Running head: INTERVIEWING ENHANCES SUGGESTIBILITY

Immediate Interviewing Increases Children's Suggestibility in the Short-term, but not in the

Long-Term

Henry Otgaar[1,2,3], Jason C.K. Chan[4], Bruna Calado[1], & David La Rooy[5]

[1]Maastricht University, The Netherlands

[2]Faculty of Psychology, Universitas Indonesia

[3]City, University of London, UK

[4]Iowa State University, USA

[5]Royal Holloway, University of London

*In Press, Legal and Criminological Psychology*

Correspondence should be sent to: Henry Otgaar, Henry.Otgaar@maastrichtuniversity.nl,

Faculty of Psychology and Neuroscience, section Forensic Psychology, Maastricht University

## Abstract

**Purpose**: Children sometimes receive misinformation after being formally interviewed about their experiences in cases of suspected abuse. Following decades of research, many guidelines have been produced for interviewers so they can obtain reliable statements in children, like for example the NICHD Protocol. One might expect that completing an early interview following research-based guidelines might guard against the incorporation of misinformation encountered later. The goal of the current experiments was to examine whether following research-based guidelines such as the NICHD Protocol might protect child witnesses against follow-up 'misinformation' or make them more vulnerable to misinformation. This increased vulnerability to misinformation has been referred to as *retrieval-enhanced suggestibility*.

**Methods**: In two experiments, children viewed a video and half of them were interviewed using the NICHD Protocol while the other half were not interviewed. The children received misinformation and a final memory test either immediately after being interviewed (Experiment 1) or one week later (Experiment 2).

**Results**: Retrieval-enhanced suggestibility was observed when misinformation was presented immediately but not when it was provided after one week.

**Conclusions**: The current experiments indicate that a well-established interview protocol can, under some circumstances, amplify levels of suggestibility in children.

**Immediate Interviewing Increases Children's Suggestibility in the Short-term, but not in the Long-Term**

A great deal of experimentation has been conducted on best practice guidelines for interviewing children (Lamb, Orbach, Hershkowitz, Esplin, & Horowitz, 2007). The impetus of this empirical endeavour is straightforward. Children are increasingly involved in legal situations where they have to report on what they witnessed or experienced (such as sexual abuse). Because children's reporting is often poor, and it is thought that they are particularly susceptible at making memory errors (Otgaar, Howe, Brackmann, & Smeets, 2016), extra assistance is needed to facilitate their memory reports. Many research-based protocols have been produced in recent decades with the NICHD Protocol (National Institute of Child Health and Human Development Protocol; La Rooy et al., 2015), in particular, receiving substantial scientific validation. Despite the extensive corpus of literature on the NICHD Protocol, it is not yet known whether this Protocol and others like it can 'immunize' children from accepting later misinformation about their experiences. This issue lies at the heart of the current experiments.

In police investigative settings, it is frequently advised that children should be interviewed soon after the purported experience, as a long delay between the event and interview might adversely affect the quality of the child's statement (Pipe, Lamb, Orbach, & Cederborg, 2007). Crucially, however, even after a well-conducted interview such as the NICHD protocol, children are sometimes exposed to misinformation from their interactions with parents or therapists (Otgaar, de Ruiter, Howe, Hoetmer, van Reekum, 2017). An important question is whether a high quality early interview protects children from falling prey to later misinformation in the form of suggestions.

## The NICHD Protocol

The fundamental tenet of the NICHD Protocol is that the protocol is constructed based on research regarding the development of memory and how children talk about abusive experiences (La Rooy et al., 2015; Lamb, Orbach, Hershkowitz, Esplin, & Horowitz, 2007). The underlying idea behind this protocol is that it takes into account developmental research concerning children's capabilities, to increase the chance that children will give an accurate and detailed account of their experience and to minimize the possibility for suggestion within an interview (La Rooy et al., 2015).  A core issue here is that the protocol emphasizes the use of open prompts (e.g., invitation questions: "Tell me what happened"). A plethora of research has found that open prompts can lead to detailed and accurate statements in children (Saywitz, Lyon, & Goodman, in press). The NICHD Protocol encourages interviewers to ask open prompts continuously until the child's statement is virtually exhausted. Interviewers accomplish this by starting with a general invitation ("Tell me what happened") and then follow-up with, for example, cued invitations (e.g., "You mentioned that it happened at your home. Tell me more about that"). Only after this, interviewers can engage in more closed questions -based on the information previously provided by the child to the open questions –   (e.g., directive questions: "Where were your clothes?" or option-posing questions: "Did he touch you in your bedroom or in his bedroom?") which intends to target recognition memory.

This emphasis on asking first open-ended questions and then resort to more focused questions is accomplished by an interview which commences with rapport building and having children practice retrieving an unrelated event before they concentrate on the experience in question. This latter phase is referred to as the substantive phase. During this phase, the child is asked about why he/she is at the interview. More importantly, by make continuous use of open

prompts such as invitations, the child is invited to retrieve what ostensibly happened. Studies have shown that the NICHD Protocol is effective in generating highly detailed and accurate statements in children with and without a history of maltreatment (Lamb et al., 2007). Because the protocol leads to detailed memories in children, one might assume that retrieving such detailed memories can act as a protective shield when misinformation is encountered on another occasion (e.g., by parents, teachers; Huff, Weinsheimer, & Bodner, 2016) which could justify the usual procedure in police investigative settings to interview children as soon as possible. Nevertheless, it is not uncommon for children to talk about their experiences after a well-conducted interview, and such follow-up occasions might contain misinformation.

Children who are interviewed by the NICHD Protocol are encouraged to retrieve relevant memories concerning the experience during the substantive phase. It is this act of retrieval during the substantive phase that is of interest for the current experiment. Interestingly, there is recent research on whether the act of retrieval might affect one's susceptibility to accepting suggestive information. Much of this research has focused on the so-called retrieval practice effect and its impact on suggestibility. The retrieval practice effect refers to the finding that people have better memory for information when they have previously been tested on it than when they have not (Roediger & Karpicke, 2006). Recently, there is experimental work revealing that, counterintuitively, retrieval might enhance eyewitnesses' susceptibility to subsequently presented misleading information. This phenomenon has been referred to as retrieval-enhanced suggestibility (Chan, Thomas, & Bulevich, 2009; Thomas, Bulevich, & Chan, 2010).

