



European Journal of Psychology and Educational Research

Volume 4, Issue 1, 13 - 24.

ISSN: 2589-949X

<https://www.ejper.com/>

The Potential Influence of To-Be-Forgotten Information on Educational Judgments

Sebastian Scholz*

University of Münster, GERMANY

Stephan Dutke 

University of Münster, GERMANY

Received: December 14, 2019 • Revised: March 4, 2021 • Accepted: May 15, 2021

Abstract: Teachers often face complex educational judgments and research has shown that teachers are prone to be influenced by unrelated information in their judgments and decisions. To investigate the influence of potential misinformation we employed a list-method directed forgetting paradigm and investigated a simulated judgment scenario, in which participants were asked to recommend a higher or lower school track for a fictitious elementary school child. Previous research using list-method directed forgetting revealed that participants can intentionally forget information but this information might still influence further judgments. In two experiments, data on recall performance, school track recommendation, and the evaluative impression of the target were analyzed to investigate whether participants were able to intentionally forget information and whether the to-be-forgotten information influenced later judgments. To-be-forgotten information was either presented before (Experiment 1) or following (Experiment 2) information instructed to be remembered. Both experiments revealed that participants did not forget information instructed to be forgotten and their judgments were not influenced by this information. Bayes factors spoke in favor of the null hypotheses, indicating that the influence of to-be-forgotten information on simulated school track recommendations is questionable. Our results revealed important boundary conditions of directed forgetting in applied contexts.

Keywords: *Directed forgetting, educational judgments, judgment bias, school track recommendation.*

To cite this article: Scholz, S., & Dutke, S. (2021). The potential influence of to-be-forgotten information on educational judgments. *European Journal of Psychology and Educational Research*, 4(1), 13-24. <https://doi.org/10.12973/ejper.4.1.13>

Introduction

Judgments and decisions are part of our everyday life but also pervasive in professional contexts like behavior in the court room (Englich & Mussweiler, 2001), sports (Kaya, 2014), and education, in which unbiased judgments are especially important. Teachers are specifically required to constantly process various incoming information from or about many students, weight them appropriately, and integrate them in adequate judgments and decisions. Consequently, research regarding biases in teachers decision-making have a long history (Bishop, 1976). It has been shown that teachers are influenced by irrelevant assessment information on, for example, in-class performance (Kaiser et al., 2013), academic achievement and academic competence (La Neal et al., 2003; McCombs & Gay, 1988; Parks & Kennedy, 2007), and test performance (Dünnebier et al., 2009). For instance, Dünnebier et al. (2009) revealed that grades given by experienced school teachers were systematically influenced by grades proposed randomly by other people. Moreover, Parks and Kennedy (2007) showed that students' ethnicity and physical attractiveness influenced teachers' ratings of students' academic and social competence. The lowest ratings were given to unattractive Black boys. Similarly, La Neal et al. (2003) reported that teachers rated students displaying the influence of Black culture in their movement styles as more aggressive, less academically competent, and more in need of special education services than other students. Next to ethnicity, teachers are also influenced by their own gender stereotypes and tend to perceive boys as having more developmental resources and talent in math classes than girls (Carlana, 2019; Tiedemann, 2002).

A key to many of these undesirable biases in judgments and decisions in educational contexts refers to the ability to neglect, ignore, or forget information that is actually unrelated to the subject of judgment. For example, intentionally forgetting randomly proposed grades would prevent them from influencing judgments on test performance (Dünnebier

* Corresponding author:

Sebastian Scholz, University of Münster, Institute for Psychology in Education, 48149 Münster, Germany. ✉ sebastian.scholz@uni-muenster.de



et al., 2009) and actively ignoring the students' ethnicity would prevent biases in judging academic competence (La Neal et al., 2003; Parks & Kennedy, 2007). Therefore, the idea of directed forgetting may play a key role in avoiding judgment and decision biases in educational contexts. A simple hypothesis would be that biases could be reduced if individuals are able to forget information identified or designated as invalid or misleading. The influence of to-be-forgotten information is investigated by research applying the list-method directed forgetting (LMDF) paradigm (R. A. Bjork, 1970). Studies in this domain typically reveal that people are able to forget previously presented information when instructed to do so (Anderson & Hanslmayr, 2014; Pastötter & Bäuml, 2010). In the LMDF task, participants are presented two lists of stimuli. Following the first list of stimuli (List 1), a cue is presented that signals to forget the presented list. Then, a second list is presented (List 2) followed by a cue signaling to remember this information and, after a delay, participants are instructed to recall all stimuli they are able to recall irrespective of the different memory cues, i.e. stimuli from both lists. Results typically show that information from List 2 is better remembered than information from List 1. These results are explained by either the retrieval inhibition account or the context-change account. The retrieval inhibition account assumes that stimuli from List 1 are inhibited by the forget cue and therefore are less accessible to the participant during recall (Anderson & Hanslmayr, 2014; Geiselman et al., 1983). The context-change account explains the forgetting of stimuli from List 1 with a change of the internal context following the forget cue that results in a mismatch between the context of List 1 and the context of the recall phase. In contrast, the context of List 2 and the context of the recall phase are the same and stimuli from List 2 are therefore easier recalled (Sahakyan & Kelley, 2002).

Interestingly however, directed forgetting effects in LMDF are limited to tests of free recall and are not generalizable to tests of recognition memory (Basden et al., 1993). This result suggests that to-be-forgotten information is still available in memory but simply not retrieved. Therefore, this information might continue to influence behavior – even more than information, which was not declared inadmissible or to be forgotten. This is since people cannot control the influence of information they are not aware of (E. L. Bjork & Bjork, 1996, 2003). For example, E. L. Bjork and Bjork (2003) cued participants to forget non-famous names and revealed that these names were later associated with more fame when they were instructed to be forgotten than when they were instructed to be remembered. Similarly, Golding and Hauselt (1994) revealed that participants were influenced in their evaluative judgments of a target by information instructed to be forgotten due to its incorrectness. Most interestingly, this information was later recalled even better than information not instructed to be forgotten. When these results would generalize to other real life settings this could have drastic consequences for complex educational judgment tasks. The first intuition that effective directed forgetting would help to avoid judgment and decision biases could turn into the opposite when forgotten but misleading information still affects judgment and decision processes. This question has not yet been investigated in an educational context. Therefore, in the present study, we chose an authentic educational judgment scenario which might be extremely prone to the influence of misinformation such as information that was later found to be false or irrelevant or information that stemmed from an unreliable source. An adaptive behavior in such situations would be to forget false, irrelevant, and unreliable information to reduce its potential influence on a decision. One authentic educational task that was previously shown to be influenced by misinformation is school track recommendation in German elementary schools.

