

## **Study 2**

Blinded<sup>1</sup>, Blinded<sup>2</sup>, Blinded<sup>1</sup>, and & Blinded<sup>1</sup>

<sup>1</sup> Blinded

<sup>2</sup> Blinded

## **Author Note**

Correspondence concerning this article should be addressed to Blinded, Blinded.  
E-mail: Blinded

## **Abstract**

Six studies etc.

*Keywords:* keywords

Word count: TBC

## Study 2

### Study 2 - Online Replication 1

Study 1 demonstrated interesting variability in responses to the critical slide depending on cognitive load. The aim of Study 2 was to assess the replicability of the results of Study 1, using an online sample. In Study 1, the experimenter was in the room with the participants. This made it more difficult for participants to cheat on the memory task. This is not possible with an online sample. An alternative cognitive load manipulation was taken from De Neys and Schaeken (De Neys & Schaeken, 2007), whereby a dot pattern is briefly presented to participants, and participants are required to reproduce the dot pattern at a later stage.

#### Study 2: Method

##### *Study 2: Participants and design*

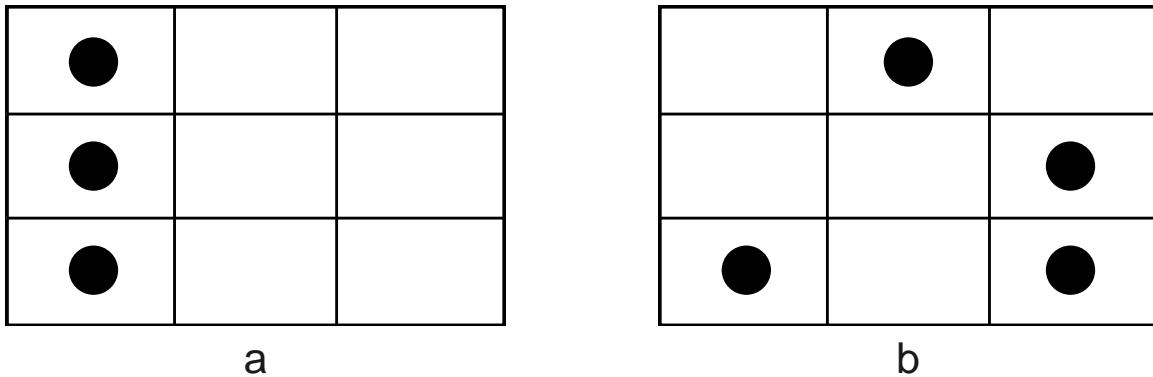
Study 2 was a between subjects design. The dependent variable was response to the critical slide. The independent variable was cognitive load with two levels: high and low. Need for Cognition (Cacioppo & Petty, 1982; Petty, Cacioppo, & Kao, 1984) was included as a potential correlate and moderator variable.

A total sample of 100 participants (56 female, 44 male;  $M_{\text{age}} = 38.38$ ,  $\text{min} = 19$ ,  $\text{max} = 72$ ,  $SD = 12.41$ ) took part. Participants in this sample were recruited using Amazon's MTurk (Amazon Web Services Inc., 2016). Participants were paid \$0.50 for their participation. Participants were recruited from English speaking countries or from countries where residents generally have a high level of English (e.g., The Netherlands, Denmark, Sweden).

### *Study 2: Procedure and materials*

Data were collected using an online questionnaire. Materials were largely the same as in Study 1, with a change to the cognitive load manipulation. Cognitive load was manipulated using a dot-pattern memory task (De Neys & Schaeken, 2007).

Participants were presented with a 3 x 3 grid containing a dot pattern. This image disappeared after one second. Participants then answered a question relating to the moral judgement task. Following this, participants were asked to reproduce the dot-pattern. All participants took part in the memory task, and cognitive load was manipulated by varying the complexity of the patterns presented (De Neys & Schaeken, 2007). The control group were presented with simple patterns, containing three dots in a line, while the experimental group were presented with more complex dot patterns containing 4 dots, see Figure 1.



