotype “elderly” and the number of words recalled.

**Keywords:** replication study, stereotype activation, memory performance, elderly, priming

### Introduction

In recent years, interest in replicating social studies has increased greatly, especially since more and more replication studies failed to find the original effect. Cesario (2014, p. 40) argues that “replication can accomplish a number of important goals, such as narrowing effect size estimates and providing information about whether an earlier published effect should be considered to be Type I error”. Even so, if we take into account only the studies that were

---

<sup>1</sup> Al. I. Cuza University, Iași, Romania, Faculty of Psychology and Educational Sciences.

\* E-mail of corresponding author: girbaemilia@yahoo.com

published because they had found significant effects, the chances of a false positive are very low, leading to a publication bias (Sterling, 1959). An example is the study of stereotype threat and women's performance on math tests (Spencer, Steele and Quinn, 1999), which was replicated by Flore and Wicherts (2014) who found that, after correcting for publication bias, there was no effect of stereotype threat on women's math performance.

Another example is the theory of ego depletion (Baumeister, Bratslavsky, Muraven, & Tice, 1998), who has been successfully replicated in independent studies over the years. Moreover, the phenomenon was attested even by a meta-analysis conducted by Hagger, Wood, Stiff and Chatzisarantis (2010), which revealed a medium-to-large effect size, but without taking into account the replication bias. The meta-analysis conducted in 2015 by Carter, Kofler, Forster and McCullough, in which they included results from unpublished experiments, revealed no significant effect of ego depletion on adults.

The same patterns can be found in the replication of studies on automaticity in attitudes and social cognition, more specifically in the elderly priming effects. Several replication studies, for example, failed to reveal the effects of elderly priming on walking speed found by Bargh, Chen and Burrows (1996). In the first experiment, Doyen, Klein, Pichon and Cleeremans (2012) found no effect of the priming on walking speed. Even so, they decided to investigate further the factors that could help to uncover the link between priming and walking speed. To this aim, conducted a second experiment in which they manipulated the experimenters' expectations about primed participants' behavior. This time the replication was a success, Doyen et al. (2012) concluding that the priming alone was not sufficient to promote a priming effect, but neither the experimenter expectation was sufficient, and that the two have to be in line in order to influence walking speed. Two experiments conducted by Hull, Slone, Meteyer and Matthews (2002) replicated the effect of elderly primes on walking speed only when they took into account participants' level of self-consciousness, as only high but not low self-conscious individuals walked more slowly following exposure to elderly primes. Another moderator of the elderly priming effects was indicated by the Cesario, Plaks and Higgins (2006) study, which found that the effect of elderly primes on walking speed appeared mainly for individuals who like, versus dislike, the elderly.

Such results show the volatility of the automatic activation effects, with failed replications becoming an opportunity to reveal more subtle explanations for the initial phenomenon. Our study focuses on the research conducted by Dijksterhuis, Aarts, Bargh and Knippenberg (2000) on the effects of elderly primes on memory performance, which has not been submitted to any published test of replication so far. The present paper presents an exact replication of Dijksterhuis et al. (2000) experiments on the effects of the activation of the stereotype towards elder people on memory performance. These relations were

tested in two experiments. The first study tested the hypothesis that activating the category of the elderly will lead to impaired memory performance in participants with frequent contact with the elderly, but not in those with little contact. The first task investigated the amount of contact participants had with the elderly, and then were presented with a word-recognition task. Each word was preceded by subliminal primes (either related or unrelated to the elderly). After the word-recognition task, participants were presented with a surprise free recall task of the words used in the word-recognition task.

In the second Dijksterhuis et al. (2000) experiment, the associative strength between the category elderly and the attribute “forgetfulness” was also examined, as well as the mediating role of this associative strength in the relation of elderly priming and impaired memory performance among participants with a lot of contact with the elderly. The first task investigated the amount of contact participants had with the elderly as in the first experiment. Second, they were presented with a word-recognition task and after this first word-recognition task participants received a 14 min filler task. The second recognition task consisted in completing twice the word-recognition task that was used in the first experiment.