During the fourth grade in Germany, students and their parents must decide whether the students continue their secondary education in a low, medium, or high school track (differing in length, academic level, and access to tertiary education). At the end of the fourth grade, teachers are required to recommend the appropriate school track for each of their students. Even though this recommendation is not legally binding in most German federal states, most parents tend to follow the teachers' recommendation (Stubbe & Bos, 2008). Teachers are advised to base their recommendation solely on grades in the most important school subjects, observed social behavior, and work behavior. However, studies showed that teacher recommendations only partially correspond to the results of standardized school performance tests. Instead, many teachers consider other factors, such as the social background of the student (Schneider, 2011) or the student's gender (Stubbe & Bos, 2008). Teachers must weight many factors when recommending a school track, and this judgment rarely involves feedback about the validity of the decision, even though this decision might severely influence the students' future. Thus, these task characteristics make school recommendation an appropriate scenario to investigate the potential influence of to-be-forgotten information on educational judgments. Therefore, we conducted two experiments in which school track recommendation was measured as one of the main dependent variables.

Experiment 1

In Experiment 1, a LMDF paradigm was applied in which participants received descriptions of a target's (a fictitious fourth-grade student) behavior. These descriptions were either consistent with a low or high school track recommendation, or were neutral with regard to school-relevant behavior. These descriptions were organized in two lists. One list consisted of high track and neutral descriptions and the other list consisted of low track and neutral descriptions. Following the presentation of the first list (List 1), participants were instructed to either forget the presented descriptions or to remember them. After presentation of the second list (List 2), participants were asked to give an evaluative impression of the target and to recommend one of the three school tracks (low, medium, or high). Finally, a free recall test for all descriptions was completed. Moreover, to receive a clearer picture of processes underlying the school track recommendation, we included a measure of confidence in the adequacy of the chosen recommendation.

Based on the presented research, two hypotheses were raised: (1) If participants show a directed forgetting effect, we expect reduced recall of List 1 compared to participants who were not instructed to forget List 1 and increased recall of List 2. (2) If to-be-forgotten information influences future judgments, List 1 descriptions will be used to evaluate the target person although these descriptions were not recalled.

Methodology

Participants

Overall, 68 student participants (16 males) across all faculties of the local university completed the experiment for either course credit or a monetary reward. Their age ranged from 18 to 32 years ($M = 23.0$, $SD = 2.8$). Participants reported no current neurological or psychological conditions and gave written informed consent. All procedures were performed in accordance with the Declaration of Helsinki and were approved by the local university's ethics board.

Stimuli

One-sentence descriptions of the target's behavior were created based on prior research regarding school track recommendations (van Ophuysen, 2006). The descriptions referred to academic performance, social behavior, and work behavior. The stimulus pool consisted of 60 sentences that were constructed to describe either a boy or a girl in the fourth grade. The target's gender was randomly varied between participants. The sentences were (a) neutral statements about the target (30 stimuli for each gender, e.g., "The girl/The boy likes to wear blue shoes"), (b) statements consistent with the higher school track (15 stimuli for each gender, e.g., "The boy/The girl is always prepared in class"), or (c) statements consistent with the lower school track (15 stimuli for each gender, e.g., "The girl/The boy has problems in following the coursework").

Design and Measures

Participants were randomly assigned to one of six experimental groups (Table 1). In two of the six groups (Groups R_{HAP} and R_{LAP} ; R = descriptions associated with a remember cue), participants received 15 randomly drawn neutral sentences and 15 sentences either consistent with higher academic performance (HAP) or consistent with lower academic performance (LAP) in random order. These one-list groups were included to serve as control groups for the HAP and LAP sentences and for later comparisons of the evaluative judgment and the school track recommendation.

In the other four groups, participants received two lists of information. Each list consisted of 15 HAP or 15 LAP sentences interspersed with 15 sentences drawn randomly from the set of 30 neutral sentences. The second list consisted of the 15 neutral statements and the HAP or LAP statements that were not part of the first list. Participants in the list sequence conditions $R_{HAP}R_{LAP}$ and $F_{HAP}R_{LAP}$ (F = descriptions associated with a forget cue) received HAP statements on List 1 and LAP statements on List 2. Analogously, participants in the list sequence conditions $R_{LAP}R_{HAP}$ and $F_{LAP}R_{HAP}$ first received LAP statements and then HAP statements.

Following the presentation of List 1, participants either received a remember cue or a forget cue. The remember cue signaled that the previous information was important and that other important information will follow. The forget cue signaled that the statements stemmed from a teacher who did not know the student but referred to another student and that the statements can thus be forgotten and should not be used to form an impression of the child. Participants were also told that important statements from a teacher who really knew the target person will be presented next.

Table 1: Overview of the experimental groups in Experiment 1

Group	List 1	Cue 1	List 2	Cue 2	n
R_{HAP}	HAP + N	Remember cue	-	-	10
R_{LAP}	LAP + N	Remember cue	-	-	10
$R_{HAP}R_{LAP}$	HAP + N	Remember cue	LAP + N	Remember cue	12
$R_{LAP}R_{HAP}$	LAP + N	Remember cue	HAP + N	Remember cue	12
$F_{HAP}R_{LAP}$	HAP + N	Forget cue	LAP + N	Remember cue	12
$F_{LAP}R_{HAP}$	LAP + N	Forget cue	HAP + N	Remember cue	12

Note. HAP = sentences consistent with higher academic performance, LAP = sentences consistent with lower academic performance, N = neutral sentences about the target.

Procedure

Participants were tested individually in the laboratory and sat in front of laptops running Windows 10. Stimuli were presented in black font in front of a gray background with PsychoPy3 (Version 3.2.3). Before the experiment started, one practice trial was accomplished to familiarize participants with the experimental procedure. No judgment or recall had to be completed in the practice phase. Following the practice trial, participants were encouraged to ask questions. In each trial a fixation cross was shown for 1 s in the center of the screen. The fixation cross was followed by the first sentence

describing the target. The sentence was presented for 5 s in the center of the screen before the next trial started. If participants received a memory cue, the cue was presented following the first list for 20 s in either green font color (remember cue for $R_{HAP}R_{LAP}$ and $R_{LAP}R_{HAP}$) or red font color (forget cue for Groups $F_{HAP}R_{LAP}$ and $F_{LAP}R_{HAP}$) in the center of the screen. Each list started and ended with two trials that were not part of the analyses and were only included to prevent primacy and recency effects.