**Figure 1**

*Sample dot patterns - more simple for the control group (a) and higher complexity for the experimental condition (b)*

Study 2 proceeded in much the same way as Study 1. There were four target questions during which participants were engaged in the memory task. A different pattern was presented before each of the following: the initial judgement, the initial opportunity to provide reasons, the critical slide, and the revised judgement. After each of these questions participants were required to reproduce the pattern. As in Study 1, dumbfounding was

measured using the critical slide.

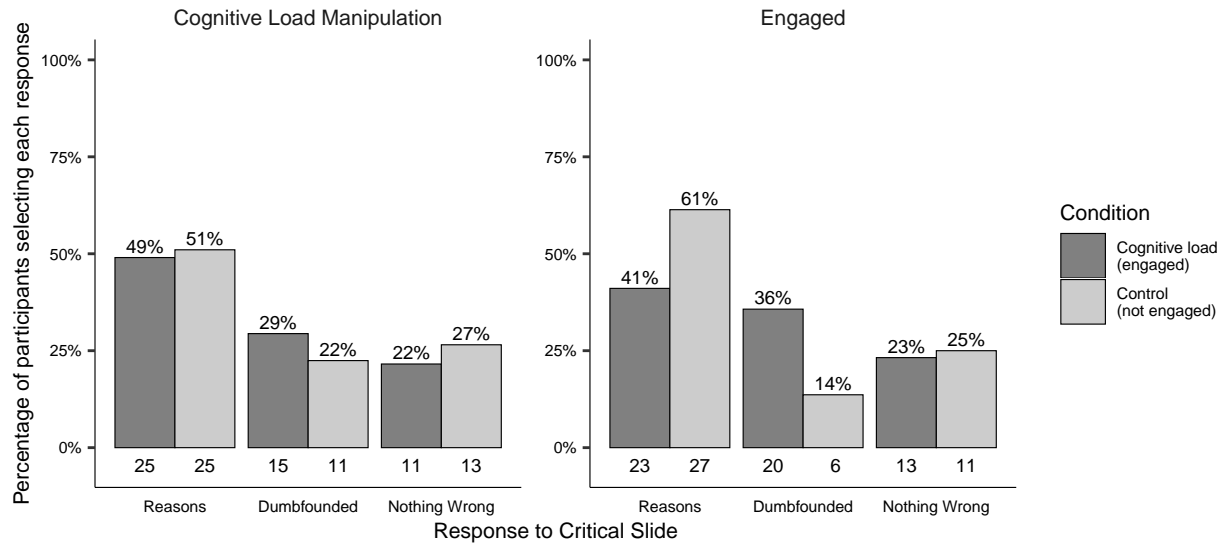
## Study 2: Results

Seventy seven participants (77%) rated the behavior of Julie and Mark as wrong initially, and seventy participants (70%) rated the behavior as wrong at the end of the task. Initial ratings ( $M = 2.13$ ,  $SD = 1.54$ ) were significantly more severe than revised ratings ( $M = 2.35$ ,  $SD = 1.65$ ),  $t(99) = -2.85$ ,  $p = .005$ ;  $d = 0.28$ . Inspection of the binned judgments revealed that ten participants changed the valence of their judgments, and all but one of these involved one judgment that was neutral (see Supplementary materials Table XX).

Participants who selected the admission of not having reasons on the critical slide were identified as dumbfounded. Twenty six participants (26%) selected “It’s wrong but I can’t think of a reason.” Fifty participants (50%) selected “It’s wrong and I can provide a valid reason”; and twenty four participants (24%) selected “There is nothing wrong.”