### Retrieval-enhanced Suggestibility

Retrieval-enhanced suggestibility refers to the memory phenomenon that participants are more likely to endorse suggestive (mis)information if they had been tested earlier on their

memory than participants who did not receive such an initial test. The first empirical demonstration of this effect was provided by Chan and colleagues (2009). In three experiments, they had adult participants view a video, and then half of the participants were tested on the content of the video by means of a cued recall test. The other participants did not receive a memory test. Following this, all participants were exposed to misinformation through an audio narrative. Next, all participants received a final cued recall test. The authors found that participants who received an intervening memory test were more prone to reporting misinformation on the final test than those who did not receive the initial test.

The retrieval-enhanced suggestibility effect has been replicated under a variety of circumstances. For example, it has been demonstrated using various stimuli (Chan, Wilford, & Hughes, 2012; Butler & Loftus, 2017), in children (Brackmann, Otgaar, Sauerland, & Howe, 2016), younger and older adults (Chan et al., 2009), with repeated tests (Chan & LaPaglia, 2011), and in free and cued recall initial tests (Wilford, Chan, & Tuhn, 2014). A proposed explanation for retrieval-enhanced suggestibility is that the initial test inadvertently enhanced learning of the misinformation (Chan, Manley, & Lang, 2017; Thomas et al., 2010). Specifically, a wealth of research has shown that retrieval has both a backward and forward effect on learning. The backward effect refers to the idea that retrieving previously learned information strengthens accessibility of that information (i.e., the testing effect), whereas the forward effect refers to the idea that retrieval can facilitate subsequent learning of new information (Chan, Meissner, & Davis, 2018; Pastötter & Bauml, 2014; Yang, Potts, & Shanks, 2018). According to this forward effect of retrieval, taking an initial test can enhance future learning. In the context of the retrieval-enhanced suggestibility paradigm, the initial interview therefore ironically enhances an eyewitness' ability to learn the misinformation that is presented following the interview.

When participants are confronted with misinformation, details related to the initially tested material capture their attention and become better encoded. This increased encoding of the misinformation in turn eases its retrieval during the final test, leading to retrieval-enhanced suggestibility (for a review, see Chan, Manley, & Lang, 2017).

Of relevance for the current experiments is research showing that retrieval-enhanced suggestibility also appears when memory is tested using an empirically well-validated interview protocol. In one experiment (LaPaglia, Wilford, Rivard, Chan, & Fisher, 2014), after watching a video, adult participants either received an immediate Cognitive Interview or not. The Cognitive Interview has been proven to be an effective interview in producing accurate memory reports from adult witnesses in experimental contexts (Fisher & Geiselman, 1992; Memon, Meissner, & Fraser, 2010). Next, misinformation was presented to all participants. During the final test, it was shown that participants who completed the Cognitive Interview became more vulnerable to misinformation relative to participants who did not receive the Cognitive Interview. This experiment clearly showed that even a well-established interview protocol does not prevent adult witnesses from being particularly prone to misinformation.

**The Current Experiments**

Although the retrieval-enhanced suggestibility effect has been observed in a variety of contexts, the current experiments aimed at assessing whether this effect is present in children when they have been interviewed following best practice recommendations such as the NICHD protocol. The relevance of this empirical exercise become evident when considering the following issues. First of all, to date, no research has focused on examining whether high quality interviews actually protect children from follow-up misinformation. There is only one study on testing and misinformation in children (Brackmann et al., 2016). Although retrieval-enhanced

suggestibility was found in this study, the testing phase used in this study was substantially different than that in a NICHD Interview. That is, in the Brackmann et al. study, after viewing a video of a mock theft, participants received only *what*-questions (e.g., What did [the thief] wear?). Such questions are called directives in the NICHD Protocol and they are used after interviewers have been engaged in a thorough questioning phase using invitations. It is unclear whether retrieval-enhanced suggestibility will be found in children when questions are asked that lead to more detailed responses than what-questions.

Second, although previous *adult* research using the Cognitive Interview (LaPaglia et al., 2014) has found retrieval-enhanced suggestibility effects, this does not necessarily mean that this should be found in children who are interviewed by the NICHD Protocol as well. Other memory phenomena exist that differ between children and adults too. For example, false memories elicited by the so-called Deese/Roediger-McDermott (Deese, 1959; Roediger & McDermott, 1995) paradigm are less likely to be induced in children than in adults (Brainerd, Reyna, & Ceci, 2008). In this paradigm, participants are presented with lists containing associatively-related words (e.g., bed, rest, awake, tired, etc) that are linked to a non-presented critical lure (i.e., sleep). During memory tests, participants oftentimes falsely remember the critical lure and this effect is largest in adults.

Also, in contrast to LaPaglia et al.'s (2014) work, Gabbert, Hope, Fisher, and Jamieson (2012) found that participants who completed the Self-Administered Interview were more resistant against later misinformation than participants who did not complete the interview. The Self-Administered Interview is built around the same memory principles of the Cognitive Interview (e.g., context reinstatement, report everything), but – as the name states – witnesses can administer the interview themselves. The explanation behind the resistance against later

misinformation is that a well-established interview, such as the NICHD Protocol, may lead to strong memories (aside from its advantage of being developmentally-sensitive to children). Given that eyewitnesses are more resistant to misinformation when they are able to detect discrepancies between the witnessed event and the misinformation narrative (Tousignant, Hall, & Loftus, 1986), it is possible that the NICHD would increase the likelihood that children would detect such discrepancies. So, based on this principle of discrepancy detection, it is unclear whether retrieval-enhanced suggestibility will be found in children using the NICHD Protocol.