Following the learning phase, a distraction task was included to prevent internal rehearsal. Participants marked differences in three puzzle pictures for 5 min.

After the distraction task, participants completed the judgment and recall phases using the EFS-Survey software (Questback GmbH, 2017). First, participants made a diagnostic judgment about the target person by rating the personal impression of the target on a likeability scale from 1 = very negative to 5 = very positive. Next, participants were asked to make a prognostic judgment by recommending one of the three school tracks (1 = lowest tier, 2 = middle tier, or 3 = highest tier). Three additional questions asked for how well the chosen school track would fit the child's academic performance, interest, and learning motivation on a scale from 1 = fits badly to 5 = fits very well. The scores of these three questions were summed, divided by the maximum value (15), and multiplied by 100, which served as a score of percentage of confidence in the individually chosen recommendation. No time limit was given for these prognostic questions.

For the free recall phase, participants were instructed to enter all statements they remembered in individually chosen order. They were asked to choose formulations as similar as possible to the wording of the original sentences. No time limit was given for the recall phase.

Statistical Analyses

We analyzed our free recall data in terms of costs and benefits (Sahakyan & Kelley, 2002). The cost of directed forgetting is the reduced recall of stimuli from List 1 in participants instructed to forget List 1 compared to participants instructed to remember List 1. The benefit of directed forgetting is the increase in recall of stimuli from List 2 for the group instructed to forget List 1 compared to participants instructed to remember List 1.

The research hypotheses were investigated using a statistical model including four dependent variables (evaluative judgment of the target person, school track recommendation, costs and benefits of directed forgetting) and two independent variables: experimental group and type of information (HAP or LAP). We used an alpha level of .05 for all our statistical analyses and non-significant results were analyzed using Bayes factors (BF_{01}) to receive an estimate of the evidence in favor of the null hypothesis (Rouder et al., 2009). BF_{01} indicates how much more likely a result is given that the null hypothesis is true.

To quantify the recall performance two trained but naïve raters scored the recalled statements. Statements were scored with either a score of 1 (statement was recalled) or a score of 0 (statement was missing). The interclass correlation (ICC) between the two judges was high ($ICC = .873$, $SD = 0.149$) and divergent ratings were settled after discussion.

School track recommendations and evaluative judgments were calculated with respect to the two one-list control groups: participants in $F_{HAP}R_{LAP}$ were instructed to forget HAP information and only remember LAP information. Thus, any differences in school track recommendations or evaluative judgments made by participants in R_{LAP} (who only received LAP information) would indicate persistent reliance on information which was instructed to be forgotten. Similarly, participants in $F_{LAP}R_{HAP}$ were instructed to forget LAP and remember HAP information. Thus, the differences in recommendations and evaluative judgments between participants in this group and participants in R_{HAP} (who only received HAP information) were also calculated.

Results

Costs and Benefits Analyses

To analyze the costs of directed forgetting, an ANOVA was calculated on the proportion of correct recall for List 1 depending on the experimental groups (averaged across type of information: HAP or LAP) and the type and order of information (Figure 1). The necessary conditions for conducting an ANOVA were given (e.g., Levene's test: $p = .653$). The analysis did neither reveal a significant effect for experimental group, $F(1, 44) = 0.12$, $p = .733$, $\eta_p^2 = .003$, $BF_{01} = 3.32$, nor for the type of information, $F(1, 44) = 0.91$, $p = .345$, $\eta_p^2 = .020$, $BF_{01} = 2.42$.

Benefits of directed forgetting were analyzed with a second ANOVA on the proportion of correct recall for List 2 in groups who received two lists during encoding. Participants in $R_{LAP}R_{HAP}$ and $R_{HAP}R_{LAP}$ (see Table 1) were instructed to remember Lists 1 and 2 and participants in $F_{LAP}R_{HAP}$ and $F_{HAP}R_{LAP}$ were instructed to forget List 1 and remember List 2. The necessary conditions for conducting an ANOVA were given (e.g., Levene's test: $p = .828$). The analysis did neither reveal a significant effect for experimental groups, $F(1, 44) = 1.12$, $p = .296$, $\eta_p^2 = .025$, $BF_{01} = 2.18$, nor for the type and order of information, $F(1, 44) = 0.65$, $p = .426$, $\eta_p^2 = .014$, $BF_{01} = 2.66$.

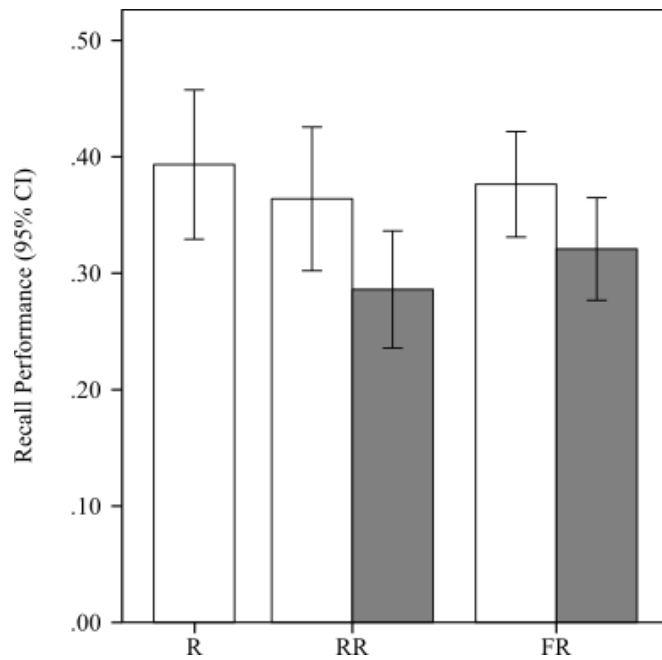


Figure 1. Recall results for Experiment 1. Performance in the free recall task averaged across the two subgroups of each experimental group ($R = R_{LAP}$ and R_{HAP} , $RR = R_{LAP}R_{HAP}$ and $R_{HAP}R_{LAP}$, $FR = F_{LAP}R_{HAP}$ and $F_{HAP}R_{LAP}$). In white, the average recall performance for List 1 and in gray the average recall performance for List 2. Error bars represent the 95% confidence interval (CI).

