



UNIVERSITY OF LEEDS

This is a repository copy of *Contribution to Alonga et al (2014). Registered replication report: Schooler & Engstler-Schooler (1990).*

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/81917/>

Version: Submitted Version

Article:

Alogna, VK, Attaya, MK, Aucoin, P et al. (87 more authors) (2014) Contribution to Alonga et al (2014). Registered replication report: Schooler & Engstler-Schooler (1990). *Perspectives on Psychological Science*, 9 (5). 556 - 578. ISSN 1745-6916

<https://doi.org/10.1177/1745691614545653>

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>



**Registered Replication Report: Schooler & Engstler-Schooler
(1990)**

Journal:	<i>Perspectives on Psychological Science</i>
Manuscript ID:	Draft
Manuscript Type:	Replication Report
Date Submitted by the Author:	n/a
Complete List of Authors:	Simons, Daniel; University of Illinois at Urbana-Champaign, Department of Psychology
Keywords:	Application: Law, Memory
User Defined Keywords:	

SCHOLARONE™
Manuscripts

Only

Registered Replication Report: Schooler & Engstler-Schooler (1990)

Multi-lab direct replication of: Study 4 (modified) and Study 1 from Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, 22, 36-71.

Proposing Authors: This proposal was initiated by the editors

Protocol vetted by: Jonathan Schooler

Protocol edited by: Daniel J. Simons

Manuscript written by: Daniel J. Simons

Contributing authors (alphabetical order): Alogna, V. K., Attaya, M. K., Aucoin, P., Bahník, S., Birch, S., Birt, A. R., Bornstein, B. H., Bouwmeester, S., Brandimonte, M. A., Brown, C., Buswell, K., Carlson, C., Carlson, M., Chu, S., Cislak, A., Colarusso, M., Colloff, M. F., Dellapaolera, K. S., Delvenne, J.-F., Di Domenico, A., Drummond, A., Echterhoff, G., Edlund, J. E., Eggleston, C., Fairfield, B., Franco, G., Gabbert, F., Gamblin, B. W., Garry, M., Gentry, R., Gilbert, E., Greenberg, D. L., Halberstadt, J., Hall, L., Hancock, P. J. B., Hirsch, D., Holt, G., Jackson, J. C., Jong, J., Kehn, A., Koch, C., Kopietz, R., Körner, U., Kunar, M. A., Lai, C., Langton, S., Leite, F. P., Mammarella, N., Marsh, J. E., McConaughy, K. A., McCoy, S., McIntyre, A., Meissner, C. A., Michael, R. B., Mitchell, A., Mugayar-Baldocchi, M., Musselman, R., Ng, C., Nichols, A. L., Nunez, N. L., Palmer, M. A., Pappagianopoulos, J. E., Petro, M., Poirier, C. R., Portch, E., Rainsford, M., Rancourt, A., Romig, C., Rubínová, E., Sanson, M., Satchell, L., Sauer, J. D., Schweitzer, K., Shaheed, B. W., Skelton, F., Sullivan, G. A., Susa, K. J., Swanner, J. K., Thompson, W. B., Todaro, R., Ulatowska, J., Valentine, T. R., Verkoeijen, P. P. J. L., Vranka, M., Wade, K. A., Was, C. A., Weatherford, D., Wiseman, K., Zaksaitė, T., Zuj, D. V., Zwaan, R. A.

Full citation: Alonga et al (2014). Registered replication report: Schooler & Engstler-Schooler (1990). *Perspectives on Psychological Science*, XX, aaa - bbb.

Citation to individual study contribution: Authorone, A., & Authortwo, M.N. (2014). Contribution to Alonga et al (2014). Registered replication report: Schooler & Engstler-Schooler (1990). *Perspectives on Psychological Science*, XX, aaa - bbb. OSF link.

Abstract

Trying to remember something now typically improves your ability to remember it later. But, after watching a video of a simulated bank robbery, participants who verbally described the robber were 25% worse at identifying the robber in a lineup than were participants who instead listed US states and capitals—the “verbal overshadowing” effect (Schooler & Engstler-Schooler, 1990). More recent studies suggested that this effect might be substantially smaller than first reported. Given uncertainty about the effect size, the influence of this finding in the memory literature, and its practical importance for police procedures, we conducted two collections of pre-registered direct replications (RRR1 and RRR2) that differed only in the order of the description task and a filler task. In RRR1, when the description task immediately followed the robbery, participants who provided a description were 4% less likely to select the robber than were those in the control condition. In RRR2, when the description was delayed by 20 minutes, they were 16% less likely to select the robber. These findings reveal a robust verbal overshadowing effect that is strongly influenced by the relative timing of the tasks. The discussion considers further implications of these replications for our understanding of verbal overshadowing.

Registered Replication Report: Schooler & Engstler-Schooler (1990)

1
2
3
4
5
6
7
8 If you want to remember something better, practice it. This mantra follows from
9
10 decades of memory research: Repeat the names of people you have just met; study
11
12 flashcards for your upcoming language test; summarize the chapter you read. Other
13
14 techniques might be even better, but this type of rehearsal cannot hurt. Or can it?

15
16
17 The results of Schooler and Engstler-Schooler (1990; henceforth S&E-S)
18
19 suggested that in some cases, when the to-be-remembered materials are visual, verbal
20
21 rehearsal hurts rather than helps memory performance. Participants in their study
22
23 witnessed a video of a simulated bank robbery. Half wrote a description of the robber
24
25 and half completed an unrelated writing task. All then tried to pick the robber out of a
26
27 photo lineup. Those who had provided a written description correctly identified that
28
29 robber approximately 25% less often than those who performed the unrelated writing
30
31 task.
32
33
34
35

36 This finding, dubbed “the verbal overshadowing effect,” suggests that verbally
37
38 describing a person impairs later recognition memory for that person. Thus, eyewitness
39
40 recollection may be impaired by asking witnesses to describe what they saw, a result
41
42 with both practical and theoretical importance. The paper has had a substantial impact
43
44 on the field: It has been cited more than 600 times and is a staple of psychology
45
46 textbooks.
47
48
49

50 Yet, the magnitude of the effect remains uncertain. Schooler has noted that the
51
52 measured effect size of the overshadowing effect reported in later studies is smaller
53
54 than that in the original report (Schooler, 2011; see also Lehrer, 2010). A meta-analysis
55
56 of verbal overshadowing studies of lineup recognition performance revealed a
57
58
59
60

1
2
3 significant but smaller (about 12%) effect of verbal description (Meissner & Brigham,
4
5
6 2001). The studies in the meta-analysis used a variety of stimuli, delays, filler tasks, and
7
8 other materials, with substantial heterogeneity in the measured effect sizes across
9
10 studies, with some studies finding no effect or even an effect in the opposite direction.¹
11
12 The studies also might have overestimated the true effect because they had small
13
14 sample sizes: The collection of studies included more statistically significant results than
15
16 would be expected based on their power to find statistical significance, indicating a
17
18 pattern of publication bias in favor of statistically significant results (Francis, 2012).² In
19
20 the presence of publication bias, the true population effect size is difficult to estimate
21
22 from a meta-analysis. Moreover, some of the differences in methods used across the
23
24 studies could have moderated the underlying effect. For example, the meta-analysis
25
26 found less verbal overshadowing with a delay between the verbal description task and
27
28 the lineup identification task. However, the size of the delay varied substantially across
29
30 studies.
31
32
33
34
35

36
37 Verbal overshadowing potentially has broad ramifications, both for our
38
39 understanding of the mechanisms of memory and for police practices. If asking a
40
41 witness to verbally describe the person they saw substantially impairs their ability to
42
43 recognize that person later, then eyewitness identification should be weighted less if the
44
45 witness had provided a description earlier. Given the importance and influence of this
46
47 finding, coupled with uncertainty about the size of the effect and the absence of any
48
49 large-scale direct replications of it, the original study merits a large-scale direct
50
51 replication to better determine the size of the effect. This registered replication report
52
53 was designed to provide an accurate estimate of the verbal overshadowing effect via a
54
55
56
57
58
59
60

1
2
3 collection of pre-registered, independently conducted direct replications of the original
4 study, all using the same materials and a common, vetted protocol.
5
6
7
8
9

10 **Protocol Development to Compare Past and Present Studies**

11
12 The protocol for a direct replication of the original verbal overshadowing study
13 was developed in collaboration with the lead author of the original article, Jonathan
14 Schooler. Once the protocol was completed, *Perspectives on Psychological Science*
15 publicly announced a call for laboratories interested in participating on May 14, 2013.
16
17 Based on the rapid response from a large number of labs, we set a deadline for
18 proposals of June 11, 2013. A total of 31 labs joined the initial replication project
19 (RRR1). All labs pre-registered the details of their plan to implement the protocol, the
20 editors verified those plans before data collection began, and each lab conducted an
21 independent replication. Of those teams, 22 completed a follow-up experiment (RRR2)
22 that reversed the order of the filler task and the description task.
23
24
25
26
27
28
29
30
31
32
33
34
35

36 We conducted RRR2 after discovering an error in the original protocol that went
37 unnoticed throughout the development process. Although we had intended to replicate
38 S&E-S Study 1, the protocol inadvertently reversed the order of the verbal description
39 task and the filler task. In S&E-S Study 1, participants saw the video, did the filler task,
40 then wrote their verbal description and moved to the lineup task. In RRR1, they wrote
41 their description immediately after seeing the video and then did the filler task, thus
42 adding a 20-minute delay before the lineup task. Previous evidence suggests that
43 introducing a delay between the verbal description task and the lineup can reduce the
44 overshadowing effect, meaning that the task order of RRR1 might not provide the
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 strongest possible test of the overshadowing effect (e.g., Finger & Pezdek, 1999;
4
5 Meissner & Brigham, 2001; note, though, that S&E-S Study 4 showed a roughly
6
7 comparable overshadowing effect using an order comparable to RRR1). After a
8
9 participating laboratory noticed the error, the editors consulted with Schooler and we
10
11 collectively decided to conduct RRR2, reversing the task order to match that of S&E-S
12
13 Study 1. That way, we could replicate the original study as intended, providing the
14
15 strongest and clearest test of the verbal overshadowing effect, and we could also
16
17 examine the effect of task order by comparing the two RRR studies. Critically, this
18
19 decision was made before data collection from RRR1 was completed or analyzed,
20
21 making the decision blind to the outcome of RRR1. Moreover, labs were not informed
22
23 about the results from any other labs until data collection from both studies had been
24
25 completed.
26
27
28
29
30

31
32 For the purposes of this report, we treat RRR1 as a fairly direct replication of
33
34 S&E-S Study 4. The studies used the same task ordering, with the filler task coming
35
36 after the verbal description. Note that S&E-S Study 4 included another between-subject
37
38 condition and that the delay was 10 minutes rather than 20 minutes. So, RRR1 is not an
39
40 exact replication of all conditions of S&E-S Study 4. But, the difference in the length of
41
42 the delay is the only substantive change in procedure from a direct replication of the
43
44 critical comparisons for a test of verbal overshadowing on face recognition. A benefit
45
46 arising out of our error is that, by using the same timing in RRR1 and RRR2, we can
47
48 provide one of the first highly powered direct comparisons of the influence of task order
49
50 on the verbal overshadowing effect.
51
52
53
54
55
56
57
58
59
60

1
2
3 Many of the teams consisted of experts on memory and eyewitness accuracy,
4 including some researchers who had previously studied verbal overshadowing. Other
5 labs had experience in conducting other types of cognitive psychology experiments, and
6 still others lacked domain-specific experience but were skilled in experimental methods
7 and were interested in replication efforts more broadly. The participating labs included
8 teams from 10 countries and from a variety of college and university settings. For labs in
9 non-English speaking countries, the associated researchers translated the instructions
10 and other materials and then independently translated them back to English to verify
11 accuracy. In some cases, the audio track on the bank robbery video was dubbed into
12 the native language of the participants. Details of this translation process and any other
13 departures from the standardized procedures are noted by the participating labs in the
14 individual study descriptions (see Table 3 and Appendix). Laboratories were responsible
15 for obtaining any necessary ethics approval from their institutions.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

34 In addition to the lab-based studies, one lab that had participated in RRR1
35 replicated the procedures of both RRR1 and RRR2 in a large-scale online experiment
36 using participants from Amazon Mechanical Turk. Except as noted in the study
37 description below, it followed the same protocols as the lab based studies. Given that it
38 adopted a different procedure, it was not included in the meta-analytic effect size
39 estimates, but it is reported alongside the lab-based results for comparison.
40
41
42
43
44
45
46
47
48
49
50

51 **Insert Figure 1 About Here**

52 (Note: If not exactly here, earlier is better than later)
53
54
55
56
57
58
59
60

Protocol Requirements

Participants

The protocol specified a minimum allowable sample size of 50 participants in each condition, but labs were encouraged to include as many participants as possible. Given that the goal was a direct replication of the original result, the protocol specified that participants be drawn from an undergraduate subject population with all participants between the ages of 18 and 25 years. It further required that participants be able to understand the instructions and have vision adequate to perceive the events in the video and to recognize people. Because the robber depicted in the original video was White, and the verbal overshadowing effect is thought to be weaker with other-race faces than with own-race faces (Fallshore & Schooler, 1995), only White participants were included in the analyses reported here. The sample in each replication study was required to be between 20% and 80% female. Many of the labs collected additional data from participants who did not meet these inclusion criteria, and data from all participants are included in the data files posted on the main project page at OSF (<https://osf.io/ybeur/>).