Thus, based on these issues and previous research demonstrating that retrieval-enhanced suggestibility was observed among children albeit without the use of a well-validated interview protocol (Brackmann et al., 2016) and when adults were interviewed by the Cognitive Interview (LaPaglia et al., 2014), our primary purpose was to examine whether retrieval-enhanced suggestibility would be observed in child participants when the NICHD Protocol was employed. As the NICHD Protocol is regarded as one of the best techniques to interview children and has been proven to support children's memory (La Rooy et al., 2015; Lamb, Orbach, Hershkowitz, Esplin, & Horowitz, 2007), if retrieval-enhanced suggestibility is found using this interview method, this would be a stronger demonstration of retrieval-enhanced suggestibility than when for example other methods are used (e.g., free recall). In two experiments, all children watched a video. We included a wide age range (5- to 10-years old) of children in our experiment. Although research shows that misinformation effects are more pronounced in 5-year-olds than 10-year olds (Ceci & Bruck, 1993; Otgaar et al., 2016), similar testing effects have been observed among different age groups such as 5-year-olds, 7-year-olds, and 11-year-olds; Brackmann et al., 2016; Brainerd & Reyna, 1996). Hence, this implies that we could find similar retrieval-enhanced suggestibility effects in different aged-children as well.

We divided children into four groups. For two groups (Experiment 1), half of the children were interviewed about the video using procedures based on the NICHD protocol. The other half did not receive an interview. Following this, all children received misinformation and then a final memory test. These two groups were tested to examine whether we could find retrieval-enhanced suggestibility in children using the standard procedure as in previous work (Chan et al., 2009). For the other two groups (Experiment 2), misinformation and the final memory test were presented one week later. The reason for presenting misinformation at a different time to interviews was for both practical and theoretical considerations. In legal cases, children, after being interviewed, can also be exposed to follow-up misinformation. For example, immediately after being interviewed, it could be normal for a guardian to question the child and they may inadvertently ask suggestive questions and provide misinformation to the child about what they think happened. Furthermore, such suggestive information can also occur after an interval when the children speak with others who are concerned about them (e.g., friends, therapists). Of particular relevance to retention intervals is research showing that the benefits of prior testing increase with delay (Chan, 2010), and that the retrieval-enhanced suggestibility effect can be reduced when misinformation is encountered after a delay (Chan & LaPaglia, 2011, 2013; but see also Chan & Langley, 2010). This implies that no retrieval-enhanced suggestibility will be found in children after a delay. It is important to stress here that, of course, in many child sexual abuse cases, children are not immediately interviewed after being mistreated (Saywitz et al., in press) and the NICHD Protocol has been in part developed to obtain accurate accounts even in such cases. For the present purpose, the relevance of the current experiments lies in those situations in which early interviews did occur after an incident, and whether such interviews might affect children's suggestibility later.

Thus, based on the findings of research on retrieval-enhanced suggestibility, one might expect that although best practice recommendations for interviewing children are a pathway to obtaining accurate accounts of children's experiences, it might also increase children's susceptibility to misinformation thereby evincing retrieval-enhanced suggestibility. We expected that this would be especially the case when misinformation was presented immediately following and interview procedure based on the NICHD Protocol. Examining these predictions is important as other research has found protective effects of the Self-Administered Interview for misinformation (Gabbert et al., 2012) and because retrieval-enhanced suggestibility has not been assessed in children using well-validated interview recommendations.

### Experiment 1

### Method

Before data collection, this experiment was pre-registered on the Open Science Framework. All data and materials can be found on https://osf.io/s4ekh/

**Participants**

For Experiments 1 and 2, an a priori power analysis was performed to determine the sample size using G*Power. To gain a power of .85, with an effect size of .25 and an α-level of .05, 128 participants had to be included in the sample. Several primary schools were approached and one gave permission to conduct the research among the children of their school. The director of the school provided consent to test 132 children for our experiment, thereby meeting our desired sample sizes. Letters were sent to the parents/caretakers of 132 children to inform them about the research. However, parents of 21 children did not want their children to take part in the study. During the experiment, another three children were excluded from the sample because they did not attend school at the times they needed to be interviewed or tested. Therefore, the

current research was carried out among 108 children between the age of 5 and 10 years, with a mean age of 7.32 years (*SD* = 1.33). The participants were 53.2% boy and 46.8% girl. The study was approved by the standing ethical committee of the Faculty of Psychology and Neuroscience, Maastricht University.

**Materials**

**Stimuli.** Children were presented with an 8-minute video clip from the movie "Return of the Pink Panther." This video has been used in previous retrieval-enhanced suggestibility experiments (Chan et al., 2012). The video depicts a burglar breaking into a museum, stealing a diamond, and escaping from guards.

**Interview Protocol.** Half of the children were interviewed using a procedure adapted from the NICHD Protocol suitable for research purposes. For example, the NICHD Protocol also includes questions concerning the disclosure of an alleged event. For the experiments, this was not necessary to ask. One female research assistant received a 1-day training on conducting an NICHD Interview using the Dutch translation of the protocol (de Ruiter, Otgaar, Serie, & van den Hombergh, 2016). The research assistant was trained in the different components of the protocol (e.g., rapport building, episodic memory training) by means of practice role-playing interviews. She was trained by a colleague who had received previous training in the use of the NICHD Protocol in child abuse investigations (first author).

**Misinformation**. Children were presented with misinformation through an audio narrative containing details about the event. We defined 15 pieces of details in the video as "critical details" for which participants would be questioned during the final memory test. The narrative presented five of these critical details as consistent details (i.e., they conformed to what was depicted; e.g., a consistent detail was that the walls were white), five as control details (i.e.,

they omitted the critical information), and five as misleading details (i.e., they were incorrect). This narrative was based on the one used in Chan et al. (2012). It was slightly translated into Dutch and adapted to be more suitable for children (i.e., we left out the part where the guards fired shots at the burglar but missed). The misinformation narrative contained 640 words and the audio narrative lasted about 3:47 minutes.