School Track Recommendation

The influence of to-be-forgotten statements on school track recommendation was calculated with respect to the one-list control groups. All participants who received only statements related to higher academic performance (R_{HAP}) recommended the highest tier. In contrast, 80% of participants who only received information consistent with lower academic performance (R_{LAP}) recommended the middle tier and the remaining 20% the lowest tier. The difference was significant, $\chi^2(2, 20) = 20.00$, $p < .001$ and represented a very strong effect, $CC = 1.00$, $p < .001$, $Cramér's V = 1.00$, $p < .001$, indicating that the information in the one-list control groups were indicative for the chosen school track recommendation and can be used as baseline measures.

We found no difference in school track recommendation between participants instructed to forget statements consistent with higher academic performance ($F_{HAP}R_{LAP}$) and participants who had not received these statements (R_{LAP}), $\chi^2(2, 22) = 1.39$, $p = .500$, $BF_{01} = 4.42$. Furthermore, these groups did not differ in the confidence in the adequacy of their recommendation, $t(20) = 1.85$, $p = .079$, $d = 0.77$. Recommendations were given with more confidence when statements were instructed to be forgotten ($M = 71.3\%$, $SD = 7.7\%$) than when they were not presented ($M = 62.7\%$, $SD = 13.4\%$).

All participants instructed to forget statements consistent with lower academic performance ($F_{LAP}R_{HAP}$) recommended the same school track as participants who had not received these statements (R_{HAP} ; 100% of these participants recommended the highest tier). These two groups also did not differ in the confidence associated with the school track recommendation $t(20) = 0.70$, $p = .491$, $d = 0.29$, $BF_{01} = 2.17$, indicating that to-be-forgotten information had no influence on school track recommendation.

Evaluative Judgment of the Target Person

The influence of to-be-forgotten information on evaluating the target person was calculated with reference to the one-list control groups. Participants who received only statements consistent with higher academic performance (R_{HAP}) rated the target with a likeability score of $M = 4.70$ ($SD = 0.48$), whereas participants who received only information consistent with lower academic performance (R_{LAP}) showed a likeability score of $M = 2.70$ ($SD = 0.48$). The means of these groups differed significantly, $t(18) = 9.26$, $p < .001$, $d = 4.14$, and demonstrate that the one-list control groups could be used as baseline measures for the evaluative judgments.

Participants instructed to forget statements associated with higher academic performance ($F_{HAP}R_{LAP}$) judged the target as likeable as participants who had not received statements associated with higher academic performance (R_{LAP}), $t(20) = -0.93$, $p = .366$, $d = 0.40$, $BF_{01} = 2.37$. Analogous to this pattern of results, participants instructed to forget statements associated with lower academic performance ($F_{LAP}R_{HAP}$) did not differ in their impression of the target from participants who never received these statements (R_{HAP}), $t(18) = -0.93$, $p = .366$, $d = 0.39$, $BF_{01} = 2.37$, meaning that participants were not influenced in their evaluative judgment by to-be-forgotten information.

Intermediate Discussion

Experiment 1 demonstrated that participants did not forget List 1 descriptions even when they were instructed to do so, despite previous research reporting that this cost of directed forgetting is a robust phenomenon (Sahakyan et al., 2013). In addition, no benefit for List 2 recall was evident for the group instructed to forget List 1. Thus, we found no support for directed forgetting in Experiment 1. Consistently, the results also revealed that descriptions from List 1 were not used to form the evaluative judgment of the target person or to recommend a school track for the target. Overall, participants seemed to be aware that the information on List 1 should not be used for further judgment processes.

What does this pattern of results mean for our goal to test the hypothesis that to-be-forgotten information can influence judgment processes (E. L. Bjork & Bjork, 1996) in this specific judgment scenario? Given that no evidence for directed forgetting was found in Experiment 1, the central condition for Bjork and Bjork's hypothesis, that to-be-forgotten information can have a larger influence on judgments than to-be-remembered information, was not given. We consistently failed to find evidence of the influence of List 1 information on our participants' judgments. Since the same information could be recalled, participants can be assumed to have metacognitive control over using or not using List 1 information. Thus, our results support the view that the participants controlled their use of List 1 information according to the instruction. They did not use this information for their judgments because they could recall this information – probably including knowledge about its source, the list they were instructed not to use. The central question resulting from these findings refers to the circumstances which are responsible for not finding a list-method directed forgetting effect. Experiment 2 was conducted to control for some factors that might be responsible for the current pattern of results and to investigate the influence of to-be-forgotten information in applied settings in more detail.

Experiment 2

Experiment 1 revealed no evidence for directed forgetting in the proposed judgment task. Therefore, the influence of to-be-forgotten information on later judgments and decisions (E. L. Bjork & Bjork, 1996) could not be evaluated. This pattern of results may be due to characteristics of the design we chose for Experiment 1. Since we were interested in an authentic application context with high external validity the design of Experiment slightly differed from traditional LMDF designs. One important difference between traditional LMDF tasks and Experiment 1 was that we did not control for output order of to-be-remembered and to-be-forgotten information. Previous research suggested that directed forgetting effects primarily emerge when to-be-remembered information is to be recalled before to-be-forgotten information is recalled. When the order of recall is not controlled in that way, inhibitory processes are supposed not to act on the to-be-forgotten information (Anderson, 2005). We refrained from controlling output order in the first experiment as this constraint would be untypical for authentic judgment scenarios like these. However, to explore to what extent this design characteristic affects directed forgetting in the present task, we controlled output order in Experiment 2 even though this tended to reduce external validity.

The second constraint of Experiment 1 lied in the fixed order of diagnostic and prognostic judgment tasks. A prognostic judgment (school track recommendation) was to be made following a diagnostic judgment (evaluative impression of the target). The explicit evaluative judgment of the target could have affected the later school track recommendation. To control for this potential influence, the order of these two judgment tasks was randomized among participants in Experiment 2.

Another critical design feature of Experiment 1 might be the fixed order of to-be-forgotten information (List 1) and to-be-remembered information (List 2). In a practical context like the present scenario, to-be-forgotten information could potentially also be encountered following information to-be-remembered. This sequence was not covered by traditional LMDF studies and was not included in the design of our first experiment. Therefore, Experiment 2 included an additional experimental group in which to-be-remembered information were presented on List 1 and to-be-forgotten information on List 2. This might be especially important since some studies revealed that presenting to-be-forgotten information on List 2 - thereby following to-be-remembered information - influenced later judgments in that this information was used to form an impression of a target (N. A. Wyer, 2010; R. S. Wyer & Unverzagt, 1985).