Testing Location

The protocol required in-person testing. Testing could occur individually or in small groups, provided that participants could not see or hear each other when viewing stimuli or responding, and that they could not communicate with each other during the study. The protocol specified that the study could not be conducted in a classroom setting. (This stipulation was included to maximize the similarity of the testing context across labs.)

Experimenters

Any trained research assistant, postdoctoral researcher, or faculty member could serve as the experimenter if they had experience collecting experimental psychology data and interacting with subjects. No special expertise was required to conduct the study, and the experimenter did not need to be blind to condition assignment (as that would be difficult to achieve and was not the case in the original study).

Materials

Schooler provided a digitized version of the original videotape that was then reformatted as a QuickTime movie file. Schooler also provided a digital version of the original lineup image (an 8-person lineup that included the robber) as well as the text of the instructions given to participants for each task. The original studies used a variety of filler tasks, but Schooler recommended using a crossword puzzle, something he had done in some of his studies. The original crossword puzzle was no longer available, so Schooler selected a comparable crossword puzzle. All of these materials are available from <https://osf.io/ybeur/>.

Data Collection

The study could be conducted by presenting the video using a computer display, television, or projector and by collecting written responses and ratings either on paper or on a computer. Participants were blind to the hypothesis about verbal overshadowing and were unaware of any experimental conditions other than their own. They also were not informed that they were participating in a recognition memory experiment—the study description used for recruiting participants described it more vaguely as a study of perception and memory. Participants were randomly assigned to the experimental and

1
2
3 control conditions with the constraint that approximately equal numbers of participants
4 were assigned to each condition. Labs differed in how they implemented the random
5 assignment, and details are provided in the individual study descriptions (see Table 3
6 below). Note that the original S&E-S studies assigned participants to conditions in small
7 groups and all members of each group were assigned to the same condition. The
8 replication protocol required individual assignment to condition to eliminate this non-
9 independence in randomization.
10
11
12
13
14
15
16
17
18

19 **Procedure**

20
21
22 Participants were told, "This experiment consists of several tasks. First, please
23 pay close attention to the following video." They then viewed a 44-second video
24 depicting a bank robbery. Participants assigned to the Experimental condition were then
25 asked to write a description of the robber:
26
27
28
29
30
31
32
33

34 Please describe the appearance of the bank robber in as much detail as possible.

35 It is important that you attempt to describe all of his different facial features.

36 Please write down everything that you can think of regarding the bank robber's
37 appearance. It is important that you try to describe him for the full 5 minutes.
38
39
40
41
42
43
44
45

46 Participants assigned to the Control condition were asked to "Please name as many
47 countries and their capitals as you can." In the original study, participants were asked to
48 list the states of the United States and their capitals, but for the replication protocol, the
49 control task was changed because participants outside the United States might not be
50 as familiar with states and capitals in the United States.
51
52
53
54
55
56
57
58
59
60

1
2
3 After 3 minutes, each group received a reminder to keep working. Participants in
4
5 the Experimental condition were told, "Please continue describing every detail of the
6
7 bank robber. It is important that you provide as full a description as possible."
8
9
10 Participants in the Control condition were told, "Please continue to list countries and
11
12 their capitals. It is important that you continue this task for the full 5 minutes." This
13
14 reminder could be spoken aloud or presented on the computer display. If the reminder
15
16 was spoken and the testing session included participants from both conditions, the
17
18 reminder was phrased to be condition-blind: "Please keep working. It's important that
19
20 you continue the task for the full 5 minutes."
21
22
23

24 After 5 minutes of writing/typing, participants spent 20 minutes working on a
25
26 printed crossword puzzle. Immediately after this filler task, participants viewed a lineup
27
28 of eight faces and heard/read the following instructions: "Next you will see a lineup with
29
30 eight faces. Please identify the individual in the lineup who you believe was the bank
31
32 robber in the video you watched earlier. If you do not believe the bank robber is present
33
34 please indicate 'not present'." If the lineup was presented on a computer or projector,
35
36 the images were numbered 1-8 to allow a keyboard response and the last sentence of
37
38 the instructions was modified to end "...please indicate 'not present' by pressing '9'.
39
40 Press 'space' to view the image." Finally, participants were asked to "Please indicate
41
42 your confidence in your selection from the lineup on a scale from 1 (guessing) to 7
43
44 (certain)."
45
46
47
48
49

50 **Data Collection Stopping Rules and Exclusions**

51
52 Each lab documented their stopping rules for data collection as part of their OSF
53
54 pre-registration (see Appendix for links), and the editors reviewed these procedures to
55
56
57
58
59
60

1
2
3 verify that they ensured random assignment to conditions and that each lab would be
4 able to meet the minimum required sample size after any exclusions necessitated by
5 the protocol requirements. Labs were permitted to exclude participants for any of the
6 following reasons: participants did not meet the age or race requirements for the study,
7 participants did not follow instructions on the experimental or control task, participants
8 did not complete all tasks, or the experimenter/computer incorrectly administered the
9 task or instructions. Labs were permitted to pre-specify additional exclusions
10 necessitated by their testing situation (e.g., failure to understand the nature of the
11 video). All decisions about whether or not to exclude data were made prior to examining
12 performance on the recognition task and were based on factors unrelated to the
13 outcome measures. All excluded data are included in the data files along with the
14 reason for exclusion.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

31 **Differences between RRR1 and RRR2**

32 All materials and procedures were identical across the two studies except for the
33 following substantive changes (see Figure 1):
34
35

36 (a) In RRR2, the crossword puzzle filler task followed immediately after the video and
37 preceded the verbal description (experimental) or countries/capitals (control) task. The
38 lineup task immediately followed the experimental/control task.
39
40
41
42
43
44

45 (b) The minimum required sample size for the study was reduced from 50 to 30
46 participants in each condition in order to accommodate smaller subject pools available
47 in the spring semester at many universities, thereby permitting participation by more
48 labs.
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 (c) When necessary, labs were permitted to use paid participant pools, funding was
4
5 provided from the Association for Psychological Science (APS) via a grant from the
6
7 Center for Open Science.
8
9

10
11
12 As noted above, S&E-S Study 1 used a 20-minute filler task, but S&E-S Study 4
13
14 used a 10-minute filler task. After we identified the error in the task ordering in RRR1,
15
16 and in consultation with Schooler, we chose to maintain the 20-minute filler task across
17
18 RRR1 and RRR2 in order to make them directly comparable.
19
20

21 22 **Online Version of the Protocol** 23

24
25 In addition to the lab based protocol adopted by all of the replicating teams,
26
27 *Perspectives* solicited and APS funded an online version of the study that was
28
29 conducted by one of the teams that had participated in RRR1 (Michael et al). The
30
31 participants for this study were drawn from Amazon Mechanical Turk, with each
32
33 participant randomly assigned to the task order from RRR1 or RRR2 and to the verbal
34
35 description or countries/capitals task. We chose to have one lab conduct a single large-
36
37 scale online experiment rather than having multiple labs conduct smaller experiments.
38
39 That approach avoids a duplication of effort and the difficulty of ensuring that a
40
41 Mechanical Turk participant did not complete multiple verbal overshadowing
42
43 experiments. We also would not have been able to collect enough independent online
44
45 replications to conduct a meta-analysis of the online-only studies, so we favored a
46
47 single, larger-scale study. The results of the Mechanical Turk study were not included in
48
49 the meta-analysis of the lab-based replications, but they are reported along with the lab
50
51 results in all Tables and Figures.
52
53
54
55
56
57
58
59
60

1
2
3 In most respects, the Mechanical Turk study was identical to the lab-based ones:
4
5 It used the same materials, the same timing and instructions, and the same measures.
6
7
8 Due to the constraints of online testing, though, the following changes were made
9
10 based on consultation between the editors and Schooler: 1) Participants were paid
11
12 USD2.00 for participation. 2) Participants were excluded for reasons beyond those in
13
14 the lab task, including: a failure to list at least 5 countries/capitals in the control
15
16 condition, a self-reported failure to engage appropriately with the filler task, having seen
17
18 the robbery video before, or reporting participation in a study just like this one. 3)
19
20 Participation was limited to participants from the United States. 4) The crossword puzzle
21
22 filler task was replaced with a set of Sudoku puzzles. 5) Participants were not given a
23
24 reminder after 3 minutes to continue writing their description of the robber or listing
25
26 countries/capitals.
27
28
29
30
31
32
33

34 Results

36 Lab Demographics and Results

37
38 Tables 1 and 2 provide demographic data for each lab, including the number of
39
40 participants tested in each condition, the number who did not meet the demographic
41
42 requirements or who were excluded for other reasons, and the number of the included
43
44 participants who made each type of lineup selection (correct ID, mistaken ID, “not
45
46 present”). For comparison, the tables include data from the original S&E-S studies. Note
47
48 that some of the S&E-S numbers were reported in the original journal article and others
49
50 were in Schooler’s dissertation (those that were not reported and are no longer
51
52 available are marked “NA”).
53
54
55
56
57
58
59
60

INSERT TABLES 1 and 2 HERE

Table 1 includes RRR1 info. Table 2 includes RRR2 info

Data Analyses – Original and Present

The S&E-S data analysis consisted of a Chi-Square test comparing the frequency of correct and incorrect identification in the Experimental and Control conditions. A secondary analysis included a Chi-Square comparing the types of errors (selecting the wrong face from the lineup or indicating “not present”) across the conditions. Finally, the original study reported a 2 (Condition) x 2 (Correct vs. incorrect/miss) ANOVA on confidence ratings. For the RRRs, each lab conducted these analyses and they are reported on the lab OSF project pages (URLs available along with each lab’s project summary in the Appendix). Given that we have access to the full data set for each study, for the meta-analysis, we used a more direct measure of the performance difference between conditions (the risk difference: the difference in percentage correct and the difference in percentages of error types). We did not meta-analyze the ANOVAs of confidence ratings.

“Verbal overshadowing” is defined as the difference in accuracy between the control condition (listing countries and capitals) and the verbal description condition (writing a description of the robber). But that difference can be measured in absolute or relative terms. The difference between 10% accuracy and 15% accuracy could be treated as a 5% increase in accuracy (the difference between the percentages) or it

1
2
3 could be treated as a 50% increase in accuracy (15% accuracy is 1.5x as large as
4
5 10%). Note how these measures differ when the baseline accuracy is different: 50%
6
7 and 55% accuracy still differ by 5%, but 55% is only 10% bigger in ratio terms ($55 = 1.1$
8
9 $* 50$). A ratio measure takes that baseline difference into account. When the baseline
10
11 accuracy varies widely across studies or when the same difference in magnitude has
12
13 different meanings (the difference between 50 and 55 has less importance than the
14
15 difference between 5 and 10), ratio measures are more appropriate. But, when
16
17 accuracy levels are roughly comparable across studies and none are extreme, the raw
18
19 difference between the percentages is more straightforward.

20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Given that accuracy levels in these studies were not extreme, for the meta-analyses, we used “risk difference” as our measure of effect size: the percentage accuracy for the Verbal Description condition minus the percentage accuracy for the Control Condition. Negative effect sizes indicate a cost of verbally describing the robber.

Effect Size Measurements

For both RRRs, we provide a forest plot showing the accuracy percentages in each condition for each lab, the effect size measured by each lab (with 95% confidence intervals), and the meta-analytic effect size estimate in a random effects model. The top-most data point in each plot shows the effect from S&E-S and the data point below that shows the effect found in the online Mechanical Turk variant of the study. Neither of those results are included in the meta-analytic effect size estimate at the bottom of each figure; the meta-analysis includes only the pre-registered, lab-based replications of the original study. To permit a visual comparison of effects across the RRR studies, the plot

1
2
3 for RRR1 identifies the subset of labs that completed both studies and separates those
4
5 from the subset that completed only the first study.
6
7
8
9

10 **INSERT FIGURES 2 and 3 HERE**
11
12
13
14

15 For RRR1 (Figure 2), the meta-analysis showed a small effect of verbally
16 describing the robber relative to listing countries and capitals: Whereas the original
17 study showed a -22% difference between the verbal description condition and the
18 control condition, the meta-analytic effect across 31 larger-scale replications was
19 substantially smaller: -4.01% [95% confidence interval: -7.15% to -0.87%]. The original
20 study had a larger absolute effect size than any of the replication studies, but that
21 estimate also was the least precise because of its smaller sample size. All of the
22 replication effect size estimates, including the online Mechanical Turk study, fell
23 between -17.54% and 14.00%. The differences in the estimated effect size among the
24 studies (i.e., heterogeneity) were consistent with what would be expected by chance
25 (Tau = 0, $I^2 = 0\%$, $H^2 = 1.00$, $Q_{30} = 29.302$, $p = 0.502$).³ Taken together, these studies
26 reveal only a small effect of verbal descriptions on lineup accuracy when the task order
27 required participants to provide their verbal description immediately after witnessing the
28 crime video and then view the lineup after a 20 minute delay (see also Finger & Pezdek,
29 1999 and Meissner & Brigham, 2001 for evidence that the verbal overshadowing effect
30 is smaller with a delay between the description and lineup task).
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52 For RRR2 (Figure 3), the meta-analysis revealed a substantially larger effect of
53 verbally describing the robber relative to listing countries and capitals. The original study
54
55
56
57
58
59
60

1
2
3 showed a -25.00% difference between the verbal description condition and the control
4
5 condition, and the meta-analysis of 22 studies showed a difference of -16.31% [95%
6
7 confidence interval: -20.47% to -12.14%]. All 22 studies as well as the online
8
9 Mechanical Turk study showed an effect in the same direction, with effect sizes ranging
10
11 from -28.99% to -10.61%. The differences in the estimated effect size among the
12
13 studies (i.e., heterogeneity) were entirely consistent with what would be expected by
14
15 chance (Tau = 0, $I^2 = 0\%$, $H^2 = 1.00$, $Q_{21} = 15.25$, $p = 0.810$). Taken together, these
16
17 studies reveal a robust and consistent effect of verbal descriptions on lineup accuracy
18
19 when the task order requires participants to wait 20 minutes before providing the verbal
20
21 description and then immediately try to identify the person they saw in a lineup.
22
23
24
25
26

27 When participants did not correctly select the robber from the lineup, they could
28
29 make one of two types of error: Selecting someone else from the lineup (false
30
31 identification) or electing not to select anyone (miss). S&E-S Study 1 reported no
32
33 difference in the proportion of errors that were false identifications between the verbal
34
35 description condition and the control condition.⁴ This breakdown of the errors into two
36
37 categories is no longer available for S&E-S Study 4.
38
39
40
41
42
43
44
45
46
47

48
49
50
51
52
53
54
55
56
57
58
59
60

INSERT FIGURES 4 and 5 HERE

58 In RRR1, the meta-analysis showed reliably higher false identification rates in the
59 control condition than in the verbal description condition (Figure 4). Across the 31 lab
60 replication studies, the meta-analytic effect size was -11.53% [-16.36% to -6.70%],
where negative numbers mean that the false alarm rate was larger in the control

1
2
3 condition. The Mechanical Turk replication showed a difference of -22.97%, consistent
4
5 with the pattern of the lab studies. The heterogeneity across studies was largely
6
7 consistent with what would be expected by chance ($\tau = 0.0462$, $I^2 = 11.41\%$, $H^2 =$
8
9 1.13 , $Q_{30} = 34.72$, $p = 0.253$).