**Memory Test.** The final memory test contained 15 questions: 5 targeted the misinformation items, 5 targeted the control items (items referring to details only presented in the video), and 5 targeted the consistent items (items involving details being presented in the video and narrative). This test was derived from the test used by Wilford et al. (2014). Again, it was slightly adapted to be more appropriate for children (i.e., the question about the form of the window was replaced by how the burglar landed on the floor). Misinformation questions contained three options, of which one was the correct answer, one was false, and one was the misinformation detail (e.g., "Which color was the letter on the glove?" White (false), gold (misinformation), silver (correct)). Control items were open questions (e.g., "The man escaped from the museum. How many guards chased him?"). Consistent items contained three options of which only one was correct. All children received the same items in the exact same order.  For all question types (i.e., misinformation, control, consistent), we coded whether the answer was correct. In addition, for the misinformation question type, we also coded whether the false (i.e., misinformation) option was chosen.

As in Wilford and colleagues' (2014) study, we could not use the same detail for the different item types due to the brief nature of our video. So, the consistent, control, and misinformation items were always about different details of the video clip.  This means that no comparison was made across item types, but because our key interest is testing, and not item

type, this is not a matter of concern.

**Design and Procedure**

Experiment 1 used a between-subjects design in which children were randomly assigned to two conditions: Adapted NICHD Protocol and immediate misinformation ($n = 27$; mean age = 7.56, $SD = 1.40$) and no interview and immediate misinformation ($n = 27$; mean age = 7.39, $SD = 1.39$). Children's age did not statistically differ between the two groups, $t(51) = 0.45$, $p = 0.66$.

The experiment took place at the primary school of the children. It was carried out in a separate classroom with only the participant and the experiment leader present, and the children were tested individually. First, the video was shown to the children. It was shown on a laptop and it took approximately eight minutes. Next, a 2-min distractor task was presented (i.e., circling plus-signs among different symbols) and then children were either interviewed using the Adapted NICHD Protocol or they played a game for 15 minutes. During the interview, children received certain ground rules (e.g., It is ok to say "I don't know") of importance for the interview. After this, children engaged in rapport building in which they were asked what they like to do. Next, children received training in episodic memory retrieval where they had to retrieve a memorable event (e.g., birthday) and talk about it. Then, the substantive part of the interview took place. Here, they received open prompts to talk about the video they just viewed ("Tell me what happened") and when needed more specific questions were asked (e.g., directives or option-posing questions). The interview concluded with asking what children would be doing after the interview. Directive questions were often not needed as many children did not provide any extra information after some invitations. This also shows that although the Adapted NICHD Protocol was used in the current experiments, it of course differs in several aspects from actual forensic interviews (e.g., length, topic discussed). Following the interview (or the game), the

same 2-min distractor task was used and then all children listened to the misinformation audio-narrative. Children were told that the narrative was a summary of the event. After the presentation of the misinformation, participants received the same distractor task and then the final memory test was presented. Completing the test took about five minutes. Children and their parents were debriefed after the experiment.

### Results and Discussion

Independent samples *Welch* t-tests were conducted on the endorsement of misinformation items. i.e., (in)correct recall of misinformation, and correct recall of consistent and control items.[1] The most important analysis concerned the misinformation items. Children who were interviewed by the Adapted NICHD Protocol were more prone to reporting misinformation ($M =$ 2.41, $SD = 1.22$) than children who were not interviewed – a retrieval-enhanced suggestibility effect ($M = 1.52$, $SD = 1.09$), $t(51.35) = 2.83$, $p = .007$, Cohen's $d = 0.77$ (see Table 1). We also performed a Bayesian analysis and using a default prior of 0.707 given in the statistical software program JASP, a Bayes Factor of 6.65 was detected. This result indicates that our data are more in line with the alternative (difference between the groups) than the null hypothesis (no difference). No significant effects of the Adapted NICHD Protocol emerged concerning the control ($t(47.62) = -1.64$, $p = .11$, Cohen's $d = -0.45$) and consistent items ($t(47.14) = 0.69$, $p = .49$, Cohen's $d = 0.19$). To also examine whether the presentation of misinformation reduced correct responding on the misinformation items, we conducted an independent samples *Welch* t-test on correct responding on the misinformation items with Condition (Adapted NICHD

---

[1] Although in the preregistration we noted to perform two-way ANOVAs on all data (Experiments 1 & 2), in retrospect, we reasoned that because in Experiment 1 the memory test was provided immediately and in Experiment 2 after one week, it would be better to conduct separate independent samples t-tests on the immediate memory data and the one week memory data

Protocol vs. no interview) as a between-subjects variable. No statistically significant effect was found ($p > .05$; see Table 1).

This experiment demonstrates that an adaptation of the NICHD Protocol that embodies accepted guidelines for interviewing children can lead to retrieval-enhanced suggestibility in children. When children were interviewed by this NICHD Protocol, they were more likely to assent to the later presented misinformation than children who were not tested. This is the first study showing that a well-established and empirically validated child interview protocol can increase children's suggestibility. The current finding is well in line with research showing that the Cognitive Interview can increase adults' suggestibility as well (LaPaglia et al., 2014). The next question is whether this effect also appears when misinformation is presented after a delay. The importance of this becomes clear considering the fact that many legal cases can draw out over the course of months or even years, and when the duration of a legal case gets extended, the likelihood that a child witness will be exposed to misinformation also increases as a necessity. Therefore, in Experiment 2, we examined whether completing an immediate interview based on the Adapted NICHD protocol can protect child witnesses from misleading suggestions that are encountered a week later. Although we discovered a research-enhanced suggestibility effect in Experiment 1, we hypothesized that the same may not occur in Experiment 2 because research in adult samples has demonstrated that retrieval-enhanced suggestibility can be reduced or even eliminated after a delay (Chan & LaPaglia, 2011).