Finally, the design of Experiment 2 was simplified in four aspects. First, only one kind of information (statements consistent with lower academic performance, LAP) was instructed to be forgotten. Second, since the target's gender had no influence in Experiment 1 it was fixed for all participants. Third, more participants per group were included to increase statistical power of the results. Fourth, since recall performance was rather low in Experiment 1 we reduced the number of descriptions on each list.

Using these modifications in the experimental paradigm of Experiment 2, we investigated the same hypotheses as in Experiment 1.

Methodology

Participants

Overall, 186 student participants across all faculties of the local university completed the experiment for either course credit or for the possibility to take part in a raffle. No participant took part in Experiment 1. The data of all participants who obviously failed to follow the instructions and recalled less than 15% of presented information were removed from the data set. This resulting sample consisted of 126 participants (29 males) and ranged in age from 18 to 48 years ($M = 23.2$, $SD = 4.1$).

All procedures were performed in accordance with the Declaration of Helsinki and were approved by the local university's ethics board.

Stimuli

We used the same stimulus material as in Experiment 1. However, the stimulus pool consisted of only 40 sentences in total (10 HAP, 10 LAP, and 20 neutral sentences) and the target's gender was male in all descriptions and conditions.

Design and Measures

Participants were randomly assigned to one of four experimental groups. Dividing the 126 participants into four groups resulted in a statistical power of 80% to detect any medium effects (α level = .05). Participants in condition R ($n = 32$; Table 2) received only one list of descriptions with 10 neutral sentences and 10 sentences consistent with higher academic performance (HAP) in random order. This one-list control group was included for later comparisons of the evaluative judgment and the school track recommendation.

In the other three groups, participants received two lists of information. Participants in the list sequence conditions RF ($n = 33$) and RR ($n = 26$) received 10 HAP sentences interspersed with 10 neutral sentences on List 1. Following the presentation of that first list, participants received a remember cue (RR). Then, List 2 followed which consisted of 10 LAP statements and 10 other neutral sentences in random order. This second list was followed by a remember cue (in case of RR) or by a forget cue (in case of RF). In contrast to RF, participants in FR ($n = 35$) first received LAP plus neutral statements, followed by a forget cue, and then HAP plus neutral statements followed by a remember cue.

For all groups, the remember cue signaled that the previously presented statements were important for the following tasks. In addition, participants in RF and RR were informed that information from a second teacher of the target will follow next. In contrast to the remember cue, the forget cue signaled that the statements stemmed from a teacher who did not know the target but referred to another student and that the statements can thus be forgotten and are unimportant for the following tasks. In addition, participants in FR were also told that important statements from a teacher who really knew the target person will be presented next.

Table 2: Overview of the experimental groups in Experiment 2

Group	List 1	Cue 1	List 2	Cue 2	n
R	10 HAP + 10 N	Remember cue	-	-	32
RR	10 HAP + 10 N	Remember cue	10 LAP + 10 N	Remember cue	26
FR	10 LAP + 10 N	Forget cue	10 HAP + 10 N	Remember cue	35
RF	10 HAP + 10 N	Remember cue	10 LAP + 10 N	Forget cue	33

Note. HAP = sentences consistent with higher academic performance, LAP = sentences consistent with lower academic performance, N= neutral sentences about the target.

Procedure

Participants were tested online via Labvanced (<https://www.labvanced.com/>), which is a cloud-based solution for web-based psychological experiments.

Following the completion of a short demographic questionnaire, participants received information about the study and gave informed consent. Explanations emphasized that participants will receive information from a teacher about a boy in the fourth grade and have to form an impression of the boy and to answer questions about him. Participation was only possible using desktop computers or laptops. Stimuli were presented in black font in front of a gray background.

Before the learning phase of the experiment started, two practice trials were accomplished to familiarize participants with the experimental procedure. No judgment or recall had to be completed in the practice phase. In each trial of the experiment a fixation cross was shown for 1 sec in the center of the screen. The fixation cross was followed by a statement describing the target. The sentence was presented for 5 secs in the center of the screen before the next trial started. Memory cues were presented following the lists for 20 secs in either green (remember cue) or red (forget cue)

font color in the center of the screen. Each list started and ended with two trials that were not part of the analyses and were only included to prevent primacy and recency effects.

Following the learning phase, a distraction task was included to prevent internal rehearsal. In this phase, participants had to solve easy mathematical equations for 5 minutes.

After the distraction task, participants completed the same diagnostic and prognostic judgment tasks as in Experiment 1. However, the order of these two tasks was randomly permuted among participants to control for mutual interference between the two judgments.

For the free recall phase, participants were first asked to enter all statements they were instructed to remember in individually chosen order. Then, on a new page, participants in FR and RF entered all statements instructed to forget. Participants were asked to choose formulations as similar as possible to the wording of the original sentences. No time limit was given for the recall phase.

Statistical Analyses

Analogous to Experiment 1, we calculated costs and benefits of forgetting for FR and RF, used an alpha level of .05 for all our statistical analyses, and presented Bayes factors for non-significant results.

Due to the high ICC in Experiment 1 and using the same stimulus material, only one trained rater who was naïve to the experimental conditions scored the recalled statements.

School track recommendations and evaluative judgments were calculated again with respect to the one-list control group (R). Any differences in school track recommendations or evaluative judgments made by participants in FR or RF compared to the control condition R would indicate persistent reliance on information that was instructed to be forgotten.

Results

Costs and Benefits Analyses

To analyze the costs of directed forgetting in FR, a t-test was calculated on the proportion of correct recall for List 1 depending on the experimental groups (FR and RR; Figure 2). The test revealed no difference in recall performance between the groups, $t(59) = -1.11$, $p = .271$, $d = .14$, $BF_{01} = 2.94$. The cost of forgetting for RF was calculated with a t-test on the proportion of recall of List 2. The analysis detected no difference between RF and RR, $t(57) = 0.50$, $p = .622$, $d = .15$, $BF_{01} = 4.54$.

Benefits of forgetting for FR were calculated with a t-test on the proportion of correct recall of List 2. The analysis showed no difference in recall between FR and RR, $t(59) = -0.9$, $p = .926$, $d = 0.13$, $BF_{01} = 5.11$. Finally, the benefits of forgetting in our condition RF were calculated on the proportion of recall of List 1 between RF and RR. The analysis demonstrated no difference between the groups, $t(57) = -0.06$, $p = .956$, $d = 0.15$, $BF_{01} = 5.07$.