A chi-squared test for independence revealed no association between experimental condition and response to the critical slide,  $\chi^2(2, N = 100) = 0.74$ ,  $p = .690$ ,  $V = 0.09$ , the observed power was 0.11. The responses to the critical slide for the experimental group ( $N = 51$ ) and the control group ( $N = 49$ ) are displayed in Figure 2.

It is possible that the difference in results observed between Study 1 and Study 2 is due to the alternative manipulation of cognitive load employed. In Study 1, the control group did not engage in any task, however, adopting De Neys and Shaeken’s procedure (2007), participants in the control group of Study 2 engaged in a memory task. It is possible that simply engaging in a memory task led to differences in responses, and that level of difficulty (the manipulation that was employed) was irrelevant. Indeed, the responding to the critical slide in the control group in Study 2 is more similar to the responding in the experimental group in 1 than to the control group in Study 1. Furthermore, rates of successful reproduction of the dot patterns in Study 2 were much

**Figure 2**

*Study 2: Responses to critical slide for (left) the experimental group (N = 51) vs the control group (N = 49); and (right) depending on engagement (N = 56) or non-engagement (N = 44) with the memory task*

lower than reported by De Neys and Shaeken (2007). It is likely that engagement with the memory task in Study 2 was not comparable to that observed by De Neys and Shaeken (2007), and this impacted the results. We propose that not engaging appropriately with the memory task undermines its efficacy as a cognitive load manipulation, and hypothesized that the effect of the cognitive load manipulation is related to the degree to which people engage with the manipulation.

To test this claim we developed a measure of engagement with the memory task. The memory task involved correctly placing dots in a 3 x 3 grid. For the scoring of this task, each of the nine places in the grid could be marked/not marked correctly or incorrectly, making 9 the total possible number of correct responses. If a person misplaced one dot in the pattern this would count for 2 incorrect places in the grid: the mark in the incorrect place, and the absence of a mark in the place it should have been. A participant who received a score of 7, could reasonably be taken to have engaged with the task, and

**Table 1**

*Study 2 – Observed counts, expected counts, and standardised residuals for each response to the critical slide depending on cognitive load*

		Engaged	Not Engaged
Observed count	Reasons	23	27
	Dumbfounded	20	6
	Nothing Wrong	13	11
Expected count	Reasons	28	22
	Dumbfounded	14.56	11.44
	Nothing Wrong	13.44	10.56
Standardised residuals	Reasons	-2.01*	2.01*
	Dumbfounded	2.5*	-2.5*
	Nothing Wrong	-0.21	0.21

*Note.* \* = sig. at  $p < .05$ ; \*\* = sig. at  $p < .001$

simply made a slip. As such, this was taken as the cut-off point for identifying engagement. This resulted in 56 participants being identified as engaging with the memory task, and 44 being identified as not engaging with the task.

Using this engagement threshold, we tested if responses to the critical were associated with engagement with the memory task. A chi-squared test for independence revealed an association between engagement in the memory task and response to the critical slide,  $\chi^2(2, N = 100) = 6.68, p = .035, V = 0.26$ , the observed power was 0.11. The responses to the critical slide for the experimental group ( $N = 51$ ) and the control group ( $N = 49$ ) are displayed in Figure 2. The observed counts, expected counts and standardized residuals are displayed in Table 1.

A multinomial logistic regression revealed no significant association between Need for Cognition and response to the critical slide,  $\chi^2(2, N = 100) = 2.19, p = .334$ , the observed power was 0.24 (see supplementary Figure XX for relative probabilities of selecting each response depending on Need for Cognition).

Amazon Web Services Inc. (2016). *Amazon Mechanical Turk*.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116–131. <https://doi.org/10.1037/0022-3514.42.1.116>

De Neys, W., & Schaeken, W. (2007). When people are more logical under cognitive load: Dual task impact on scalar implicature. *Experimental Psychology*, 54(2), 128–133. <https://doi.org/10.1027/1618-3169.54.2.128>

Petty, R. E., Cacioppo, J. T., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48(3), 306–307.