## Experiment 2

### Method

**Participants**

See Experiment 1. Children were randomly assigned to a group receiving the Adapted

NICHD Protocol and misinformation after a 1-week delay (mean age = 7.56, *SD* = 1.48) or a group who was not tested and received misinformation after a 1-week delay (mean age = 7.64, *SD* = 1.22). Children's age did not statistically differ between the two groups ($t(53)$ = -0.24, *p* = 0.81).

**Materials**

The same materials were used as in Experiment 1.

**Design and Procedure**

A similar design and procedure was employed as in Experiment 1. The only difference was that misinformation and the final memory test were presented after 1-week.

## Results and Discussion

In contrast to Experiment 1, no evidence was found for a retrieval-enhanced suggestibility effect, $t(48.87)$ = -1.18, *p* = 0.25, Cohen's *d* = -0.32 (see Table 2). Also, regarding the control and consistent items, no statistical differences were detected between the groups (control items: ($t(51.70)$ = 0.99, *p* = 0.33, Cohen's *d* = -0.27; consistent items: ($t(51.64)$ = 1.37, *p* = 0.18, Cohen's *d* = 0.37).

To investigate the effect of the Adapted NICHD interview on correct responding of misinformation items, we conducted an independent samples *Welch* t-test with Condition (Adapted NICHD Protocol vs. no interview) as a between-subjects variable. We found that the NICHD Protocol led to higher correct responding scores for the misinformation items, $t(43.62)$ = 2.51, *p* = 0.02, Cohen's *d* = 0.68; see Table 2).

**Exploratory Analyses**

**Age**. We also conducted several exploratory analyses. Specifically, we divided children in two age groups (5-7-year-olds vs. 8-10-year-olds) and examined whether retrieval-enhanced

suggestibility would occur independent of age. We deemed it important to conduct this analysis because previous research has indicated developmental differences in memory accuracy, reporting of misinformation and false memory propensity between these age groups (e.g., Brainerd, Reyna, & Ceci, 2008; Goodman, 2006; Otgaar et al., 2016). For children that received misinformation immediately (i.e., data from Experiment 1), we performed a 2 (Adapted NICHD Protocol vs. no test) x 2(5-7-year-olds ($n = 22$) vs. 8-10-year-olds ($n = 32$) factorial ANOVA. No statistically significant interaction emerged ($F(1,50) = 3.13$, $p = .08$, $\omega^2 = 0.03$). Also, no statistically significant age effect was detected ($F(1,50) = 1.05$, $p = .31$, $\omega^2 = 0.001$). We did find a retrieval-enhanced suggestibility effect for all children ($F(1,50) = 10.67$, $p = .006$, $\omega^2 = 0.12$) showing that children who were interviewed had higher misinformation scores (($M = 2.41$, $SD = 1.22$) than children who were not tested ($M = 1.52$, $SD = 1.09$).

We also conducted a similar factorial ANOVA but now included children (5-7-year-olds ($n = 20$), 7-9-year-olds ($n = 34$)) who received misinformation one week after being presented with the video (i.e., data from Experiment 2). We only found a statistically significant age effect ($F(1,50) = 13.58$, $p < .001$, $\omega^2 = 0.19$) with older children ($M = 1.94$, $SD = 1.25$) endorsing more information than younger children ($M = 1.21$, $SD = 1.22$).

We also included age as an extra factor to examine the effect of the Adapted NICHD Protocol on correct responding of control, consistent, and misinformation items. We first looked at data from Experiment 1. For control items, we found a statistically significant interaction between age and prior interview ($F(1,50) = 10.20$, $p = .002$, $\omega^2 = 0.13$). Simple effect analyses using independent samples t-tests showed the following. For children receiving the Adapted NICHD Protocol, younger ($M = 2.45$, $SD = 1.04$) and older children ($M = 2.38$, $SD = 0.96$) did not differ in their accuracy for control items ($t(25) = 0.21$, $p = 0.84$, $d = 0.07$). However, for

children who did not receive an initial interview, younger children ($M = 1.91$, $SD = 1.04$) were less likely to be correct on control items than older children ($M = 3.63$, $SD = 1.03$, $t(25) = -4.24$, $p < 0.001$, $d = 1.66$). For consistent items, we found a statistically significant main effect of age ($F(1,50) = 10.53$, $p = .002$, $\omega^2 = 0.14$), with younger children ($M = 2.55$, $SD = 1.38$) have lower accuracy levels than older children ($M = 3.50$, $SD = 0.76$). For correct responding of misinformation items, no statistically significant effects emerged (all $p$s $> .05$).

When we examined the data from Experiment 2, the following effects were observed. For control items, we only demonstrated that younger children ($M = 1.25$, $SD = 0.98$) had lower accuracy than older children ($M = 2.35$, $SD = 1.19$; $F(1,50) = 12.17$, $p = .001$, $\omega^2 = 0.17$). For consistent items, younger children ($M = 2.10$, $SD = 1.28$) also displayed lower accuracy than older children ($M = 2.88$, $SD = 1.05$; $F(1,50) = 5.29$, $p = .02$, $\omega^2 = 0.08$). For correct responding of misinformation items, we found that children receiving the Adapted NICHD Interview ($M = 1.94$, $SD = 1.29$) were more accurate than children not receiving the interview ($M = 1.23$, $SD = 0.80$; $F(1,50) = 5.37$, $p = .03$, $\omega^2 = 0.08$).

**Interview Performance.** We also examined whether children's memory performance when being interviewed using the Adapted NICHD Protocol correlated with performance on the final memory test. To examine this, we first made a scoring sheet of the video and identified 71 unique details that could be scored. We then scored children's statements and calculated accuracy levels (i.e., how many of the correct 71 details are stated by children) for each of them. The mean proportion accuracy was 0.18 ($SD = .07$). When we correlated children's accuracy levels with the final memory test performance, we found the following. First, we found a statistically significant positive correlation between accuracy levels and reporting of misinformation ($r(52) = 0.33$, $p = .02$; two children did not have any interview performance).

Second, a statistically significant positive correlation was detected between accuracy levels and correct responding of consistent questions ($r(52) = 0.27$, $p = .049$; two children did not have any interview performance).