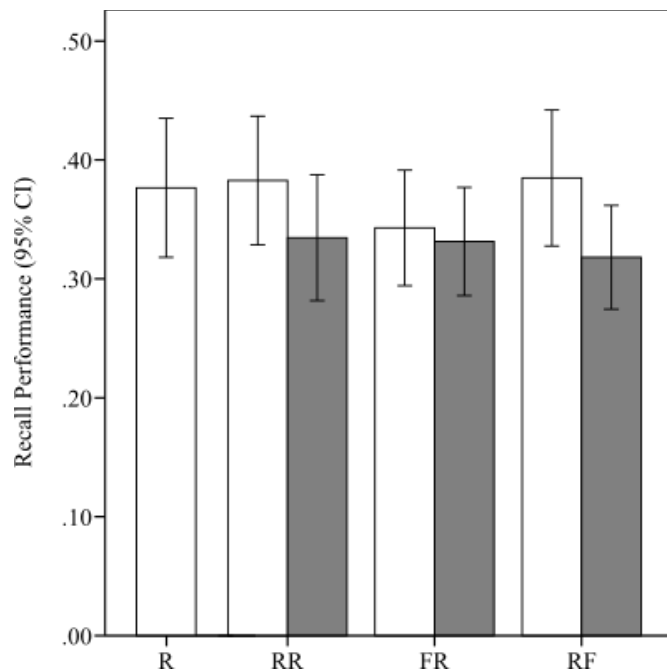


Figure 2. Recall results for Experiment 2. Performance in the free recall task for the experimental groups. In white, the average recall performance for List 1 and in gray the average recall performance for List 2. Error bars represent the 95% CI.

School Track Recommendation

Thirty-one participants (96.9%) who only received HAP information recommended the highest tier (R) and one participant the middle tier. They had an average rating of confidence in the adequacy of their recommendation of 81.3% ($SD = 14.0\%$). Seventeen participants (65.4%) who were instructed to remember LAP and HAP information (RR) recommended the highest tier and the remaining nine participants the middle tier. These participants had an average confidence rating of 74.1% ($SD = 7.8\%$).

Thirty-two participants (91.4%) who were instructed to forget LAP information on List 1 (FR) recommended the highest tier, two participants the middle tier, and one participant the lowest tier. There was no difference in school track recommendation between R and FR, $\chi^2(2, 67) = 1.22, p = .544, BF_{01} = 32.64$. The confidence in the adequacy of the recommendation also did not differ between these groups (R: $M = 81.3\%, SD = 14.0\%$; FR: $M = 86.1\%, SD = 9.9\%$), $t(65) = 1.65, p = .104, d = 0.40, BF_{01} = 1.27$.

Thirty-one participants (93.9%) who were instructed to forget LAP statements on List 2 (RF) recommended the highest tier and two participants the middle tier. There was no difference between RF and R, $\chi^2(1, 65) = 0.32, p = .573, BF_{01} = 6.58$. However, participants in RF ($M = 88.7\%, SD = 9.6\%$) differed in confidence regarding the recommendation from participants in R ($M = 81.3\%, SD = 14.0\%$), $t(63) = 2.50, p = .015, d = 0.62$.

Evaluative Judgment of the Target Person

The influence of to-be-forgotten information on the evaluative judgment of the described student was calculated for the one-list control groups. Participants who received only HAP statements (R) rated the target with a likeability score of $M = 4.66 (SD = 0.48)$. Participants instructed to remember HAP and LAP statements (RR) showed a likeability score of $M = 3.85 (SD = 0.37)$.

Participants instructed to forget LAP statements on List 1 (FR; $M = 4.60, SD = 0.55$) judged the target as likeable as participants who had not received statements associated with lower academic performance (R), $t(65) = -0.44, p = .660, d = -0.12, BF_{01} = 4.94$. Similarly, there was no difference in evaluative judgment between participants instructed to forget LAP statements on List 2 (RF, $M = 4.58, SD = 0.56$) and participants in R, $t(63) = -0.62, p = .538, d = -0.15, BF_{01} = 4.47$.

Intermediate Discussion

Experiment 2 replicated the results of Experiment 1 although we controlled for output order effects, the order of the two judgment tasks, simplified the design and the learning task, and increased statistical power. Nevertheless, we did not find any evidence of directed forgetting in the chosen context and no influence of this information on prognostic (school track recommendation) or diagnostic judgments (likeability rating). This pattern of results was also found for to-be-forgotten information on List 2. Bayes factors offered some support for the null hypotheses, especially regarding benefits of forgetting and school track recommendations. Consequently, to-be-forgotten information seems not to play a role in the chosen judgment scenario.

Discussion

In two experiments we did not find evidence for directed forgetting in an applied educational judgment context and since no directed forgetting was found, the prerequisite for an influence of this information on later judgments was not given (E. L. Bjork & Bjork, 2003). Participants instructed to forget some information recommended the same school track as participants who never saw the to-be-forgotten information. There was also no difference in evaluative judgment between these groups. Bayes factors offered some support for the finding that to-be-forgotten information had no influence on later diagnostic or prognostic judgments.

Given the consistent pattern of results in the two experiments described here, it is important to state which features differentiated our experiments from studies applying the classical LMDF paradigm (Abel & Bäuml, 2019; Sahakyan & Kelley, 2002). In these studies, the to-be-forgotten items are independent of each other and have no common context and thus are less likely to be integrated into a common mental model. In contrast, the to-be-forgotten stimuli in our research could have been easily thematically integrated because all information referred to the same person and therefore might have become less forgettable. Further support for the influence of the integration of information on directed forgetting was already presented by research that contrasted to-be-forgotten sentences that could be thematically integrated with to-be-forgotten sentences that could not be integrated (Delaney et al., 2009). These authors found directed forgetting costs only for the latter category of sentences and attributed this result to the integration of information. Participants in the experiments described in the present study were instructed to integrate information in order to form an impression of a target and to recommend a school track. This instruction could have prevented directed forgetting and an influence of to-be-forgotten information on later judgments. Other research however, found an influence of to-be-forgotten information on later judgments, even when information was instructed to be integrated (R. S. Wyer & Unverzagt, 1985). In this research however, to-be-forgotten information only had an influence when it was presented on List 2. The reason might be that participants in the latter condition had already formed a stable impression of the target based on the information on List 1 (Srull & Wyer, 1989). Thus, forgetting irrelevant information presented later might be easier

because it is not connected to the previously formed impression. Our Experiment 2 however, did not present any evidence for an influence of to-be-forgotten information even when the to-be-forgotten information was presented on List 2. Thus, we conclude that the sequence of to-be-remembered and to-be-forgotten information, in the present context, was not a critical condition for the influence of to-be-forgotten information on judgments.