## Method

### Pre-test of the prime words

The present replication includes, additionally, a phase of elaboration and verification of the prime words on Romanian population. The pre-test of the prime words was necessary because we believe that concepts like *bingo* or *Florida*, used in Dijksterhuis et al. (2000) are not associated with the elderly in the Romanian culture, and thus, the cultural adaptation of the set of elderly related primes was imperiously needed (the original words were: *worried, Florida, old, lonely, grey, selfishly, careful, sentimental, wise, stubborn, courteous, bingo, withdraw, forgetful retired, wrinkle, rigid, traditional bitter, obedient, conservative, knits, dependent, ancient, helpless, gullible, cautious, and alone*).

The generation of the elderly prime words was effectuated with the help of forty-five (five females) undergraduate students at the Faculty of Psychology and Education Sciences, “Alexandru Ioan Cuza” University, aged 20-38 ( $M = 22.82$ ,  $SD = 4.19$ ) against partial course credits. The subjects had to elaborate as many words possible associated with the concept *elder*. A total of one hundred-one words were elaborated, from which we selected thirty words with frequencies ranged from 4 to 22.

The associative link of these notions with the concept *elder* was checked for on a sample fifty undergraduate students, aged 20-46 ( $M = 22.12$ ,  $SD = 4.15$ ). Each word was evaluated using a 10-points Likert scale concerning his match with the *elder* concept. For each prime word we calculated the mean, and

the fifteen words with the highest mean were selected. The word with the highest mean was *grey* ( $M = 9.42$ ,  $SD = 0.88$ ), and the word with the smallest mean in this set was *sentimental* ( $M = 6.76$ ,  $SD = 2.31$ ).

The selected words are: *grey*, *elderly*, *grandpa*, *pensioner*, *wrinkles*, *stories*, *memories*, *wise*, *experienced*, *traditional*, *carefully*, *glasses*, *slow*, *cane*, *sentimental*.

## **Experiment 1**

### **Method**

***Participants and Design.*** Forty undergraduates students at the Faculty of Psychology and Education Sciences, “Alexandru Ioan Cuza” University, aged 19-41 ( $M = 22.75$ ,  $SD = 4.83$ ) participated for partial course credits. All subjects had normal or adjusted to normal vision and were right handed.

***Procedure and stimulus materials.*** The procedure was the same as in Dijsterhuis et al. (2000) experiments. Participants were seated in front of a computer and told that they will participate in two different experiments: one that investigated the social life of undergraduate students and one that investigated word recognition. All the instructions and experimental tasks were on the computer. Before the tasks started, the computer randomly assigned each participant in one of the two experimental conditions: the control group or the elderly prime condition. The first task investigated the amount of contact participants had with the elderly by answering on a 9-point scale ranging from *very little time* (1) to *very much time* (9) to the question “How much time do you generally spend among the elderly?”. Seven other filler questions about the time spent among different social groups were addressed in order to mask the aim of the study. The question about the time spent among the elderly was always in the fifth position.

Next, participants received the instructions for the word-recognition task, which consisted in 30 trials.

A trial began with the presentation of a subliminal prime words 17 ms. In the elderly prime condition, the prime words were the 15 words related to elderly elaborated previously (e.g., *gray*); in the control condition, the prime words were not related to the elderly (e.g., *word*). Then the prime words were masked by a row of X's for 225 ms and the word-recognition task started. Their task was to indicate if a target word was an existing word or not by pressing the corresponding buttons (*L* - “yes” or *A* - “no”) as fast as possible. The target words were randomly selected from a set of 15 existing (e.g., *plane*) words and 15 nonsense words (e.g., *acknur*). After the participants responded the trial ended, and, after a 1500 ms inter-trial pause, a new trial began. The trials were randomized for each participant.

After the word-recognition task, participants were presented with a surprise free recall task, in which they were asked to recall as many of the

existing words from the word-recognition task as possible in 3 min. After completion of the surprise recall task, participants were thanked and debriefed.