In contrast to Experiment 1, Experiment 2 showed no signs of retrieval-enhanced suggestibility when misinformation was delivered after a 1-week interval. This shows that, at least in children, retrieval-enhanced suggestibility effects are perhaps only limited to situations when misinformation is presented soon after children are interviewed (cf. Chan & Langley, 2011). We did find that children receiving an Adapted NICHD interview had higher true memory scores thereby demonstrating a testing effect.

**General Discussion**

The foremost aim of the current experiments was to investigate whether research-based interview guidelines such as the NICHD Protocol could reduce or increase suggestibility in children. To address this aim, we conducted two experiments. In the first experiment, we presented misinformation immediately after children were interviewed using the Adapted NICHD Protocol. In the second experiment, children were exposed to misinformation one week later. We showed that immediate interviewing increases children's suggestibility albeit only when misinformation was presented immediately after the interview.

The finding that children displayed higher suggestibility when they were interviewed by the Adapted NICHD Protocols parallels and extends earlier research on retrieval-enhanced suggestibility. That is, our result of the first experiment is well in line with studies showing that immediate testing adversely affects memory in that it makes people more prone to accept external suggestions (e.g., Chan et al., 2009). It is also analogous with recent research showing that retrieval-enhanced suggestibility takes place in children as well (Brackmann et al., 2016).

Furthermore, it extends research showing that the empirically-validated Cognitive Interview can produce retrieval-enhanced suggestibility (LaPaglia et al., 2014). This prior research has focused on an interview protocol commonly used with adults. We showed that a well-established protocol designed specifically for children can also lead to increased levels of suggestibility for misinformation that is encountered soon after the interview.

One mechanism that has been proposed to account for retrieval-enhanced suggestibility is that testing facilitates the learning and encoding of subsequent misinformation. Prior retrieval may enhance learning of the later-presented misinformation by several means. Of most relevance to the present study, retrieval can enhance new learning by reducing inattention globally (Pastotter, Schicker, Niedernhuber, & Bauml, 2011; Szpunar, Khan, & Schacter, 2013), such that children who had received an earlier interview were more likely to sustain their attention during the post-event narrative. Further, the initial interview may enhance learning of the postevent information by making encoding more selective (Carpenter & Toftness, 2017; Davis & Chan, 2015). In the context of RES, the children who have received the initial interview may be more likely to detect conflicts between their recollection and the postevent information, thereby enhancing their learning towards the misinformation (Butler & Loftus, 2017; Gordon & Thomas, 2014). Our finding that the Adapted NICHD Protocol was positively related to misinformation reporting is consistent with this interpretation. Our experiments, combined with previous research (Brackmann et al., 2016), show that this mechanism might also underlie retrieval-enhanced suggestibility within children. One finding bears mentioning here. Although the testing effect (that is, the beneficial effects of prior retrieval on later memory) has been demonstrated in children and adults (Bouwmeester & Verkoeijen, 2011), research directed at retrieval-enhanced suggestibility has often failed to find a testing effect when the final test occurs in the same

session as the initial test (Chan et al., 2009). Consistent with this pattern, in the first experiment, we observed a retrieval-enhanced suggestibility effect but not a testing effect. More puzzling was the data from Experiment 2. Here, a one-week delay separated the NICHD interview and the final test, and the long retention interval is known to increase the testing effect (Wheeler, Ewers, & Buonanno, 2003), yet we still did not detect a testing effect. However, we caution against over-interpreting this null effect, given that recognition is not always sensitive to the benefits of prior testing (Chan & McDermott, 2007; Rowland, 2014). Regardless, further empirical scrutiny is needed to unravel the mechanisms behind when a testing effect can be expected in children in the context of the RES paradigm.

Although we found evidence for elevated levels of suggestibility when children were interviewed using the Adapted NICHD Protocol, this was only revealed when misinformation was presented directly after the interview. Our second experiment showed that this effect vanished when misinformation was provided 1-week later. To some extent, this is related to research by Chan and LaPaglia (2011). They found that retrieval-enhanced suggestibility could be eliminated after a delay but only when a within-subjects design was used. In such a design, all participant witnesses are tested initially, but retrieval-enhanced suggestibility is assessed by comparing the misinformation recall probabilities between items that were included in the initial test with items that were omitted from the initial test. Our finding of the second experiment might be seen as encouraging in terms of practical issues. That is, for misinformation that is encountered after a delay, and not immediately, after someone is interviewed, retrieval may not increase suggestibility (Otgaar et al., 2017; Otgaar & Howe, 2017). This finding is consistent with the prediction emerging from the memory reconsolidation framework, which has been applied as a secondary explanation for retrieval-enhanced suggestibility (Chan & LaPaglia, 2013;

Chan et al., 2017). The idea is that a recently retrieved memory needs to be reconsolidated. During this reconsolidation period, the memory remains in a temporary labile state, ready for updating. When misinformation is encountered during this period, the original event memory may be updated with the misinformation, leading to retrieval-enhanced suggestibility. However, when misinformation is presented outside of the period of lability, it should not affect reconsolidation of the original event memory. Consistent with this idea, the data from Experiment 2 show that when participants are interviewed one week before they encounter misinformation, the interview should not increase suggestibility.