Another aspect which differentiated the stimuli used in the current investigation from stimuli in traditional LMDF studies is the evaluative distinctiveness of List 1 and List 2. Even though we included neutral descriptions to the lists to decrease their distinctiveness, the source of the information (whether it should have been remembered or forgotten) was rather easy to recall. This might have decreased the likelihood to find directed forgetting and an influence of the to-be-forgotten stimuli because participants were able to differentiate which information to use or to avoid for their judgments and hence to metacognitively control the influence of to-be-forgotten information. However, evaluatively distinct positive and negative descriptions were already used in other research in which an enduring influence of to-be-forgotten information was found (Golding & Hauselt, 1994). Thus, we conclude that the distinctiveness of the lists was not crucial for the lack of directed forgetting in the present experiments.

Perhaps, LMDF effects are limited to strictly controlled contexts with thematically independent pieces of information. If this assumption is correct, judgments and decisions in the context of school track recommendations are potentially less prone to the influence of to-be-forgotten information than initially hypothesized. Similar practical judgment tasks as employed in our research were already investigated in the context of the continued influence effect (Johnson & Seifert, 1994, 1998; Lewandowsky et al., 2012), in which previously given information is later corrected but participants still use the misinformation although it was corrected. Complex educational judgments like school track recommendations could also be investigated using this experimental task, which could potentially reveal in more detail how misinformation influences judgments in educational contexts and, perhaps, even suggest how this influence could be minimized. Given that judgments and decision-making are central topics in educational contexts, prospective teachers should be educated about these processes and potential biases thereof.

With regard to theory, the present experiments explored some boundary conditions of experimental paradigms originating from the directed forgetting context. If the validity of the hypothesized mechanisms in LMDF depends on the structure of the material (independent pieces of information that could not easily be integrated), the application of LMDF in other applied contexts is not expected to be successful since almost all decisions in the real world are based on coherent information. This is particularly true for judgment and decision situations in the educational context. Thus, the usefulness of applying LMDF to real life situations remains questionable.

The current results regarding school track recommendation can also inform us about other practical contexts in education in which information should be better neglected, ignored, or forgotten. The influence of ethnicity and gender on judgments of academic competence and performance (La Neal et al., 2003; Parks & Kennedy, 2007; Tiedemann, 2002) are not decreased by simply instructing teachers to forget the unrelated information. However, previous research revealed that repeated practice of counterstereotypes (e.g., *Girls are more mathematically gifted than boys*) and increasing self-awareness for stereotypical beliefs can potentially help to decrease the influence of unrelated information on educational judgments and decisions (Burns et al., 2017). Future research could potentially investigate if repeated instruction to forget stereotypes can also be used to counter stereotypical beliefs in education.

Limitations and Recommendations

In favor of the theoretical approach it should be considered that generalizing the present experiments to real classroom situations is somewhat limited since we did not include experienced elementary school teachers who regularly face school track recommendations. Based on previous research, however, we expect that directed forgetting effects would be even lower in experienced teachers since these participants were shown to integrate even more information for recommendations than control participants (van Ophuysen, 2006). However, cognitive processes related to school track recommendations in experienced teachers could also differ in other aspects that limit the generalization of current results to the applied context. Therefore, future research could potentially investigate the direct or indirect influence of to-be-forgotten information on experienced teachers.

Practitioners in educational contexts, especially elementary school teachers, should be trained to increase their awareness about factors that have the potential to influence complex judgments and decisions in their professional contexts. As previous research reported, many factors such as gender, ethnicity, or economic status can influence unrelated judgments and thus drastically influence the future of the target person.

Conclusions

The aim of the current study was to investigate the influence of to-be-forgotten information in an educational context on measures of evaluative impression and school track recommendation. We revealed that neither presenting to-be-forgotten information before (Experiment 1) nor after (Experiment 2) to-be-remembered information lead to an influence of the to-be-forgotten information on the dependent variables. This result was interpreted to represent boundary conditions of list-method directed forgetting in applied contexts, in which information needs to be integrated

to form an impressions of a target and to obtain an informed judgment. Interestingly, our dependent variable school track recommendation was previously shown to be prone to be influenced by many different factors like gender of the child or the economic background of the parents (Stubbe & Bos, 2008). However, the influence of to-be-forgotten information does not seem to be one of those influencing factors.

Authorship Contribution Statement

Sebastian Scholz: Conceptualization, design, analysis, writing. Stephan Dutke: Editing/reviewing, supervision.