10
11
12 This pattern was similar in RRR2 (Figure 5), with a meta-analytic difference of -
13
14 15.49% [-22.91% to -8.06%]. The Mechanical Turk study showed a similar effect of -
15
16 19.29% . Although the overall pattern and size of the effect was consistent across
17
18 studies, the results from individual labs were more heterogeneous in RRR2 ($\tau =$
19
20 0.1113 , $I^2 = 39.77\%$, $H^2 = 1.66$, $Q_{21} = 34.06$, $p = 0.036$), ranging from a minimum of -
21
22 50.95% to a maximum of 12.47% . Note, though, that the minimum required sample
23
24 size in RRR2 was smaller than in RRR1, meaning that the effect size estimates from
25
26 each lab are less precise.
27
28
29
30
31
32
33
34

35 Conclusions

36
37 The results of this large-scale, multiple-lab direct replication of S&E-S Study 4
38
39 and S&E-S Study 1 shows that verbally describing the robber in a video can impair
40
41 successful selection of that person from a subsequent lineup. The effect was larger
42
43 when the verbal description happened immediately before the lineup selection than
44
45 when it happened immediately after viewing the video. For RRR1, all of the replication
46
47 studies produced a smaller effect size estimate than S&E-S Study 4, but the sample
48
49 size in S&E-S was small enough that its large confidence interval included most of the
50
51 replication studies. For RRR2, the original result from S&E-S Study 1 was close to that
52
53 of the average replication study.
54
55
56
57
58
59
60

1
2
3 Although S&E-S reported no difference across conditions in proportion of errors
4 that were false identifications as opposed to responding “not present”, both replication
5
6 studies found a robust difference, with a higher proportion of false identification errors in
7
8 the control condition than in the verbal description condition. This difference in the types
9
10 of errors across conditions might reflect a difference in the response bias—the
11
12 willingness to select someone from the lineup—induced by the critical manipulation (see
13
14 Clare & Lewandowsky, 2004; see also Chin & Schooler, 2008 for further discussion).
15
16 Alternatively, the pattern of errors might reflect a memory distortion caused by eliciting a
17
18 verbal description. Further studies that include both target present and target absent
19
20 lineups could help distinguish between these alternatives.
21
22
23
24
25
26
27

28 **Effect of Delay**

29
30 The only published meta-analysis of the verbal overshadowing effect (Meissner &
31
32 Brigham, 2001) found that the effect of providing a verbal description is reduced with a
33
34 delay between providing the description and completing the lineup identification task
35
36 (see also Finger & Pezdek, 1999). The present results are consistent with that
37
38 conclusion. The studies included in the meta-analysis varied in the length of the delay
39
40 and the materials used, meaning that the inference of a delay effect depended on
41
42 averaging across a number of other differences among the studies. The comparison of
43
44 RRR1 to RRR2 provides more compelling support for the conclusion that task order
45
46 alone, keeping all other aspects of the protocol constant, moderates the effect of
47
48 providing a verbal description on lineup accuracy.
49
50
51
52
53

54 Future research is needed to better understand the memory process responsible
55
56 for this difference. Switching the task order affects two aspects of the design: The delay
57
58
59
60

1
2
3 between witnessing the robbery and providing a description and the delay between
4 providing a description and completing the lineup identification task. Because the
5 lengths of these delays are confounded, it is impossible to separate the contributions of
6 these two types of delay to the difference between the two studies. Future research
7 could systematically vary the delays between the witnessed event, the description
8 tasks, and the identification task to see which contribute to the change in the effect of
9 verbally describing the robber.
10
11
12
13
14
15
16
17
18

19
20 A better of understanding of how each type of delay affects lineup identification
21 accuracy would be of both theoretical and practical importance. If verbally describing a
22 person impairs subsequent lineup identification accuracy under some delays but not
23 others, then those differences could inform police practices. For example, perhaps the
24 effect of providing a verbal description depends critically on having the lineup
25 identification task immediately follow the description. If so, then verbal overshadowing
26 would have less practical relevance: In most cases, the verbal description witnesses
27 provide to police does not immediately precede the lineup task. However, if the effect
28 instead depends only on the presence of a delay between witnessing an event and
29 describing the suspect, then the verbal overshadowing effect could have broad practical
30 importance: Witnesses rarely provide a verbal description immediately after witnessing
31 a crime, so verbal overshadowing could come into play in most eyewitness situations.
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48 **Reliability of Effect Size Over Time**

49
50 One of the central motivations for this RRR was the claim that the verbal
51 overshadowing effect had declined in size over the past 20 years, the so-called “decline
52 effect” (Schooler, 2011). Assessing whether or not an effect has diminished in size
53
54
55
56
57
58
59
60

1
2
3 depends critically on having a robust estimate of the effect size, both initially and later.
4
5 The RRR was designed to provide a robust estimate of the effect, one that could be
6
7 compared to that from the original study. However, the original studies used small
8
9 samples, meaning that the estimates were not precise. For example, the confidence
10
11 interval around the effect size for S&E-S Study 4 ranged from -44% to -0.79%. Although
12
13 that original effect size estimate falls outside the confidence intervals of our meta-
14
15 analytic effect size for that study, it is unclear whether the effect actually declined in size
16
17 or whether the original estimate was just an inaccurate estimate of the effect. Moreover,
18
19 RRR2 produced a meta-analytic effect size that was in line with that of the original
20
21 study, providing no compelling evidence for a change in the true effect over time with
22
23 that task order.
24
25
26
27
28

29 By providing a precise meta-analytic estimate of the true effect size, the RRR
30
31 studies provide guidance on the sample sizes needed to reliably detect the effect of
32
33 providing a verbal description on lineup identification performance. An analysis of the
34
35 sample sizes of earlier verbal overshadowing studies suggested that they were, on the
36
37 whole, substantially underpowered (Francis, 2012). The results of this RRR are largely
38
39 consistent with that conclusion. Only by combining across many larger-scale studies
40
41 could we detect the effect of providing a verbal description in RRR1. The confidence
42
43 intervals around an individual lab's effect size estimate are large (see the intervals
44
45 around individual lab studies in Figures 1 and 2 — even those studies with the largest
46
47 samples do not provide a highly precise estimate of the effect size). Even the
48
49 Mechanical Turk study, with nearly 200 participants in each condition, produced a
50
51 confidence interval with a range of approximately 12%. In other words, it could not have
52
53
54
55
56
57
58
59
60

1
2
3 reliably detected a significant difference from no effect with a true effect size of about
4
5 4%. In fact, all of the confidence intervals for the individual replications in RRR1
6
7 included 0. Had we simply tallied the number of studies providing clear evidence for an
8
9 effect in RRR1, we would have concluded in favor of a robust failure to replicate—a
10
11 misleading conclusion. Moreover, our understanding of the size of the effect would not
12
13 have improved. The purpose of the RRR approach is to better understand the true size
14
15 of important effects and not to make global succeed/fail judgments about individual
16
17 replication studies.
18
19
20
21

22 **Summary**

23
24 RRR1 and RRR2 combine the results of multiple, independent, direct
25
26 replications, to determine the size of the verbal overshadowing effect. In doing so, they
27
28 provide clear evidence for verbal overshadowing, particularly in the original task order
29
30 used by S&E-S Study 1 (description after a delay and just before line-up). Moreover, the
31
32 effect size estimates the RRRs provide can guide future research on verbal
33
34 overshadowing, both by suggesting new experimental questions and by indicating the
35
36 sample sizes needed to test those questions.
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgments

1
2
3
4
5
6
7
8 Geoff Cumming, Daniel Lakens, Joanne Yaffe, and John Protzko all provided helpful
9
10 guidance on the choice of a meta-analytic approach. Chris Meissner, Maryanne Garry,
11
12 Robert Michael, and Kim Wade spotted the erroneous task order in the protocol for
13
14 RRR1 and prompted a second replication study for this manuscript. Maryanne Garry
15
16 and Chris Meissner both reviewed the protocol for RRR2 to ensure that it matched the
17
18 parameters of Study 1 from S&E-S. Chris Meissner, Kim Wade, and Robert Michael
19
20 provided feedback on a preliminary draft of the manuscript, and Meissner also provided
21
22 the data he had compiled for the Meissner & Brigham (2001) meta-analysis for re-
23
24 analysis. Thanks to all 31 of the participating labs for devoting their time, efforts, and
25
26 resources to this replication project and to Brian Nosek and Jeffrey Spies for their
27
28 assistance with the registration process at OpenScienceFramework and for making it
29
30 possible for us to use OSF as the home for all materials for Registered Replication
31
32 Reports at *Perspectives*. Funding for participant payments was provided to individual
33
34 labs by the Association for Psychological Science via a grant from the Center for Open
35
36 Science. Finally, and most importantly, thanks to Jonathan Schooler for his cooperation
37
38 in developing the protocol and for his input and assistance throughout the process.
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- 1
2
3
4
5
6
7
8 Chin, J. M., & Schooler, J. W. (2008). Why do words hurt? Content, process, and
9
10 criterion shift accounts of verbal overshadowing. *European Journal of Cognitive*
11
12 *Psychology*, 20, 396-413.
13
14
15
16
17 Clare, J., & Lewandowsky, S. (2004). Verbalizing facial memory: Criterion effects in
18
19 verbal overshadowing. *Journal of Experimental Psychology: Learning, Memory,*
20
21 *and Cognition*, 30, 739-755.
22
23
24
25
26
27 Fallshore, M., & Schooler, J. W. (1995). Verbal vulnerability of perceptual expertise.
28
29 *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21,
30
31 1608-1623.
32
33
34
35
36
37 Finger, K., & Pezdek, K. (1999). The effect of the cognitive interview on face
38
39 identification accuracy: Release from verbal overshadowing. *Journal of Applied*
40
41 *Psychology*, 84, 340-348.
42
43
44
45
46 Francis, G. (2012). Too good to be true: Publication bias in two prominent studies from
47
48 experimental psychology. *Psychonomic Bulletin & Review*, 19, 151-156.
49
50
51
52
53 Lehrer, J. (2010, December 13). The truth wears off: Is there something wrong with the
54
55 scientific method? *The New Yorker*.
56
57
58
59
60

1
2
3
4
5
6 Meissner, C. A., & Brigham, J. C. (2001). A meta-analysis of the verbal overshadowing
7
8 effect in face identification. *Applied Cognitive Psychology*, 15, 603-616.
9

10
11
12 Schooler, J. W. (2011). Unpublished results hide the decline effect. *Nature*, 470, 437.
13
14

15
16
17 Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual
18
19 memories: Some things are better left unsaid. *Cognitive Psychology*, 22, 36-71.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Footnotes

1
2
3
4
5
6
7
8 1. This effect size estimate was based on a re-analysis of the data from the Meissner &
9
10 Brigham (2001) meta-analysis using the same effect size measure used in this RRR.
11
12 The data, a forest plot, and the R code used to conduct this analysis are available at
13
14 <https://osf.io/ybeur/>.
15
16

17
18
19
20 2. Traditional measures of the “file drawer” problem did not reveal substantial
21
22 publication bias in the verbal overshadowing literature (Meissner & Brigham, 2001), but
23
24 the power-based analysis likely is more sensitive in measuring the existence of
25
26 publication bias in the face of studies with small samples.
27
28

29
30
31
32 3. Tau is essentially the standard deviation of the total heterogeneity. In this case, Tau
33
34 is 1.07%. It is a measure of the distribution of the true effects. I^2 is an estimate of the
35
36 proportion of the heterogeneity that goes beyond what would be expected by chance. It
37
38 is the total heterogeneity divided by the total variability. H^2 is the total variability divided
39
40 by the sampling variability. The closer it is to 1, the more that the variability across effect
41
42 size estimates is consistent with sampling variability rather than meaningful
43
44 heterogeneity. Q is null-hypothesis test of whether there is meaningful heterogeneity.
45
46
47
48

49
50
51 4. S&E-S reported that errors consisted of 59% false alarms in the verbal description
52
53 condition and 60% false alarms in the control condition. Based on the raw numbers
54
55 provided in Jonathan Schooler’s dissertation data, the actual percentages were 59.3%
56
57
58
59
60

1
2
3 and 62.5%. In Figure 4, we used the raw numbers rather than the percentages reported
4
5
6 in S&E-S.
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Review Only

Figure Captions

Figure 1: Illustration of the task sequence for RRR1 (S&E-S Study 4) and RRR2 (S&E-S Study 1). Note that S&E-S used a different filler task, but a crossword puzzle was used in the replication studies at Jonathan Schooler's suggestion. Also, in S&E-S Study 4, the filler task lasted 10 minutes rather than 20 minutes. For the replication studies, we kept the duration of the filler task constant.