We also conducted some exploratory analyses. One noteworthy finding was that in the second experiment, we found misinformation effects to be stronger in the older than younger children. This is in contrast with a plethora of research demonstrating that younger children are more vulnerable to external suggestions than older children (e.g., Ceci & Bruck, 1993; Loftus, 2005; Otgaar, Candel, Smeets, & Merckelbach, 2010; Sutherland & Hayne, 2001). However, this result is in line with recent research showing that under certain conditions, susceptibility to suggestion increases with age (Otgaar, Howe, Brackmann, & Smeets, 2016; Otgaar, Howe, Brackmann, & van Helvoort, 2017). This research shows that such reversals in suggestibility take place when external suggestive pressure capture the underlying meaning/associative structure of the event. Specifically, research shows that younger children are less able to grasp the underlying meaning of events than older and children and hence, are less likely to spontaneously make incorrect associations ending up into the formation of false memories. This means that when stimuli are associative in nature and suggestion is presented that preserves the underlying meaning of the events, false memories will increase with age. This phenomenon is also called a developmental reversal (Brainerd, Reyna, & Ceci, 2008). Our experiments were not

designed to examine such developmental reversals in suggestibility, but future experiments should examine whether our procedure is likely to engender such developmental reversals. In general, our exploratory age analyses should be treated with caution as the sample size of the separate age groups were quite small. Our experiments show that retrieval-enhanced suggestibility is present in different age groups and given an abundance of work showing that younger children are more suggestible than older children (e.g., Bruck & Ceci, 1999; Leichtman & Ceci, 1995; Goodman & Reed, 1986), future work could attempt to replicate whether retrieval-enhanced suggestibility is developmentally invariant.

It might be appealing to postulate that our finding that retrieval-enhanced suggestibility takes places when a procedure similar to the NICHD Protocol is used is perilous in interviewing settings. However, there are several assertions against this view. First, as has been mentioned before, we only found retrieval-enhanced suggestibility when children were confronted with misinformation immediately. Of course, in many legal cases, such suggestive pressure is not immediately imposed but often occurs after a significant period of time (Howe, Knott, & Conway, 2018). Relatedly, in Experiment 2, misinformation was presented after 1 week. The absence of retrieval-enhanced suggestibility in the second experiment could be due to the misinformation and/or memory test after 1 week. Future experiments should focus on presenting misinformation immediately and varying the interval of the memory test. Second, in our study, we compared memory performance between children who received a NICHD Interview and children that did not receive any interview. Given the lack of another interview comparison condition, we cannot ascertain whether the NICHD Protocol specifically was leading to retrieval-enhanced suggestibility. To examine this issue, future studies should add a third group of children that, for example, is instructed to freely recall what they experienced (in addition to a

group of participants who receives the NICHD interview). Of importance, this has been done in previous research using the Cognitive Interview in adults (LaPaglia et al., 2014). Here, the authors found that although the Cognitive Interview elicited more correct details than the free recall group, the size of retrieval-enhanced suggestibility effect did not differ between the two groups.

Third, in all studies on retrieval-enhanced suggestibility, the witnessed event is a video or a slideshow but not a personally experienced event (e.g., physical or sexual abuse). Hence, it is vital that prospective research examines whether retrieval-enhanced suggestibility will also be observed when participants are involved in a more personal experience (e.g., interactive event). Fourth, in the present experiments, the final memory test consisted of a mixture of open-ended and forced-choice questions. Although retrieval-enhanced suggestibility has been found for cued recall and free recall (Wilford et al., 2014), this has not been examined in children. Considering that children in interview settings often have to provide a free recall of the experienced event, it is imperative to know whether retrieval-enhanced suggestibility would occur in such a situation as well. Fifth, even if misinformation would be conveyed soon after the interview, this does not preclude the use of research-based interview procedures such as the NICHD Protocol. Following research-based interview procedures is currently the best method for obtaining reliable statements in children and they should ideally always be followed when children need to talk about traumatic experiences such as child abuse. Hence, we agree with LaPaglia and colleagues (2014, p.7) who stated, "The idea that the process of conducting an effective initial interview increases eyewitnesses' sensitivity to post-interview misinformation is an unfortunate byproduct, but this should not deter criminal investigators from conducting the most effective initial interview possible".

To recap, our experiments show that when empirically recommended best practice guidelines are followed, children are more susceptible to suggestion when they receive misinformation immediately after being interviewed than when children are not interviewed. This retrieval-enhanced suggestibility did not consistently emerge when misinformation was presented after 1-week. This research, combined with recent research on age-related increases in suggestibility (Otgaar et al., 2016), clearly shows that the field of children's suggestibility is much more complex than what has previously been thought.

**References**

Bouwmeester, S., & Verkoeijen, P. P. J. L. (2010). Latent variable modelling of cognitive

processes in true and false recognition of words: A developmental perspective. *Journal of*

*Experimental Psychology*: *General, 139*, 365–381

Brackmann, N., Otgaar, H., Sauerland, M., & Howe, M.L. (2016). The effect of testing on the

formation of children and adults' false memories. *Applied Cognitive Psychology*, *30*, 785-

794.

Brainerd, C.J., Reyna, V.F., & Ceci, S.J. (2008). Developmental reversals in false memory: A

review of data and theory. *Psychological Bulletin*, *134*, 343–382.

Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of*

*Psychology, 50,* 419–439.

Butler, B. J., & Loftus, E. F. (2017). Discrepancy detection in the retrieval-enhanced

suggestibility paradigm. *Memory*, *26*, 483–492.

Carpenter, S. K., & Toftness, A. R. (2017). The effect of prequestions on learning from video

presentations. *Journal of Applied Research in Memory and Cognition*, *6*, 104–109.

Ceci, S. J., & Bruck, M. (1993). The suggestibility of the child witness: A historical review and

synthesis. *Psychological Bulletin*, 113, 403–439.

Chan, J.C.K. (2010). Long-term effects of testing on the recall of nontested materials. *Memory,*

*18*, 49-57.

Chan, J. C. K., & Langley, M. (2011). Paradoxical effects of testing: Retrieval enhances both

accurate recall and suggestibility in eyewitnesses. *Journal of Experimental Psychology:*

*Learning, Memory, and Cognition, 37,* 248-255.

Chan, J. C. K., & LaPaglia, J. A. (2011). The dark side of testing memory: Repeated retrieval can enhance eyewitness suggestibility. *Journal of Experimental Psychology: Applied*, *17*, 418-432.

Chan, J. C. K., & LaPaglia, J. A. (2013). Impairing existing declarative memory in humans by disrupting reconsolidation. *Proceedings of the National Academy of Sciences of the United States of America, 110,* 9309-9313. doi:10.1073/pnas.1218472110

Chan, J. C., Manley, K. D., & Lang, K. (2017). Retrieval-enhanced suggestibility: A retrospective and new investigation. *Journal of Applied Research in Memory and Cognition, 6,* 213-229. doi: 10.1016/j.jarmac.2017.07.003

Chan, J. C. K., & McDermott, K. B. (2007). The testing effect in recognition memory: A dual process account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 431–437.