### **Results and Discussion**

Participants were divided into two groups based on a median split on the answers to the amount of time spend among elderly question: participants with little contact ( $N = 18$ ) and participants with a lot of contact with the elderly ( $N = 18$ ).

We used a 2 (contact with elderly: a lot vs. a little) X 2 (prime: elderly prime vs. control prime) mixed ANOVA, using the number of words recalled correctly as dependent variable. Results shows there is no significant effect of the independent variables on the dependent variable (all  $p$ 's > .05). Supplementary, we used ANCOVA to analyze the effect of the subliminal activation on the number of words recalled correctly, by controlling the continuous variable contact with elderly, but there was no significant effect ( $p > .05$ ).

### **Experiment 2**

Even if the first experiment replication failed to confirm the results of the original study we continued with the replication of the second study of Dijksterhuis et al. (2000), due to two differences in procedure from the first experiment: first of all, the second study took into account the associative strength of the attribute “forgetfulness” with the concept “elderly”. Secondly, the lexical decision task was applied twice in the second study; hence the prime words and the target words were repeated, facilitating the recollection of the targets words in the surprise recall task at the end of the experiment.

### **Method**

**Participants and design.** Sixty-two undergraduate students (fifty-three females), 1st year of study at the Faculty of Psychology and Education Sciences, “Alexandru Ioan Cuza” University, aged 22-47 ( $M = 27.25$ ,  $SD = 7.06$ ) participated against partial course credits. All subjects had normal or adjusted to normal vision and were right handed. None of the participants in this study participated in the first experiment.

### **Procedure and stimulus materials.**

The procedure was the same as in Dijksterhuis et al. (2000) experiments. The overall procedure included the following steps: measurement of previous contact participants had with elderly people, the measurement of the associative strength of “forgetfulness” with “elderly” through a lexical decision task, a filler task, and the measurement of the effect of elderly priming on the reminding ability through a lexical decision task.

Participants were seated in front of a computer and told that they will participate in several different experiments. First, the time spend among elderly people was measured as in Experiment 1. After this, participants were instructed regarding the first lexical decision task that measures the associative strength between forgetfulness and elderly.

A trial began with the presentation of a subliminal prime words for 17 ms. In the experimental condition, the prime word was “old”, and in the control condition, the prime words were not related to “old” (e.g., mug). Then the prime words were masked by a row of X’s for 225 ms and the word-recognition task started. Their task was to indicate if the target word was an existing word or not by pressing the corresponding buttons (“yes” or “no”) as fast as possible. The target words consisted of 6 existing words and 6 nonsense words (e.g., ajmfui). Of the six existing words, three were related to the concept of forgetfulness [forgetful (in Romanian *uituc*), forgetting (*uitare*) and memory loss (*omitere*)], and three were not related to forgetfulness (sea, eagle, bicycle). After the participants responded the trial ended, and, after 1500 ms inter-trial pause, a new trial began. The trials were randomized for each participant.

Next, participants completed a translation of the Ambivalent Sexism Inventory (Glick and Fiske, 1996). The objective of this task, that lasted approximately 14 min to complete, was to wipe out the effects of the semantic priming induced in the previous lexical decision task. After completion of the questionnaire, the procedure was almost the same as in Experiment 1. Participants were administered the same word-recognition task used in Experiment 1, but this time the task was repeated twice, in the same form as in Experiment 1. At the end of it the free recall task was administered, after which participants were thanked and debriefed.

## Results

Participants were divided into two groups based on a median split on the answers to the amount of time spend among elderly question: participants with little contact ( $N = 17$ ) and participants a lot of contact with the elderly ( $N = 29$ ).

We used a 2 (contact with elderly: a lot vs. a little) X 2 (prime: elderly prime vs. control prime) mixed ANOVA, using the number of words recalled correctly as dependent variable. Results shows there is only a significant effect of the priming condition on the number of words recalled correctly [ $F(1,42) = 5.26$ ;  $p = .027$ ;  $\eta_p^2 = 0.11$ ], with participants in the elderly prime condition recalling fewer words than the participants in the no-prime condition (means and standard deviations are presented in *Table 1*).