Figure 2: Forest plot of the verbal overshadowing effect in RRR1, with negative effects indicating lower accuracy for participants who verbally described the robber. This study replicates the task ordering from S&E-S Study 4. The data are listed in alphabetical order by the name of the first author from each replicating team.

Figure 3: Forest plot of the verbal overshadowing effect in RRR2, with negative effects indicating lower accuracy for participants who verbally described the robber. This study replicates the task ordering from S&E-S Study 1. The data are listed in alphabetical order by the name of the first author from each replicating team.

Figure 4: Forest plot of the difference in false identification rate on error trials between the verbal description condition and the countries/capitals condition for RRR1. Negative effects, those to the left of the vertical dashed line, constitute evidence that people who verbally described the robber were more likely to make the error of selecting the wrong

1
2
3 person from a lineup from a lineup. The data are listed in alphabetical order by the
4
5 name of the first author from each replicating team.
6
7
8
9

10 **Figure 5:** Forest plot of the difference in false identification rate on error trials between
11 the verbal description condition and the countries/capitals condition for RRR2. Negative
12 effects, those to the left of the vertical dashed line, constitute evidence that people who
13 verbally described the robber were more likely to make the error of selecting the wrong
14 person from a lineup from a lineup. The data are listed in alphabetical order by the
15 name of the first author from each replicating team.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Appendix: Individual Lab Details

Below, each lab briefly describes the characteristics of their sample and notes any substantial departures from the standard protocol. Each lab description identifies the authors and their affiliations in the order of their contributions to the project. Each also provides a link to that lab's OSF project page for the study where readers can see all of the details of the study including more complete method and results descriptions as well as the raw data. Labs are listed in the same order as in the tables and figures.

Amazon Mechanical Turk Variant

Robert B. Michael, Victoria University of Wellington

Gregory Franco, Victoria University of Wellington

Mevagh Sanson, Victoria University of Wellington

Maryanne Garry, Victoria University of Wellington

OSF: <https://osf.io/ez4w3/>

For both RRR Study 1 and 2, participants were recruited from Amazon Mechanical Turk, and were paid US\$2. Participants were randomly assigned to Study 1 or Study 2, and to the Control condition or the Experimental condition. We first collected 800 subjects with no pre-screened exclusion criteria. We then collected an additional 350 subjects with pre-screened exclusion criteria. These participants were required to self-report race as White and age between 18-25 to be eligible. We use custom software (see Turkitron.com) to track Mechanical Turk workers, preventing subjects from taking

1
2
3 the study multiple times. All participants were from the United States. We tracked and
4
5 excluded participants who did or reported any of the following: (1) failed to complete the
6
7 experiment, (2) failed to follow instructions, (3) failed an attention check, (4) failed to
8
9 give at least 5 countries and capitals in the control condition, (5) failed to engage
10
11 appropriately with the filler task, (6) had seen the robbery video before, or (7) had
12
13 already participated in a study just like this one.
14
15
16

17
18
19 Because the experiment was run online, subject behavior was not subject to the same
20
21 degree of control as a lab-based experiment. Specifically, MTurk workers have the
22
23 freedom to engage in other tasks or communicate with other people. We aimed to
24
25 reduce this undesirable activity by providing instructions to MTurk workers before they
26
27 began the experiment. These instructions asked that workers complete the experiment
28
29 in an environment free from distraction, that they give the experiment their full attention,
30
31 and that they have functioning audio. We also followed these instructions up with a
32
33 series of questions at the end of the experiment. These questions asked whether the
34
35 worker did in fact follow the instructions, with the assurance that they would receive
36
37 compensation regardless of their answers.
38
39
40
41
42
43
44

45
46 We also embedded an attention check question. This question requested that subjects
47
48 select "No" as their response to the question, and that they remember the word "horse"
49
50 to be entered on the following page. If subjects selected "Yes" as their response, or
51
52 failed to enter the word "horse", they were tagged for exclusion. At the end of the
53
54 experiment whether they had seen the video of the robbery before, and if they had
55
56
57
58
59
60

1
2
3 participated in a study like this one before. A response of "Yes" to either of these
4
5 questions resulted in an exclusion tag. Our filler task was a series of Sudoku puzzles.
6
7 We asked subjects at the end of the experiment whether they gave this task their full
8
9 attention. A response of "No" to this question resulted in an exclusion tag. Because of
10
11 technical limitations, we did not give our subjects a reminder at the 3 minute mark of the
12
13 experimental or control task. Our procedures, other than the deviations listed above,
14
15 followed the approved protocol.
16
17
18
19
20
21
22
23

24 **Labs that Completed Both RRR1 and RRR2**

25
26 Victoria K. Alogna, University of Otago

27
28 Jamin Halberstadt, University of Otago

29
30
31 Jonathon Jong, Institute Cognitive and Evolutionary Anthropology, University of Oxford

32
33
34 Joshua C. Jackson, University of Otago

35
36
37 Cathy Ng, University of Otago

38
39 OSF: <https://osf.io/sqzuq/>
40
41
42

43 For RRR Study 1, participants were 236 first and second year psychology
44
45 undergraduates (86 men, 150 women, M age = 20.2, SD=2.27) at the University of
46
47 Otago. One hundred and sixty-two took part during the school term, in exchange for
48
49 course credit. The remainder took part after their classes had ended, and were
50
51 remunerated NZ\$15 as reimbursement for travel expenses. Nine participants were
52
53 excluded due to computer software failures, and two because they did not complete the
54
55
56
57
58
59
60

1
2
3 experiment. Of the remaining 225 participants, 171 reported their race as “European”
4
5 (i.e., Caucasian). Per our registered experimental protocol, the analyses are based only
6
7 on these participants, though all data are available online.
8
9

10
11
12 For RRR Study 2, participants were 137 first and second year psychology
13
14 undergraduates (43 men, 94 women, M age = 20.43, SD = 3.99) at the University of
15
16 Otago. One hundred and thirty-one of these students took part during the school term,
17
18 in exchange for course credit; the remainder were remunerated NZ\$15 as
19
20 reimbursement for travel expenses. One hundred reported their race as “European”
21
22 (i.e., Caucasian). Per our registered experimental protocol, the analyses are based only
23
24 on these participants, though all data are available online. Our procedures followed the
25
26 approved protocol and did not deviate from our pre-registered plan.
27
28
29
30
31
32
33
34
35

36 Stacy L. Birch, College at Brockport SUNY

37 OSF: <https://osf.io/9zu4g/>
38
39
40
41
42

43 For RRR Study 1, participants were recruited from the introductory psychology
44
45 participant pool at the College at Brockport, State University of New York. They
46
47 participated as part of one option for course credit. Of the 156 participants in Study 1,
48
49 106 met inclusion criteria for the study according to their responses on the demographic
50
51 form. For RRR Study 2, 159 participants were recruited from the introductory
52
53 psychology pool at the College at Brockport (none of whom had participated in Study 1).
54
55
56
57
58
59
60

1
2
3 All received participation credit, but only 107 met inclusion criteria for the study based
4 on responses to the demographic questions. All data are available on the OSF page for
5 our studies. Our procedures followed the approved protocol and did not deviate from the
6 pre-registered plan.
7
8
9
10
11
12
13
14
15
16

17 Angela R. Birt, Mount Saint Vincent University

18 Philip Aucoin, Mount Saint Vincent University

19 OSF: <https://osf.io/y3xtf/>
20
21
22
23
24
25
26

27 For RRR Study 1, participants were recruited from undergraduate courses at Mount
28 Saint Vincent University in Halifax, NS, Canada. At the discretion of course instructors,
29 most (92.30%) received course points for participating. Participants in RRR Study 2
30 were recruited in the same way as Study 1, but were compensated \$8 for participating.
31 Recruitment in both studies did not include restrictions on race or age; therefore, the
32 overall samples included participants who did not meet inclusion criteria for this RRR.
33 Data from participants excluded from analyses are included on our OSF page.
34 Following our pre-registered plans for both studies, we included a self-report question of
35 visual acuity, a brief test to measure potential effects of demand characteristics on the
36 results, and collected data on reaction times for making identifications and confidence
37 ratings. Otherwise, our procedures followed the standard protocols. Analyses of the
38 additional data can be found on our OSF page.
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6 Maria A. Brandimonte, Suor Orsola Benincasa University, Naples Italy
7

8 OSF: <https://osf.io/gfyyd/>
9
10

11 12 13 Summary

14
15 For RRR Study 1, 140 participants were recruited from the participant pool at Suor
16
17 Orsola Benincasa University of Naples, in Italy, and they participated as part of one
18
19 option for course credit. None of them was enrolled in a Psychology course. For RRR
20
21 Study 2, 100 participants were recruited. All participants were White. All participants
22
23 were able to understand the instructions and had vision adequate to watch the video
24
25 and see the images. Given that our participants were not native English speakers,
26
27 instructions were translated into Italian and then translated back into English
28
29 independently by the two labs participating in this replication effort from Italy (ours and a
30
31 lab at the University of Chieti). We also replaced the English audio track with an Italian
32
33 translation. Both Laboratories used the same translated materials and dubbed video on
34
35 which there had been full agreement. Finally, we added an additional question at the
36
37 end of the study to verify that participants understood that the video depicted a bank
38
39 robbery. No participant was excluded. In all other respects, our procedure followed the
40
41 standard protocol.
42
43
44
45
46
47
48
49
50
51
52

53 Curt A. Carlson, Texas A&M University - Commerce

54
55 Dawn R. Weatherford, Arkansas State University
56
57
58
59
60

1
2
3 Maria A. Carlson, Texas A&M University - Commerce
4

5 OSF: <https://osf.io/s73uq/>
6
7
8
9

10 For RRR Study 1, participants were recruited from the psychology participant pool at
11 Texas A&M University - Commerce, and they participated as part of one option for
12 course credit. For RRR Study 2, participants were recruited from the psychology
13 participant pool at Arkansas State University, and they also participated for course
14 credit. Each participant pool uses a pre-screening process that allowed us to recruit only
15 participants who met the specified inclusion criteria, so only participants who failed to
16 complete the study were excluded. Our procedures followed the approved protocol and
17 did not deviate from our pre-registered plan.
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

34 Kimberly S. Dellapaolera, University of Nebraska-Lincoln
35

36 Brian H. Bornstein, University of Nebraska-Lincoln
37

38 OSF:
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For RRR Study 1, participants were recruited from the psychology participant pool at the University of Nebraska-Lincoln, and they participated as part of one option for course credit. For RRR Study 2, participants were recruited from the same participant pool; participants could only participate in one of the two studies. We recruited from our participant pool without specifying restrictions on race or age, so our total sample for Study 1 included an additional 22 participants and our total sample for Study 2 included

1
2
3 an additional 26 participants who did not meet inclusion criteria. All data are provided on
4 our OSF page. Our procedures followed the approved protocol and did not deviate from
5
6 our pre-registered plan.
7
8
9

10
11
12
13
14
15 Jean-Francois Delvenne, University of Leeds

16
17
18 Charity Brown, University of Leeds

19
20 Emma Portch, University of Leeds

21
22 Tara Zaksaitė, University of Leeds

23
24 OSF: <https://osf.io/vucan/>
25
26
27
28

29 For RRR Study 1, 93 participants were recruited from the participant pool at the
30 University of Leeds (they participated as part as one option for course credit) and 37
31
32 participants were recruited from the broad campus and were compensated £5 for
33
34 participating. For RRR Study 2, 43 participants were recruited from our participant pool
35
36 and 55 participants were recruited from the broad campus and were compensated £5
37
38 for participating. Our participant pool uses a pre-screening process that allowed us to
39
40 recruit only participants who met the specified inclusion criteria, so only participants who
41
42 failed to complete the study (i.e., 10 in study 1; 4 in study 2) were excluded. Our
43
44 procedures followed the approved protocol and did not deviate from our pre-registered
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60 plan.