References

- Abel, M., & Bäuml, K.-H. T. (2019). List-method directed forgetting after prolonged retention interval: Further challenges to contemporary accounts. *Journal of Memory and Language, 106*, 18–28. <https://doi.org/10.1016/j.jml.2019.02.002>
- Anderson, M. C. (2005). The role of inhibitory control in forgetting unwanted memories: A consideration of three methods. In N. Ohta, C. M. MacLeod & B. Utzl (Eds.), *Dynamic Cognitive Processes* (pp. 159–189). Springer. https://doi.org/10.1007/4-431-27431-6_8
- Anderson, M. C., & Hanslmayr, S. (2014). Neural mechanisms of motivated forgetting. *Trends in Cognitive Sciences, 18*(6), 279–292. <https://doi.org/10.1016/j.tics.2014.03.002>
- Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(3), 603–616. <https://doi.org/10.1037/0278-7393.19.3.603>
- Bishop, A. J. (1976). Decision-making, the intervening variable. *Educational Studies in Mathematics, 7*(1/2), 41–47. <https://doi.org/10.1007/BF00144357>
- Bjork, E. L., & Bjork, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness and Cognition, 5*(1–2), 176–196. <https://doi.org/10.1006/ccog.1996.0011>
- Bjork, E. L., & Bjork, R. A. (2003). Intentional forgetting can increase, not decrease, residual influences of to-be-forgotten information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 524–531. <https://doi.org/10.1037/0278-7393.29.4.524>
- Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior, 9*(3), 255–268. [https://doi.org/10.1016/S0022-5371\(70\)80059-7](https://doi.org/10.1016/S0022-5371(70)80059-7)
- Burns, M. D., Monteith, M. J., & Parker, L. R. (2017). Training away bias: The differential effects of counterstereotype training and self-regulation on stereotype activation and application. *Journal of Experimental Social Psychology, 73*, 97–110. <https://doi.org/10.1016/j.jesp.2017.06.003>
- Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics, 134*(3), 1163–1224. <https://doi.org/10.1093/qje/qjz008>
- Delaney, P. F., Nghiem, K. N., & Waldum, E. R. (2009). The selective directed forgetting effect: Can people forget only part of a text? *Quarterly Journal of Experimental Psychology, 62*(8), 1542–1550. <https://doi.org/10.1080/17470210902770049>
- Dünnebier, K., Gräsel, C., & Krolak-Schwerdt, S. (2009). Urteilsverzerrungen in der schulischen Leistungsbeurteilung: Eine experimentelle Studie zu Ankereffekten [Biases in teacher's assessments of student performance: An experimental study of anchoring effects]. *German Journal of Educational Psychology/ Zeitschrift Für Pädagogische Psychologie, 23*(34), 187–195. <https://doi.org/10.1024/1010-0652.23.34.187>
- Englich, B., & Mussweiler, T. (2001). Sentencing under uncertainty: Anchoring effects in the courtroom. *Journal of Applied Social Psychology, 31*(7), 1535–1551. <https://doi.org/10.1111/j.1559-1816.2001.tb02687.x>
- Geiselman, R. E., Bjork, R. A., & Fishman, D. L. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. *Journal of Experimental Psychology: General, 112*(1), 58–72. <https://doi.org/10.1037/0096-3445.112.1.58>
- Golding, J. M., & Hauselt, J. (1994). When instructions to forget become instructions to remember. *Personality and Social Psychology Bulletin, 20*(2), 178–183. <https://doi.org/10.1177/0146167294202005>
- Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(6), 1420–1436. <https://doi.org/10.1037/0278-7393.20.6.1420>
- Johnson, H. M., & Seifert, C. M. (1998). Updating accounts following a correction of misinformation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(6), 1483–1494. <https://doi.org/10.1037/0278-7393.24.6.1483>

- Kaiser, J., Retelsdorf, J., Südkamp, A., & Möller, J. (2013). Achievement and engagement: How student characteristics influence teacher judgments. *Learning and Instruction, 28*, 73–84. <https://doi.org/10.1016/j.learninstruc.2013.06.001>
- Kaya, A. (2014). Decision making by coaches and athletes in sport. *Procedia - Social and Behavioral Sciences, 152*, 333–338. <https://doi.org/10.1016/j.sbspro.2014.09.205>
- La Neal, V. I., McCray, A. D., Webb-Johnson, G., & Bridgest, S. T. (2003). The effects of African American movement styles on teachers' perceptions and reactions. *The Journal of Special Education, 37*(1), 49–57. <https://doi.org/10.1177/00224669030370010501>
- Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest, 13*(3), 106–131. <https://doi.org/10.1177/1529100612451018>
- McCombs, R. C., & Gay, J. (1988). Effects of race, class, and IQ information on judgments of parochial grade school teachers. *The Journal of Social Psychology, 128*(5), 647–652. <https://doi.org/10.1080/00224545.1988.9922918>
- Parks, F. R., & Kennedy, J. H. (2007). The Impact of race, physical attractiveness, and gender on education majors' and teachers' perceptions of student competence. *Journal of Black Studies, 37*(6), 936–943. <https://doi.org/10.1177/0021934705285955>
- Pastötter, B., & Bäuml, K.-H. T. (2010). Amount of postcue encoding predicts amount of directed forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(1), 54–65. <https://doi.org/10.1037/a0017406>
- Questback GmbH. (2017). *EFS Survey* [Computer software]. <https://www.unipark.de/>
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225–237. <https://doi.org/10.3758/PBR.16.2.225>
- Sahakyan, L., Delaney, P. F., Foster, N. L., & Abushanab, B. (2013). List-method directed forgetting in cognitive and clinical research. In B. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 59, pp. 131–189). Elsevier. <https://doi.org/10.1016/B978-0-12-407187-2.00004-6>
- Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(6), 1064–1072. <https://doi.org/10.1037/0278-7393.28.6.1064>
- Schneider, T. (2011). Die Bedeutung der sozialen Herkunft und des Migrationshintergrundes für Lehrerurteile am Beispiel der Grundschulempfehlung [The importance of social and migrant backgrounds for school recommendations in Germany at the end of primary school]. *German Journal for Educational Science/ Zeitschrift Für Erziehungswissenschaft, 14*(3), 371–396. <https://doi.org/10.1007/s11618-011-0221-4>
- Srull, T. K., & Wyer, R. S. (1989). Person memory and judgment. *Psychological Review, 96*(1), 58–83. <https://doi.org/10.1037/0033-295X.96.1.58>
- Stubbe, T. C., & Bos, W. (2008). Schullaufbahneempfehlungen von Lehrkräften und Schullaufbahnentscheidungen von Eltern am Ende der vierten Jahrgangsstufe [Teacher's school track recommendations and parents' school track decisions at the end of primary school]. *Empirical Pedagogy/ Empirische Pädagogik, 22*(1), 49–63.
- Tiedemann, J. (2002). Teachers' gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educational Studies in Mathematics, 50*(1), 49–62. <https://doi.org/10.1023/A:1020518104346>
- van Ophuysen, S. (2006). Vergleich diagnostischer Entscheidungen von Novizen und Experten am Beispiel der Schullaufbahneempfehlung [Comparison of diagnostic decisions between novices and experts: The example of school career recommendation]. *German Journal for Developmental Psychology and Educational Psychology/ Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie, 38*(4), 154–161. <https://doi.org/10.1026/0049-8637.38.4.154>
- Wyer, N. A. (2010). You never get a second chance to make a first (implicit) impression: The role of elaboration in the formation and revision of implicit impressions. *Social Cognition, 28*(1), 1–19. <https://doi.org/10.1521/soco.2010.28.1.1>
- Wyer, R. S., & Unverzagt, W. H. (1985). Effects of instructions to disregard information on its subsequent recall and use in making judgments. *Journal of Personality and Social Psychology, 48*(3), 533–549. <https://doi.org/10.1037/0022-3514.48.3.533>