Supplementary, we used ANCOVA to analyze the effect of the subliminal activation on the number of words recalled correctly while controlling for the effect of participants’ contact with elderly. Results show that even when controlling for the effect of this variable the effect of the priming condition on

the number of words recalled correctly is significant [ $F(1,59) = 4.58; p = .03; \eta_p^2 = .072$ ]. Participants in the elderly prime condition recalled fewer words than those in the no-prime condition (see *Table 1*).

**Table 1.** Means and SDs of number of recalled words in the two priming conditions, with and without controlling (adjusting) for participants' contact with the elderly.

<i>Prime</i>	Unadjusted			Adjusted		
	<i>M</i>	<i>SE</i>	<i>N</i>	<i>M</i>	<i>SE</i>	<i>N</i>
No-prime	6.17	0.50	25	5.94	0.36	35
Elderly prime	4.58	0.47	21	4.73	0.41	27

To assess associative strength, we subtracted the mean reaction time of the three trials related to memory impairment from the mean of the three control trials. Hence, the higher the resulting score, the stronger the association between the elderly concept and memory impairment. We subjected the scores to a 2 (prime: elderly prime versus no prime) X 2 (contact with the elderly: a lot versus a little) mixed ANOVA. We found no significant effect of the independent variables over the dependent variable (all  $p$ 's > .05). Thus, the two groups differentiated by their amount of contact with the elderly did not vary significantly in their associative strength between elderly and memory impairment. Moreover, the Pearson correlation between associative strength and number of recalled words was not significant effect ( $r = -0.07, p = 0.58$ ).

## Discussion

Major concerns have been raised in the past years about the replicability and even plausibility of behavioral priming effects, such as those reported in Dijksterhuis et al. (2000) (e.g. Spellman, 2013). Consequently, replication studies are necessary in this line of research in order to test the robustness of the initial findings. This would ideally require using samples from the same population in order to qualify as exact replications, as behavioral priming effects have been found to be highly sensitive to variations in methodology (Cesario, 2014). Nevertheless, replications in different cultures than that of the original investigation could also be informative, especially when focusing on the behavioral priming effects of culturally – dependent primes such as social stereotypes.

The present replications of the study by Dijksterhuis et al. (2000) only partially support the initial findings of the authors. The first study does not confirm the influence of semantic activation by subliminal methods of the stereotype of “elderly” on the number of recalled words. Contrarily, the influence of the subliminal activation of this stereotype on the number of

recalled words emerged as significant in the second study. The difference between the first study and the second was that participants' exposure to the subliminal and target stimuli (the words recalled at the end of the study) was doubled. The authors' motivation was that in the initial study there were several participants who could not even remember even a word they met during the experiment, indicating that the single exposure was not enough to generate a substantial retention effect. But the same argument can also apply to the subliminal exposure, i.e., the lack of influence of the activation of the "elderly" in the first study might be due to the insufficient exposure to the stimulus. In fact, Dijksterhuis and van Knippenberg (1998) showed that there is a link between the magnitude of activation and the influence on behavior. In this research, participants who were exposed to the word "teacher" for nine minutes had significantly higher scores in a comprehensive general culture test compared to the subjects who were exposed to the same word for only two minutes.

The present study failed to replicate the mediation effect of the strength of association between the concept of elderly and forgetfulness in the relationship between the activation of the stereotype "elderly" and the number of words recalled. This mediation effect was put forth by the authors of the initial study as the explanation of the link between the elderly activation and memory impairment. Specifically, in Dijksterhuis et al.' (2000) view, this stereotype would activate the representation of the elders as having memory deficits, which would subsequently influence participants' own ability to remember the concepts they are exposed to during the experiment. The fact that this mediation effect did not emerge as significant in our results may be due to the activation of a different representation. It is possible that the participants in the present study do not see the elders as having trouble accessing information stored in their memory, but instead as having a reduced attentional capacity, which would lead to a lower ability to encode information. In fact, this alternative hypothesis would also explain the fact that when the number of exposures to the stimuli was doubled (i.e. in the second experiment), the activation of the stereotype "elderly" significantly affected the number of recalled words, in line with the improvement effect that multiple exposure generally has on encoding.