1
2
3 Gerald Echterhoff, University of Münster

4
5 René Kopietz, University of Münster

6
7 OSF: <https://osf.io/dmuqj/>

8
9
10
11
12 For RRR Study 1, participants were recruited from the introductory psychology
13 participant pool at the University of Münster in Germany, and they participated as part of
14 one option for course credit. For RRR Study 2, approx. 20% of the participants were
15 recruited from the participant pool and the remaining 80% were recruited from the
16 broader campus and were compensated €6 for participating. We recruited from our
17 participant pool and community without specifying restrictions on race or age, so our
18 total sample included an additional 39 participants—14 in Study 1 and 25 in Study 2—
19 who did not meet inclusion criteria for this RRR. Data from those participants are
20 included on our OSF page. Additionally, data on the OSF page includes participants
21 who did not understand the nature of the event as well as a sample (N = 36) from our
22 initial attempt to run Study 1 with the original English-language version of the video.
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

41 Given that our participants were not native English speakers, the second author
42 translated all instructions to German, and a bilingual student assistant independently
43 translated them back to English to verify the accuracy of the translation. Based on a
44 small, informal pretest, we initially assumed that our participants would be able to
45 understand the video with the original sound track and therefore did not dub it.
46
47
48
49
50
51
52
53 However, we added an additional question at the end of the study to verify that
54 participants understood that the video depicted a bank robbery.
55
56
57
58
59
60

1
2
3 Because many participants' did not understand the nature of the event depicted in the
4 original video, we changed the protocol to replace the English audio track with a
5 German translation. We informed the editors about this breach of protocol, and
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Because many participants' did not understand the nature of the event depicted in the original video, we changed the protocol to replace the English audio track with a German translation. We informed the editors about this breach of protocol, and excluded all participants who watched the original version of the video from the final sample. Based on our pre-registered plan, we excluded any participants who did not understand the nature of the video. Because of the need for this change to Study 1, we were unable to reach the pre-registered 50 participants per condition (final sample: $n=46$ in the control and $n=41$ in the experimental condition). Similarly, because of the need to mainly recruit participants outside the psychology department for Study 2, we did not reach our goal of 50 participants per condition after exclusion of problem cases (final sample: $n=47$ in the control and $n=46$ in the experimental condition). In all other respects, our procedure followed the standard protocol.

Casey M. Eggleston, University of Virginia

Calvin K. Lai, University of Virginia

Elizabeth A. Gilbert, University of Virginia

OSF: <https://osf.io/b4g79/>

For RRR Studies 1 and 2, participants were recruited from the introductory psychology participant pool at the University of Virginia, and they participated as part of one option for course credit. Of the 180 participants who partook in Study 1, 25 were excluded prior to data analysis based on a priori criteria (e.g., failing to meet the target study

1
2
3 demographics, improperly answering the attention catch question), and 5 participants
4
5 were unexpectedly excluded for failing to sign a proper consent form. Of the 94
6
7 participants who partook in Study 2, 10 were excluded prior to data analysis based on a
8
9 priori criteria. Data from all participants who completed the study and gave informed
10
11 consent are provided on our OSF page. Our procedures otherwise followed the
12
13 approved protocol and did not deviate from our pre-registered plan.
14
15
16
17
18
19
20
21

22 Daniel L. Greenberg, College of Charleston

23
24 Marino A. Mugayar-Baldocchi, College of Charleston

25
26 OSF: <https://osf.io/sieea/>
27
28
29
30
31

32 For both RRR Study 1 and RRR Study 2, participants were recruited from the
33
34 introductory psychology participant pool at the College of Charleston, and they
35
36 participated as part of one option for course credit. For both studies, recruitment was
37
38 conducted without specifying restrictions on race or age, so our total sample included
39
40 participants who did not meet inclusion criteria for this RRR (19 in Study 1 and 14 in
41
42 Study 2). Data from those participants are included on our OSF page but were excluded
43
44 from the analyses reported here. Our procedures followed the approved protocol and
45
46 did not deviate from our pre-registered plan.
47
48
49
50
51
52
53
54

55 Andre Kehn, University of North Dakota
56
57
58
59
60

1
2
3 Kimberly Schweitzer, University of Wyoming

4
5
6 Bradley W. Gamblin, University of North Dakota

7
8 Kimberly Wiseman, University of Wyoming

9
10 Narina L. Nunez, University of Wyoming

11
12 OSF: <https://osf.io/mkz84/>

13
14
15
16
17 For RRR Study 1 and 2, participants were recruited from the psychology participant
18 pools at the University of North Dakota and the University of Wyoming, and they
19 participated to receive course credit or extra credit. We oversampled for both studies in
20 order to reach the minimum participant numbers. Participants were excluded if they did
21 not meet the age or race requirements (n=10 in Study 1, n=7 in Study 2) . Further,
22 participants were also excluded if they did not complete the study. Our procedures
23 followed the approved protocol and did not deviate from our pre-registered plan.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38

39 Christopher Koch, George Fox University

40
41 Remi Gentry, George Fox University

42
43 Jennifer Shaheed, George Fox University

44
45 Kelsi Buswell, George Fox University

46
47 OSF: <https://osf.io/bym2a/>

48
49
50
51
52 For RRR Study 1, participants were recruited from General Psychology courses at
53 George Fox University for research participation credit. A total of 109 participants
54
55
56
57
58
59
60

1
2
3 completed the study. However, 13 participants were removed from the analysis for not
4
5 meeting the RRR inclusion criteria. The remaining participants (62 females and 34
6
7 males) were equally divided between the control and experimental conditions. Age
8
9 ranged between 18 to 23 years with a mean of 19.27 (SD = 1.14). For RRR Study 2, 46
10
11 participants were recruited from General Psychology courses for research participation
12
13 credit. An additional 21 volunteers were recruited by asking participants who had just
14
15 completed the study to suggest other people who might be willing to volunteer
16
17 (“snowball” recruiting). Five participants were removed from the analysis for not meeting
18
19 the RRR inclusion criteria and two were removed for invalid responses. The remaining
20
21 participants (45 females and 15 males) were equally divided between the control and
22
23 experimental conditions. Age ranged between 18 to 23 years with a mean of 20.08 (SD
24
25 = 1.83). Data from all participants are reported on our OSF page. Other than the use of
26
27 snowball recruiting to meet the specified sample size for Study 2, our procedure for both
28
29 studies followed the approved protocol and did not deviate from our pre-registered plan.
30
31
32
33
34
35
36
37
38
39
40

41 Nicola Mammarella, University of Chieti

42
43 Beth Fairfield, University of Chieti

44
45 Alberto Di Domenico, University of Chieti

46
47 OSF: <https://osf.io/edsrz/>

48
49
50
51
52
53 For both RRR Study 1 and 2, participants were recruited from an introductory
54
55 psychology course participant pool at the University of Chieti in Italy, and they
56
57
58
59
60

1
2
3 participated for course credit. In both studies we recruited without specifying restrictions
4
5 on race or age. Of the 232 participants in Study 1, 12 did not meet the inclusion criteria.
6
7 Of the 104 participants in Study 2, 4 did not meet the exclusion criteria. Data from all
8
9 participants are provided on our OSF page. Given that our participants were not native
10
11 English speakers, one of the authors translated all instructions to Italian, and a second
12
13 author independently translated them back to English to verify the accuracy of the
14
15 translation. We also replaced the English audio track with an Italian translation. In all
16
17 other respects, our procedure followed the standard protocol.
18
19
20
21
22
23
24
25
26

27 Shannon K. McCoy, University of Maine

28
29 Arielle Rancourt, University of Maine

30
31 OSF: <https://osf.io/ejj7d/>
32
33
34
35

36 For RRR Study 1 and Study 2, participants were recruited from the introductory
37
38 psychology participant pool at the University of Maine, and they participated for course
39
40 credit. We recruited from our participant pool without specifying restrictions on race or
41
42 age, so our total sample included an additional 32 participants who did not meet
43
44 inclusion criteria for the RRR (N = 25 from Study 1; N = 7 from Study 2). Data from
45
46 those participants are included on our OSF page. Our procedures followed the
47
48 approved protocol and did not deviate from our pre-registered plan.
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Abigail A. Mitchell, Nebraska Wesleyan University

4
5 Marilyn S. Petro, Nebraska Wesleyan University

6
7 OSF: <https://osf.io/zqnjb/>

8
9
10
11
12 For both RRR Study 1 and Study 2, participants were recruited from the Nebraska
13 Wesleyan Psychology Department's participant pool. Students participated as part of
14 one option for course credit. We recruited without specifying age or race restrictions.
15 For Study 1, 127 were recruited, however 15 did not meet the inclusion criteria. For
16 Study 2, 109 participated, but data from 17 were excluded due to not meeting inclusion
17 criteria. Our procedures followed the approved protocol and did not deviate from our
18 pre-registered plan which included open-ended debriefing questions concerning
19 perceptions of the study.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

36 Robin Musselman, Lehigh Carbon Community College

37
38 Michael Colarusso, Lehigh Carbon Community College

39
40 OSF: <https://osf.io/ybfmu/>

41
42
43
44
45 For RRR Study 1, 101 participants were recruited from Introduction to Psychology
46 courses at Lehigh Carbon Community College, and they participated in most cases for
47 extra credit in their course (whether participants received course credit was determined
48 by the course instructors and was not under the experimenters' control). For RRR Study
49 2, 60 participants were recruited from Lehigh Carbon Community College and 15 were
50
51
52
53
54
55
56
57
58
59
60

1
2
3 recruited from Cedar Crest College, with students receiving extra credit for participating.
4
5 We recruited without specifying restrictions on race, so our total sample included an
6
7
8 additional 41 participants (23 in Study 1 and 18 in Study 2) whose data are reported on
9
10 our OSF page. Our procedures followed the protocol and we did not deviate from our
11
12 pre-registered plan, with the exception of recruiting at a neighboring college to meet our
13
14 specified sample size for Study 2.
15
16
17
18
19
20
21

22 Christopher R. Poirier, Stonehill College

23
24 Matthew K. Attaya, Stonehill College

25
26
27 Kathleen A. McConaughy, Stonehill College

28
29 Jessica E. Pappagianopoulos, Stonehill College

30
31
32 Griffin A. Sullivan, Stonehill College

33
34 OSF: <https://osf.io/zgmex/>
35
36
37
38

39 For RRR Study 1 and Study 2, participants were recruited from the psychology
40
41 department's participant pool at Stonehill College, and they participated as part of one
42
43 option for course credit. Our participant pool uses a pre-screening process that allowed
44
45 us to recruit only participants who met the specified inclusion criteria; however, a
46
47 participant in Study 2 was excluded because he identified as both White and Black
48
49 during the study. Our procedures followed the approved protocol and did not deviate
50
51 from our pre-registered plan.
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6 Eva Rubínová, Masaryk University
7

8 Marek A. Vranka, Charles University in Prague
9

10 Štěpán Bahník, University of Würzburg
11

12 OSF: <https://osf.io/ikuh7/>
13
14
15
16

17 For both RRR Studies, participants were recruited from our laboratory subject pool
18 consisting of students of Czech universities, and they were compensated 100 CZK
19 (approx. \$5) for participation. Our participant database allows us to use a pre-screening
20 process, so we invited only participants who met the specified inclusion criteria. Given
21 that our participants were not native English speakers, the authors translated all
22 instructions to Czech, and an independent bilingual speaker translated them back into
23 English to verify the accuracy of the translation. We also replaced the English audio
24 track with its Czech translation and used a Czech crossword puzzle similar to the one
25 used in the original study. We did not include any comprehension checks as all of our
26 participants were native or fluent Czech speakers. The study was run on computers
27 (except for the crossword puzzle and robber description/capitals listing, which were
28 completed on paper), and we added some procedural instructions to be able to run the
29 study without additional instructions from the experimenter during the main part of the
30 session. Following our pre-registered plan, we added three questions at the end of the
31 session to check participants' knowledge of the experiment. In Study 1, based on the
32 answers, we excluded participants who stated that they: (i) knew about the experimental
33 procedure or hypothesis from other participants, (ii) knew the tested hypothesis, or (iii)
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 knew what the verbal overshadowing effect is (n=12 excluded). In Study 2, participants
4
5 also had to write down the hypothesis and/or what verbal overshadowing effect is, and
6
7
8 we excluded only those who answered correctly (n=8 excluded). In all other respects,
9
10 our procedure followed the standard protocol.
11
12
13
14
15
16

17 Kyle J. Susa, University of Texas at El Paso

18
19 Jessica K. Swanner, Iowa State University

20
21
22 Christian A. Meissner, Iowa State University

23
24
25 OSF: <https://osf.io/5vunt/>
26
27
28

29 For RRR Studies 1 and 2 participants were recruited from the introductory psychology
30
31 participant pool at Iowa State University and they participated as part of one option for
32
33 course credit. For RRR Study 1, 145 participants were recruited. For RRR Study 2, 111
34
35 participants were recruited. In accordance with IRB approval we did not restrict our
36
37 participants by race or age, however only participants who met the inclusion criteria
38
39 were evaluated in our analyses. In RRR Study 1, 35 participants did not meet the
40
41 inclusion criteria, in RRR Study 2, 11 participants did not meet the inclusion criteria.
42
43
44 Data from all participants are reported on our OSF page. Our procedures followed the
45
46 approved protocol and did not deviate from our pre-registered plan.
47
48
49
50
51
52
53
54

55 W. Burt Thompson, Niagara University
56
57
58
59
60

1
2
3 OSF: <https://osf.io/4ijas/>
4
5
6
7

8 For RRR Study 1, participants were recruited from psychology classes at Niagara
9 University, and they participated in return for course credit. The primary sample consists
10 of the first 100 participants, 50 per condition, who met all criteria for inclusion in the
11 study. An additional 31 participants either did not meet one or more of the study criteria
12 (e.g., age, ethnicity) or were tested after the primary data set had been collected. For
13 RRR Study 2, the primary sample consists of the first 77 students who met all of the
14 criteria for the study, 38 in the description (experimental) condition, and 39 in the
15 capitals (control) condition. An additional 25 students were tested but did not meet all
16 study criteria. All participants were recruited from Niagara University psychology and
17 criminal justice classes. Fifteen of the participants were compensated \$5 and the others
18 received course credit. For both studies, our procedures followed the approved protocol
19 and did not deviate from our pre-registered plan.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

41 Joanna Ulatowska, Academy of Special Education, Warsaw, Poland

42 Aleksandra Cislak, University of Social Sciences and Humanities, Warsaw, Poland

43 OSF: <https://osf.io/bzhvf/>
44
45
46
47
48
49

50 For RRR Study 1, participants were recruited among social sciences students through
51 study advertisements and personally by research assistants at the campus of Academy
52 of Special Education in Poland. For RRR Study 2, participants were recruited at the
53
54
55
56
57
58
59
60

1
2
3 campuses of Academy of Special Education and University of Social Sciences and
4 Humanities in Poland. They participated in return for a gift voucher (25 PLN,
5
6 approximately \$8.16). Only undergraduate students within age range 18-25, who
7
8 claimed to speak English were recruited. Given that our participants were not native
9
10 English speakers, all of the instructions were been translated to Polish by one of the
11
12 experimenters and then translated back to English by a fluent English speaker. The
13
14 independent translator was blind to the study topic. All Polish instructions were also
15
16 verified using Google Translate. At the end of the study, we asked an additional
17
18 question to verify that participants understood that the video depicted a bank robbery.
19
20 Based on our pre-registered plan, we excluded any participants who did not understand
21
22 the nature of the video. In Study 1, 15 participants (10 women) were excluded from
23
24 further analyses as they did not understand the sense of robber's words and two more
25
26 women were excluded as they exceeded the age limit. In Study 2, 12 participants (11
27
28 women) were excluded from further analyses as they did not understand the sense of
29
30 robber's words. Data from those participants are included on our OSF page. In all other
31
32 respects, our procedure followed the standard protocol.
33
34
35
36
37
38
39
40
41
42
43
44
45

46 Kimberley A. Wade, University of Warwick

47
48 Ulrike Körner, Heinrich Heine University

49
50 Melissa F. Colloff, University of Warwick

51
52 Melina A. Kunar, University of Warwick

53
54
55 OSF: <https://osf.io/dbxv4/>
56
57
58
59
60

1
2
3
4
5
6 For RRR Study 1, 68 of the participants were recruited from the introductory psychology
7
8 participant pool at the University of Warwick, and they participated as part of one option
9
10 for course credit. The other 52 participants were recruited from the broader campus via
11
12 a university-wide participant pool and were compensated £3. For RRR Study 2,
13
14 participants were recruited from across the University of Warwick campus via the
15
16 university-wide participant pool and were compensated £3. Our participant pool uses a
17
18 pre-screening process that allowed us to recruit only participants who met the specified
19
20 inclusion criteria, so only 1 participant in Study 2 who failed to complete the study was
21
22 excluded. Our procedures followed the approved protocol and did not deviate from our
23
24 pre-registered plan.
25
26
27
28
29
30
31
32
33

34 **Labs that Completed Only RRR1**

35
36 Simon Chu, Ashworth Research Centre

37
38 John E. Marsh, University of Central Lancashire

39
40 Faye Skelton, University of Central Lancashire

41
42 OSF: <https://osf.io/qu3zp/>
43
44
45
46
47

48 For RRR Study 1, 79 participants were recruited from the undergraduate psychology
49
50 participant pool at the University of Central Lancashire and participated as one option in
51
52 return for course credit. Twenty-two participants from across the broader university
53
54 campus were also recruited through a university online bulletin board. Participants
55
56
57
58
59
60

1
2
3 recruited from outside the psychology department volunteered their time. We operated a
4 pre-screening process that allowed us to recruit only participants who met the specified
5 inclusion criteria, so only participants who failed to complete the study were excluded.
6
7
8 Owing to time constraints, we were forced to close the study before meeting our original
9
10 recruitment target of 120. Our experimental procedure followed the approved protocol.
11
12
13
14
15
16
17
18
19

20 John E. Edlund, Rochester Institute of Technology

21 Austin Lee Nichols, University of Navarra

22 OSF: <https://osf.io/ybswb/>
23
24
25
26
27
28

29 For RRR Study 1, participants were recruited from the introductory psychology
30 participant pool at the Rochester Institute of Technology, and they participated as part
31 of one option for course credit. Due to our limited participant pool, we were unable to
32 complete RRR Study 2. We recruited from our participant pool without specifying
33 restrictions on race or age, so our total sample for Study 1 included an additional 22
34 participants who did not meet the inclusion criteria for this RRR. Data from those
35 participants are included on our OSF page. Our total included sample consisted of 61
36 participants in the control condition and 51 participants in the experimental condition.
37
38 Our procedures followed the approved protocol and did not deviate from our pre-
39 registered plan.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Fiona Gabbert, Goldsmiths University of London

4
5 Tim Valentine, Goldsmiths University of London

6
7 OSF: <https://osf.io/rmdz7/>

8
9
10
11
12 For RRR Study 1, participants were recruited as part of their Research & Methods
13 laboratory class on chi-square analysis at Goldsmiths University of London. They were
14 not required to take part, but everyone did. No compensation was given. We recruited
15 without specifying restrictions on race or age, so our total sample included an additional
16 45 participants who did not meet inclusion criteria for this RRR. Our procedures
17 followed the approved protocol and did not deviate from our pre-registered plan. Due to
18 having tested all of our first year Psychology students for Study 1, we were unable to
19 complete RRR Study 2.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

36 Fabio P. Leite, The Ohio State University at Lima

37 OSF: <https://osf.io/kmibs/>

38
39
40
41
42 For RRR Study 1, 128 participants were recruited from the introductory psychology
43 participant pool at the Ohio State University at Lima, and they participated as part of
44 one option for course credit. We recruited from our participant pool without specifying
45 restrictions on race or age. Twenty seven participants did not meet inclusion criteria for
46 this RRR, and their data are included on our OSF page. Due to our limited participant
47 pool, we were unable to complete RRR Study 2. The incomplete data set for Study 2 is
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 also included on our OSF page. Our procedures followed the approved protocol and did
4
5 not deviate from our pre-registered plan.
6
7
8
9

10
11
12 Alex H. McIntyre, University of Stirling

13
14
15 Stephen R. H. Langton, University of Stirling

16
17
18 Peter J. B. Hancock, University of Stirling

19
20 OSF: <https://osf.io/3rn5f/>
21
22
23

24 For RRR Study 1, 103 participants were recruited from the introductory psychology
25 participant pool at the University of Stirling in Scotland, and they participated as part of
26 one option for course credit. A further 7 participants were excluded due to age criteria,
27 and 1 was excluded in line with race criteria. Data from the excluded participants are
28 included on our OSF page. For RRR Study 2, we were unable to recruit the required
29 sample of 30 participants in each group and just 24 participants were recruited from the
30 participant pool. Data from all participants are available on our OSF page. Our
31 procedures followed the approved protocols and did not deviate from our pre-registered
32 plan.
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

50 Robert B. Michael, Victoria University of Wellington

51
52 Gregory Franco, Victoria University of Wellington

53
54
55 Mevagh Sanson, Victoria University of Wellington
56
57
58
59
60

1
2
3 Maryanne Garry, Victoria University of Wellington
4

5 OSF: <https://osf.io/bnzzrj/>
6
7
8
9

10 Participants were recruited from the introductory psychology participant pool at Victoria
11 University of Wellington, and participated for course credit. Our participant pool uses a
12 pre-screening process that allowed us to exclude, post-hoc, participants who did not
13 meet the specified inclusion criteria. We also excluded subjects who failed to complete
14 the experiment, or when there were other procedural difficulties, such as sound
15 malfunctions on the video. The results we report are from a dataset with strict exclusion
16 criteria, but we have additional datasets available on our OSF page with less strict
17 exclusion criteria that may be of interest to researchers. Our procedures followed the
18 approved protocol and did not deviate from our pre-registered plan.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35

36 Matthew A. Palmer, University of Tasmania

37
38 Aaron Drummond, Flinders University

39
40 James D. Sauer, University of Portsmouth

41
42 Daniel V. Zuj, University of Tasmania

43
44 Glenys A. Holt, University of Tasmania

45
46 Miriam Rainsford, University of Tasmania

47
48 Lauren Hall, Flinders University

49
50 Liam Satchell, University of Portsmouth
51
52
53
54
55
56
57
58
59
60

1
2
3 For RRR Study 1, 107 participants were recruited from three locations: the University of
4
5 Tasmania (comprising 19 recruited from the introductory psychology participant pool
6
7 who received course credit and 17 from the broader campus community who were
8
9 compensated \$10); Flinders University (19 recruited from the broader campus
10
11 community who were compensated \$15); and The University of Portsmouth (52
12
13 recruited from the broader campus community who volunteered their time). Due to our
14
15 limited participant pool, we were unable to complete RRR Study 2 (we recruited 29
16
17 participants from the introductory psychology participant pool at the University of
18
19 Tasmania). We recruited from our participant pool and community without specifying
20
21 restrictions on race or age, so our total sample included an additional 16 participants in
22
23 Study 1 and 23 participants in Study 2 who did not meet inclusion criteria for this RRR.
24
25 Data from all participants are included on our OSF page. Due to experimenter error, 41
26
27 participants in Study 1 received a version of the response questionnaire in which
28
29 subjects made their identification response and identification confidence rating on the
30
31 same page, rather than different pages. This had minimal effect on identification
32
33 accuracy and the results of the main analyses. Details of these extra analyses are
34
35 included on our OSF page.
36
37
38
39
40
41
42
43
44
45
46
47

48 Peter P. J. L. Verhoeven, Erasmus University Rotterdam

49 Samantha Bouwmeester, Erasmus University Rotterdam

50 Rolf A. Zwaan, Erasmus University Rotterdam

51 OSF: <https://osf.io/wtbkp/>
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6 The results of RRR Study 1 were obtained by exactly executing the sampling plan and
7
8 procedure described on our lab's project page at the Open Science Framework. We
9
10 tested 115 Dutch-speaking Erasmus University undergraduates (most of them were
11
12 Psychology undergraduates) who took part in the experiment to meet their course
13
14 requirements. Ten participants did not meet the inclusion criteria: 5 were non-white, 1
15
16 had heard about crucial experiment characteristics prior to participation (note that 3
17
18 other participants also indicated they heard about the experiment before, but 2 reported
19
20 the characteristics of a different unrelated experiment and 1 had only heard about the
21
22 crossword puzzle; these 3 participants were not excluded) and 4 failed to adhere to the
23
24 experimental instructions. After exclusion, the sample consisted of 105 participants, with
25
26 51 participants in the experimental (i.e., verbal overshadowing) condition and 54 in the
27
28 control condition. Because our participants were not native English speakers, we used
29
30 translated instructions in the present study. To obtain the Dutch instructions, one of the
31
32 members of the research team (Verkoeijen) translated the English instructions from the
33
34 approved protocol and a colleague at the Department of Psychology of the Erasmus
35
36 University Rotterdam checked whether the translated versions matched the meaning of
37
38 their English counterparts. The translations were adjusted based on this feedback.
39
40
41
42
43
44
45
46
47
48
49

50 Christopher A. Was, Kent State University

51 Dale Hirsch, Kent State University

52 Rachael Todaro, Kent State University
53
54
55
56
57
58
59
60

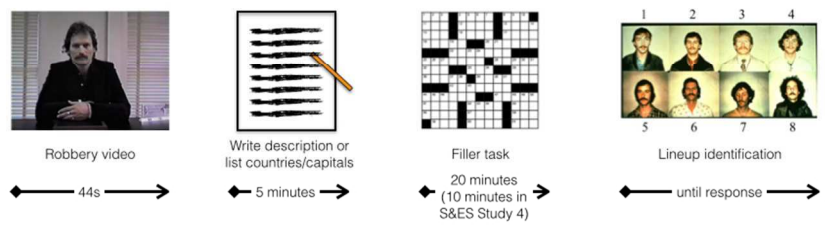
1
2
3 Connie Romig, Kent State University
4

5 OSF: <https://osf.io/fub7j/>
6
7
8
9

10 For RRR Study 1, 145 participants were recruited from the educational psychology
11 participant pool at the Kent State University, and they participated as part of one option
12 for course credit. Our participant pool uses a pre-screening process that allowed us to
13 recruit only participants who met the specified inclusion criteria. Ten participants who
14 failed to complete the study were excluded from analyses. Our procedures followed the
15 approved protocol and did not deviate from our pre-registered plan.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Sequence for RRR Study 1 and S&E-S Study 4



Sequence for RRR Study 2 and S&E-S Study 1

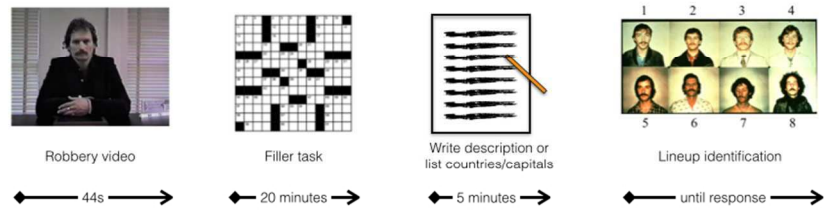


Illustration of the task sequence for RRR1 (S&E-S Study 4) and RRR2 (S&E-S Study 1). Note that S&E-S used a different filler task, but a crossword puzzle was used in the replication studies at Jonathan Schooler's suggestion. Also, in S&E-S Study 4, the filler task lasted 10 minutes rather than 20 minutes. For the replication studies, we kept the duration of the filler task constant.