Also, another difference between our replication results and those in the initial study is that Dijksterhuis et al. (2000) found an interaction effect between activating the stereotype "elderly" and the amount of time spent with the elderly, while this interaction effect did not emerge as significant in the replication study. The influence of the subliminal activation of the stereotype "elderly" was significant in our second experiment regardless of the time participants had spent with the elderly. According to Dijksterhuis et al. (2000), an increase in interaction time leads to an increase in the observation of behaviors and features specific to the target social category, which leads to more detailed mental representations becomes. In this specific case, it is possible that subjects do not

need a long interaction time with the elderly in order to form detailed representations, for several reasons. First, the elderly do not represent a low-frequency population, so that its specific features are frequently encountered and thus sufficiently relevant to our young participants. Another possibility may be that there are often indirect interactions with the elderly, through behaviors described by a third person who has frequent direct interaction or through contact with the media and cultural representations of the typical behaviors and traits of this social category.

In conclusion, our results partially replicate the findings reported in Dijksterhuis et al. (2000), indicating a detrimental effect of activating the stereotype towards the elderly on memory when the material to be retrieved is exposed twice, as opposed to once, irrespective of the amount of personal contact with the elderly. These differences between our pattern of results and those originally reported might stem from the cultural specificities of the elderly stereotype in the two countries in which the studies were conducted (the Netherlands vs. Romania). Specifically, the Romanian content of this stereotype might focus on impairments in encoding rather than retrieval in the elderly, and this association might be generalized rather than formed through direct personal contact with representatives of this social category. Further studies should test these explanatory hypotheses and consequently contribute towards answering the general question of whether behavioral priming effects are culturally variable.

## References

- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of personality and social psychology*, *71*(2), 230.
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource?. *Journal of personality and social psychology*, *74*(5), 1252.
- Carter, E. C., Kofler, L. M., Forster, D. E., & McCullough, M. E. (2015). A series of meta-analytic tests of the depletion effect: self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General*, *144*(4), 796.
- Cesario, J. (2014). Priming, replication, and the hardest science. *Perspectives on Psychological Science*, *9*(1), 40-48.
- Cesario, J., Plaks, J. E., & Higgins, E. T. (2006). Automatic social behavior as motivated preparation to interact. *Journal of personality and social psychology*, *90*(6), 893.
- Dijksterhuis, A., & Van Knippenberg, A. (1998). The relation between perception and behavior, or how to win a game of trivial pursuit. *Journal of personality and social psychology*, *74*(4), 865.
- Dijksterhuis, A., Aarts, H., Bargh, J. A., & Van Knippenberg, A. (2000). On the relation between associative strength and automatic behavior. *Journal of Experimental Social Psychology*, *36*(5), 531-544.

- Doyen, S., Klein, O., Pichon, C. L., & Cleeremans, A. (2012). Behavioral priming: it's all in the mind, but whose mind?. *PloS one*, 7(1), e29081.
- Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of school psychology*, 53(1), 25-44.
- Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. (2010). Ego depletion and the strength model of self-control: a meta-analysis. *Psychological bulletin*, 136(4), 495.
- Hull, J. G., Slone, L. B., Meteyer, K. B., & Matthews, A. R. (2002). The nonconsciousness of self-consciousness. *Journal of personality and social psychology*, 83(2), 406.
- Spellman, B. A. (2013). Introduction to the special section on advancing science. *Perspectives on Psychological Science*, 8, 412–413.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of experimental social psychology*, 35(1), 4-28.
- Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance-or vice versa. *Journal of the American statistical association*, 54(285), 30-34.