361x270mm (72 x 72 DPI)

Only

Perspectives on Psychological Science

Study	Verbal	Control		Difference [95% CI]
-------	--------	---------	--	---------------------

1 S&E-S STUDY 4	48.6	71.1		-0.22 [-0.44 , -0.01]
2				
3 ONLINE MTURK - Michael, Franco, Sanson, Garry	54.9	48.9		0.06 [-0.04 , 0.16]
4				

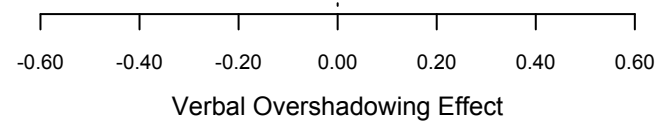
Completed Both RRR Studies

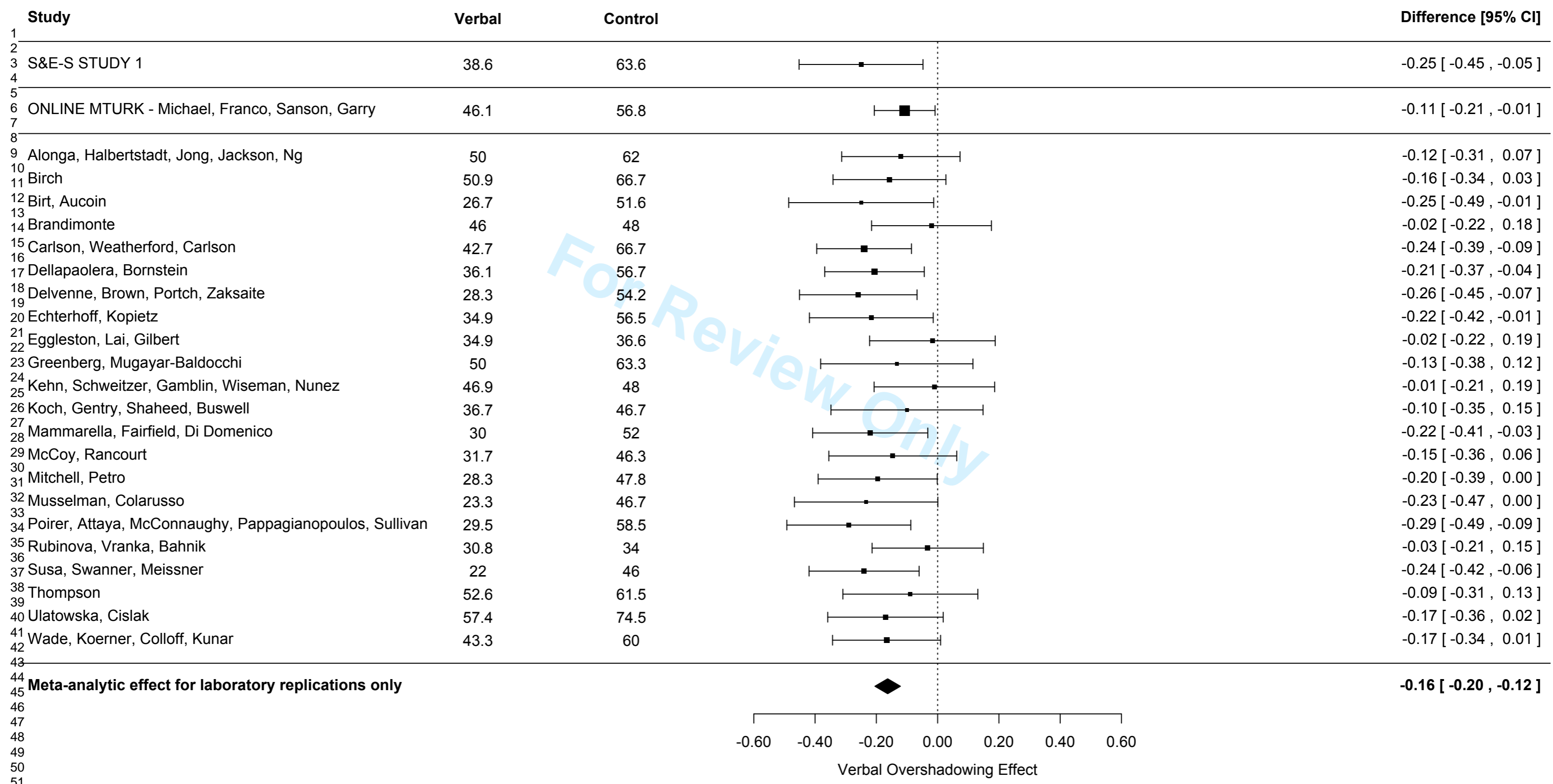
7 Alonga, Halbertstadt, Jong, Jackson, Ng	57.1	70.2		-0.13 [-0.31 , 0.05]
8 Birch	66.1	66		0.00 [-0.18 , 0.18]
9 Birt, Aucoin	33.3	36		-0.03 [-0.21 , 0.16]
10 Brandimonte	48.6	38.6		0.10 [-0.06 , 0.26]
11 Carlson, Weatherford, Carlson	42.7	56		-0.13 [-0.29 , 0.03]
12 Dellapaolera, Bornstein	39.2	53.9		-0.15 [-0.31 , 0.01]
13 Delvenne, Brown, Portch, Zaksait	67.8	65.6		0.02 [-0.15 , 0.19]
14 Echterhoff, Kopietz	36.6	52.2		-0.16 [-0.36 , 0.05]
15 Eggleston, Lai, Gilbert	49.4	56.5		-0.07 [-0.23 , 0.09]
16 Greenberg, Mugayar-Baldocchi	48	56		-0.08 [-0.28 , 0.12]
17 Kehn, Schweitzer, Gamblin, Wiseman, Nunez	52.4	57.4		-0.05 [-0.22 , 0.12]
18 Koch, Gentry, Shaheed, Buswell	54.2	45.8		0.08 [-0.12 , 0.28]
19 Mammarella, Fairfield, Di Domenico	40.9	52.7		-0.12 [-0.25 , 0.01]
20 McCoy, Rancourt	41.9	57.4		-0.15 [-0.33 , 0.02]
21 Mitchell, Petro	54.8	51		0.04 [-0.15 , 0.22]
22 Musselman, Colarusso	23.5	40		-0.16 [-0.34 , 0.01]
23 Poirer, Attaya, McConaughy, Pappagianopoulos, Sullivan	43.6	56.4		-0.13 [-0.31 , 0.06]
24 Rubinova, Vranka, Bahnik	52.9	46		0.07 [-0.11 , 0.25]
25 Susa, Swanner, Meissner	41.8	41.8		0.00 [-0.18 , 0.18]
26 Thompson	56	42		0.14 [-0.05 , 0.33]
27 Ulatowska, Cislak	72.7	69.1		0.04 [-0.13 , 0.21]
27 Wade, Koerner, Colloff, Kunar	56.7	65		-0.08 [-0.26 , 0.09]

Completed RRR Study 1 Only

30 Chu, Marsh, Skelton	54	62.7		-0.09 [-0.28 , 0.10]
31 Edlund, Nichols	54.9	42.6		0.12 [-0.06 , 0.31]
32 Gabbert, Valentine	61.8	58.5		0.03 [-0.15 , 0.22]
33 Leite	38.5	56		-0.18 [-0.37 , 0.02]
34 McIntyre, Langton, Hancock	71.2	72.5		-0.01 [-0.19 , 0.16]
35 Michael, Franco, Sanson, Garry	56.3	56.1		0.00 [-0.14 , 0.15]
36 Palmer et al	57.1	58.8		-0.02 [-0.20 , 0.17]
37 Verhoeijen, Bouwmeester, Zwaan	51	55.6		-0.05 [-0.24 , 0.15]
38 Was, Hirsch, Todaro, Romig	53.7	55.9		-0.02 [-0.19 , 0.15]

41 Meta-analytic effect for laboratory replications only				-0.04 [-0.07 , -0.01]
--	--	--	--	------------------------



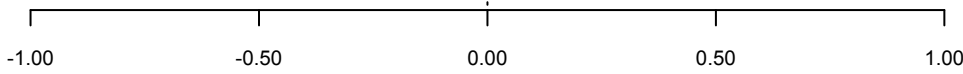
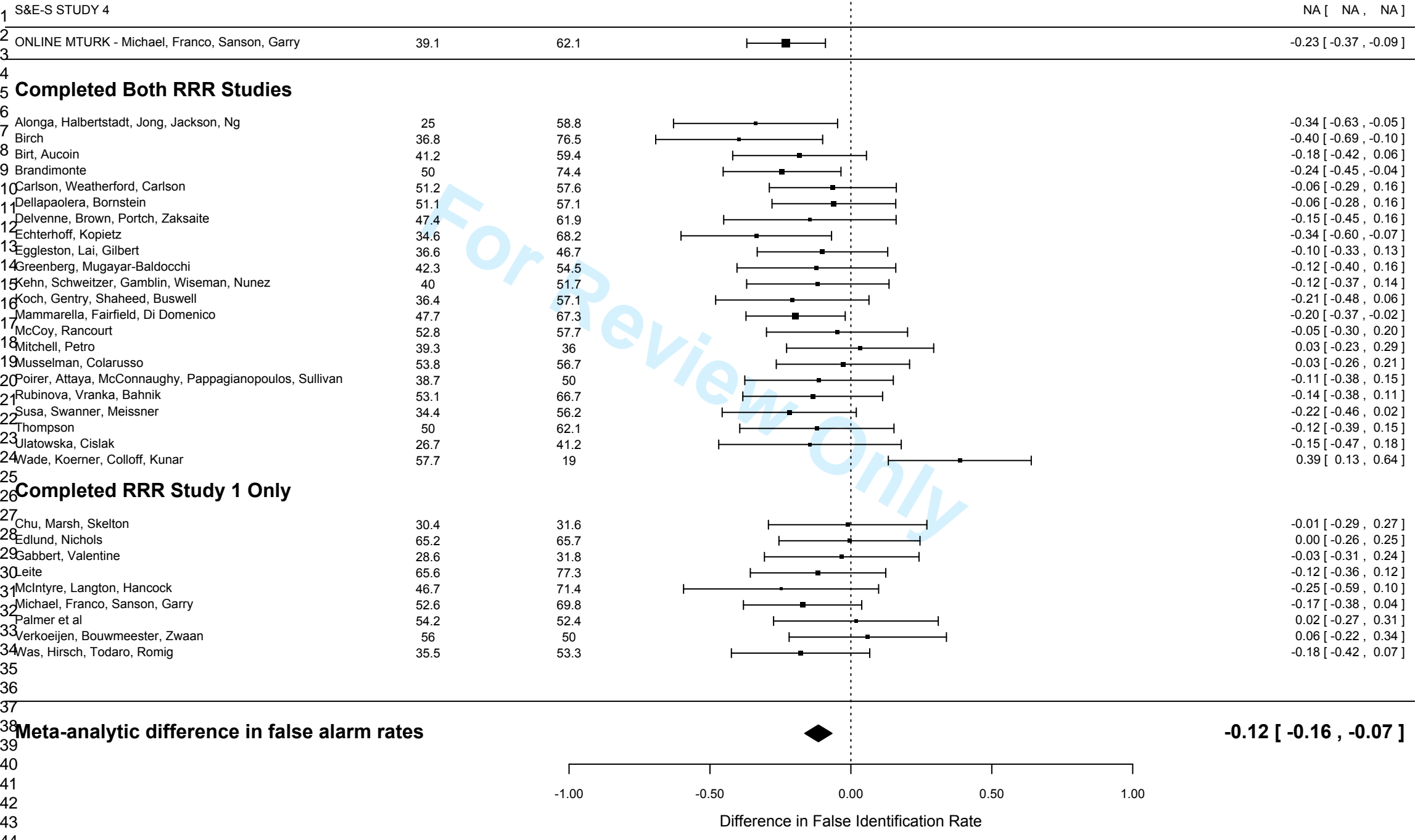


Perspectives on Psychological Science

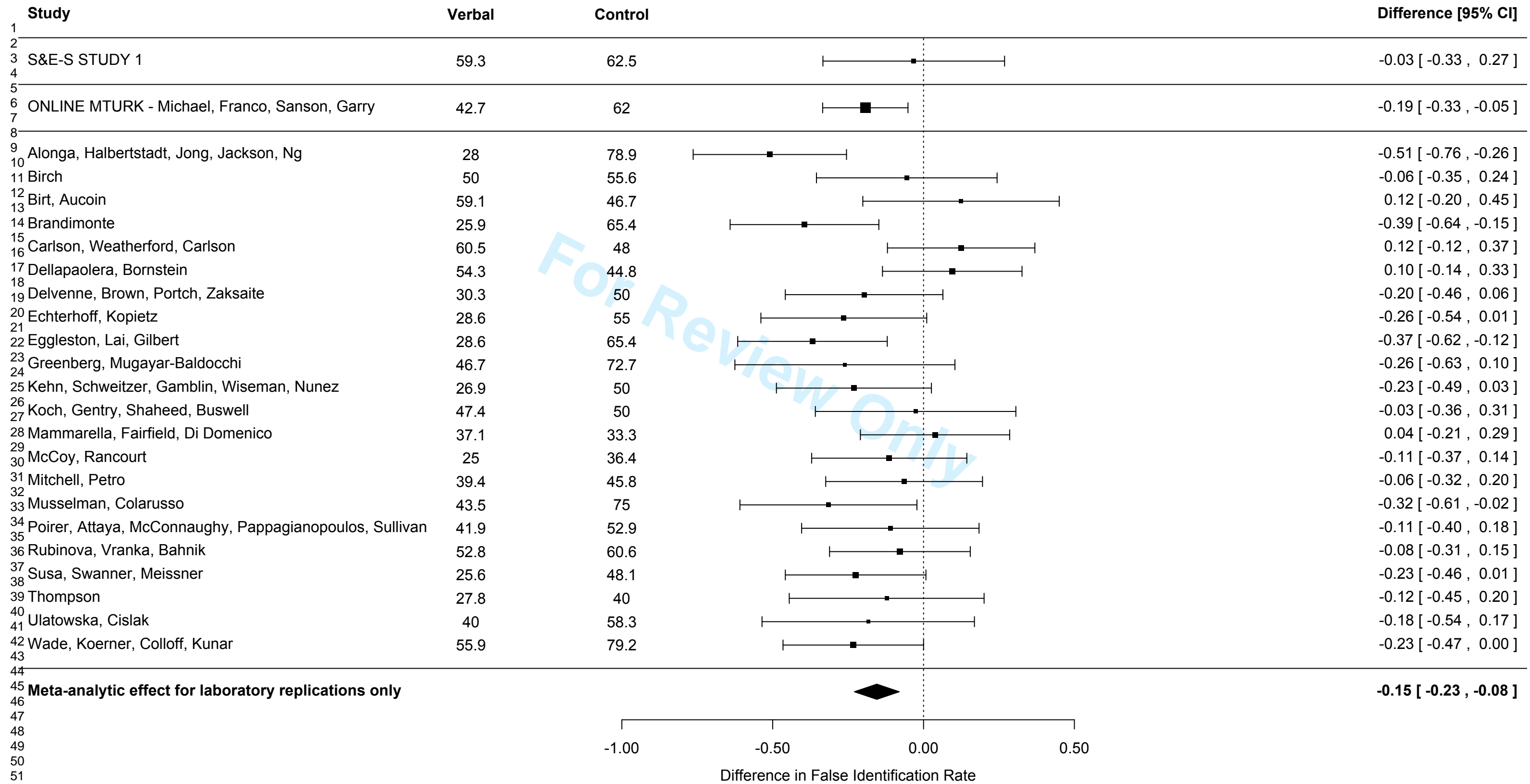
Study

Verbal

Control



Difference in False Identification Rate



Authors

ORIGINAL STUDY - Schooler & Engstler-Schooler (1990), Study 4

ONLINE STUDY (MTURK) - Robert B. Michael, Gregory Franco, Mevagh Sanson, Maryanne Garry

Victoria K. Alogna, Jamin Halberstadt, Jonathan Jong, Joshua C. Jackson, Cathy Ng

Stacy Birch

Angela R. Birt, Philip Aucoin

Maria A. Brandimonte

Curt Carlson, Dawn Weatherford, Maria Carlson

Kimberly S. Dellapaolera, Brian H. Bornstein

Jean-Francois Delvenne, Charity Brown, Emma Portch, Tara Zaksaitė

Gerald Echterhoff, René Kopietz

Casey Eggleston, Elizabeth Gilbert, Calvin Lai

Daniel L. Greenberg, Marino Mugayar-Baldocchi

Andre Kehn, Kimberly Schweitzer, Bradlee W. Gamblin, Kimberly Wiseman, Narina L. Nunez

Chris Koch, Remi Gentry, Jennifer Shaheed, Kelsi Buswell

Nicola Mammarella, Beth Fairfield, Alberto Di Domenico

Shannon McCoy, Arielle Rancourt

Abigail Mitchell, Marilyn Petro

Robin Musselman, Michael Colarusso

Christopher R. Poirier, Matthew K. Attaya, Kathleen A. McConaughy, Jessica E. Pappagianopoulos, Grif

Eva Rubínová, Marek Vranka, Štěpán Bahník

Kyle J. Susa, Jessica K. Swanner, Christian A. Meissner

W. Burt Thompson

Joanna Ulatowska, Aleksandra Cislak

Kimberley A. Wade, Ulrike Körner, Melissa, F. Colloff, Melina A. Kunar

Simon Chu, John E. Marsh, Faye Skelton

John E. Edlund, Austin Lee Nichols

Fiona Gabbert, Tim R. Valentine

Fabio P. Leite

Alexandra McIntyre, Stephen Langton, Peter J. B. Hancock

Robert B. Michael, Gregory Franco, Mevagh Sanson, Maryanne Garry

Matthew A. Palmer, Aaron Drummond, James D. Sauer, Daniel V. Zuj, Lauren Hall, Liam Satchell, Glenys

Verkoeijen, P. P. J. L., Bouwmeester, S., Zwaan, R. A.

Christopher A. Was, Dale Hirsch, Rachel Todaro, Connie Romig

Country	Language	Total N	Excluded - Race
USA	English	37	0
New Zealand	English	313	0
New Zealand	English	78	0
USA	English	69	0
Canada	English	53	0
Italy	Italian	70	0
USA	English	84	0
USA	English	86	0
United Kingdom	English	63	0
Germany	German	54	10
USA	English	93	2
USA	English	59	0
USA	English	73	4
USA	English	54	2
Italy	Italian	117	2
USA	English	75	2
USA	English	71	0
USA	English	65	0
USA	English	56	0
Czech Republic	Czech	80	0
USA	English	69	0
USA	English	66	2
Poland	Polish	59	0
United Kingdom	English	60	0
United Kingdom	English	50	0
USA	English	64	5
United Kingdom	English	83	6
USA	English	63	1
United Kingdom	English	54	2
New Zealand	English	184	0
USA	English	65	4
Netherlands	Dutch	56	0
USA	English	71	0

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Verbal Description Condition			
Excluded - Age	Excluded - Other	Total Included	Correct
0	0	37	18
0	109	204	112
17	5	56	32
13	0	56	37
2	0	51	17
0	0	70	34
0	9	75	32
12	0	74	29
2	2	59	40
0	3	41	15
0	10	81	40
9	0	50	24
6	0	63	33
4	0	48	26
0	5	110	45
11	0	62	26
9	0	62	34
14	0	51	12
0	1	55	24
0	12	68	36
13	1	55	23
11	3	50	28
0	4	55	40
0	0	60	34
0	0	50	27
9	0	51	28
22	0	55	34
7	3	52	20
0	0	52	37
67	30	87	49
5	0	56	32
1	4	51	26
4	0	67	36

False ID	Not Present	Total N	Excluded - Race	Excluded - Age
NA	NA	38	0	0
36	56	313	0	0
6	18	81	0	20
7	12	71	0	21
14	20	52	0	2
18	18	70	0	0
22	21	79	0	0
23	22	86	0	10
9	10	67	0	0
9	17	53	4	0
15	26	78	5	1
11	15	60	0	10
12	18	75	2	5
8	14	54	1	5
31	34	115	0	0
19	17	73	2	10
11	17	57	0	6
21	18	59	1	8
12	19	59	0	0
17	15	58	0	0
11	21	76	0	18
11	11	65	1	14
4	11	68	2	0
15	11	60	0	0
7	16	51	0	0
15	8	70	1	8
6	15	83	9	21
21	11	65	1	4
7	8	57	5	1
20	18	191	0	55
13	11	58	6	1
14	11	59	0	5
11	20	71	0	3

Control Condition				
Excluded - Other	Total Included	Correct	False ID	
0	38	27	NA	
127	186	91	59	
4	57	40	10	
0	50	33	13	
0	50	18	19	
0	70	27	32	
4	75	42	19	
0	76	41	20	
6	61	40	13	
3	46	24	15	
3	69	39	14	
0	50	28	12	
0	68	39	15	
0	48	22	16	
5	110	58	35	
0	61	35	15	
0	51	26	9	
0	50	20	17	
4	55	31	12	
8	50	23	18	
3	55	23	18	
0	50	21	18	
11	55	38	7	
0	60	39	4	
0	51	32	6	
0	61	26	23	
0	53	31	7	
10	50	28	17	
0	51	37	10	
38	98	55	30	
0	51	30	11	
0	54	30	12	
0	68	38	16	

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	Not Present
	NA
	36
	7
	4
	13
	11
	14
	15
	8
	7
	16
	10
	14
	12
	17
	11
	16
	13
	12
	9
	14
	11
	10
	17
	13
	12
	15
	5
	4
	13
	10
	12
	14

For Review Only

Lab

ORIGINAL STUDY - Schooler & Engstler-Schooler (1990), Study 1

ONLINE STUDY (MTURK) - Robert B. Michael, Gregory Franco, Mevagh Sanson, Maryanne Garry

Victoria K. Alogna, Jamin Halberstadt, Jonathan Jong, Joshua C. Jackson, Cathy Ng

Stacy Birch

Angela R. Birt, Philip Aucoin"

Maria A. Brandimonte

Curt Carlson, Dawn Weatherford, Maria Carlson

Kimberly S. Dellapaolera, Brian H. Bornstein

Jean-Francois Delvenne, Charity Brown, Emma Portch, Tara Zaksaitė

Gerald Echterhoff & René Kopietz

Casey Eggleston, Elizabeth Gilbert, Calvin Lai

Daniel L. Greenberg

Andre Kehn, Kimberly Schweitzer, Bradlee W. Gamblin, Kimberly Wiseman, Narina L. Nunez

Christopher Koch, Remi Gentry, Jennifer Shaheed, Kelsi Buswell

Nicola Mammarella, Beth Fairfield, Alberto Di Domenico

Shannon K. McCoy, Arielle Rancourt

Abigail Mitchell, Marilyn Petro

Robin Musselman, Michael Colarusso

Christopher R. Poirier, Matthew K. Attaya, Griffin A. Sullivan, Kathleen A. McConaughy, Jessica E. Papp

Eva Rubínová, Marek Vranka, Štěpán Bahník

Kyle J. Susa, Jessica K. Swanner, Christian A. Meissner

W. Burt Thompson

Joanna Ulatowska, Aleksandra Cislak

Kimberley A. Wade, Ulrike Körner, Melissa, F. Colloff, Melina A. Kunar

Country	Language	Total N	Excluded - Race
USA	English	44	0
New Zealand	English	302	0
New Zealand	English	70	0
USA	English	83	1
Canada	English	33	0
Italy	Italian	50	0
USA	English	81	4
USA	English	82	0
United Kingdom	English	48	0
Germany	German	58	10
USA	English	49	1
USA	English	37	0
USA	English	55	1
USA	English	35	1
Italy	Italian	50	0
USA	English	45	1
USA	English	57	0
USA	English	38	0
USA	English	46	0
Czech Republic	Czech	56	0
USA	English	53	0
USA	English	51	1
Poland	Polish	51	0
United Kingdom	English	61	0

Verbal Description Condition		
Excluded - Age	Excluded - Other	Total Included
0	0	44
0	98	204
20	0	50
25	4	53
3	0	30
0	0	50
0	2	75
10	0	72
1	1	46
0	5	43
0	5	43
7	0	30
5	0	49
3	1	30
0	0	50
3	0	41
9	2	46
8	0	30
1	1	44
0	4	52
3	0	50
12	0	38
0	4	47
0	1	60

only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	Correct	False ID	Not Present	Total N	Excluded - Race
	17	16	11	44	0
	94	47	63	313	0
	25	7	18	67	0
	27	13	13	73	1
	8	13	9	32	0
	23	7	20	50	0
	32	26	17	79	3
	26	25	21	82	0
	13	10	23	50	0
	15	8	20	66	15
	15	8	20	44	0
	15	7	8	38	0
	23	7	19	58	1
	11	9	10	32	1
	15	13	22	54	4
	13	7	21	44	1
	13	13	20	52	1
	7	10	13	40	0
	13	13	18	49	0
	16	19	17	54	0
	11	10	29	58	0
	20	5	13	51	1
	27	8	12	55	0
	26	19	15	60	0

only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Control Condition		
Excluded - Age	Excluded - Other	Total Included
0	0	44
0	130	183
17	0	50
14	4	54
1	0	31
0	0	50
0	1	75
15	0	67
1	1	48
0	5	46
0	3	41
7	1	30
7	0	50
1	0	30
0	0	50
2	0	41
2	3	46
10	0	30
0	8	41
0	4	50
8	0	50
11	0	39
0	8	47
0	0	60

only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	Correct	False ID	Not Present
1			
2			
3			
4			
5			
6	28	10	6
7	104	49	30
8			
9	31	15	4
10	36	10	8
11	16	7	8
12	24	17	9
13	50	12	13
14	38	13	16
15	26	11	11
16	26	11	9
17	15	17	9
18	19	8	3
19	24	13	13
20	14	8	8
21	26	8	16
22	19	8	14
23	22	11	13
24	14	12	4
25	24	9	8
26	17	20	13
27	23	13	14
28	24	6	9
29	35	7	5
30	36	19	5
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			

Only