Chan, J. C. K., Meissner, C. A., & Davis, S. D. (2018). Retrieval potentiates new learning: A theoretical and meta-analytic review. *Psychological Bulletin.*

Chan, J. C. K., Thomas, A. K., & Bulevich, J. B. (2009). Recalling a witnessed event increases eyewitness suggestibility: The reversed testing effect. *Psychological Science*, *20*, 66-73.

Chan, J. C. K., Wilford, M. M., & Hughes, K. L. (2012). Testing can increase or reduce suggestibility depending on how memory is tested: The importance of source complexity. *Journal of Memory and Language*, *67*, 78-85. doi: 10.1016/j.jml.2012.02.006

Davis, S. D., & Chan, J. C. K. (2015). Studying on borrowed time: how does testing impair new learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1741–1754.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate

recall. *Journal of Experimental Psychology, 58*, 17-22.

Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative*

*interviewing: The cognitive interview*. Springfield, IL: Charles C. Thomas.

Gabbert F., Hope L., Fisher R. P., & Jamieson K. (2012). Protecting against misleading post-

event information with a Self-Administered Interview. *Applied Cognitive Psychology*, *26*,

568–575. DOI:10.1002/acp.2828.

Goodman, G.S. (2006). Children's eyewitness memory: A modern history and contemporary

commentary. *Journal of Social Issues, 62*, 811-832.

Goodman, G. S., & Reed, R. (1986). Age differences in eyewitness testimony. *Law and Human*

*Behavior, 10*, 317–332.

Gordon, L. T., & Thomas, A. K. (2014). Testing potentiates new learning in the misinformation

paradigm. *Memory & Cognition*, *42*, 186–197.

Huff, M. J., Weinsheimer, C. C., & Bodner, G. E. (2016). Reducing the misinformation effect

through initial testing: Take two tests and recall me in the morning? *Applied Cognitive*

*Psychology*, *30,* 61-69.

Lamb, M.E., Orbach, Y., Hershkowitz, I., Esplin, P.W., & Horowitz, D. (2007). Structured

forensic interview protocols improve the quality and informativeness of investigative

interviews with children: A review of research using the NICHD Investigative Interview

Protocol. *Child Abuse and Neglect, 31*, 1201–1231. doi: 10.1016/j.chiabu.2007.03.021.

LaPaglia, J. A., Wilford, M. M., Rivard, J. R., Chan, J. C. K., & Fisher, R. P. (2014). Misleading

suggestions can alter later memory reportseven following a Cognitive Interview. *Applied*

*Cognitive Psychology, 28*, 1–9.

La Rooy, D., Brubacher, S. B., Arom€aki-Stratos, A., Cyr, M., Hershkowitz, I., Korkman, J., Myklebust, T., Naka, M., Peixoto, C. E. S., Roberts, K. P., Stewart, H. L., & Lamb, M. E. (2015). The NICHD Protocol: A review of an internationally-used evidence-based tool for training child forensic interviewers. *Journal of Criminological Research, Policy and Practice*, 1, 76–89. doi:10.1108/ JCRPP-01-2015-0001

Leichtman, M.D. & Ceci, S.J. (1995). The effects of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology, 31*,568-578.

Loftus, E.F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning and Memory, 12*, 361-366.

Memon, A., Meissner, C.A., & Fraser, J. (2010). The cognitive interview: A meta- analysis and study space analysis of the past 25 years. *Psychology, Public Policy, & Law*, *16*, 350–372 doi: 10.1037/a0020518.

Otgaar, H., Candel, I., Smeets, T., & Merckelbach, H. (2010). "You didn't take Lucy's skirt off": The effect of misleading information on omissions and commissions in children's memory reports. *Legal & Criminological Psychology*, *15*, 229-241.

Otgaar, H., & Howe, M.L. (2017). *Finding the truth in the courtroom: Dealing with deception, lies, and memories*. Oxford University Press, UK.

Otgaar, H., Howe, M. L., Brackmann, N., & van Helvoort, D. (2017). Eliminating age differences in children's and adults' memory conformity and suggestibility. *Developmental Psychology, 53,* 962-970.

Otgaar, H., Howe M. L., Brackmann N., & Smeets T. (2016). The malleability of developmental trends in neutral and negative memory illusions. *Journal of Experimental Psychology*: *General, 145*, 31–55. DOI:10.1037/xge0000127.

Otgaar, H., de Ruiter , C., Howe, M.L., Hoetmer, L., & Van Reekum, P. (2017). A case concerning children's false memories of abuse: Recommendations regarding expert witness work. *Psychiatry, Psychology, and Law*, *24*, 365-378.

Pastötter, B., & Bauml, K. H. (2014). Retrieval practice enhances new learning: the forward effect of testing. *Frontiers in Psychology*, *5*, 1–5.

Pastötter, B., Schicker, S., Niedernhuber, J., & Bauml, K. H. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 287–297.

Pipe, M.E., Lamb, M.E., Orbach, Y., & Cederborg, A.C. (2007). *Child sexual abuse: Disclosure, delay and denial.* Mahwah, NJ: Lawrence Erlbaum Associates.

Roediger, H.L., III, & Karpicke, J.D. (2006). The power of testing memory: Basic research and implications for educational practice. *Psychological Science, 1,* 181-210.

Roediger, H. L., & McDermott, K. (1995). Creating false memories: Remembering words not presented in a list. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803-814.

Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychological Bulletin*, *140*, 1432–1463.

de Ruiter, C., Otgaar, H.,  Serie, C., & van den Hombergh, M. (2016*). A Dutch translation of the NICHD Protocol.* Unpublished manuscript.

Saywitz, K. J., Lyon, T., & Goodman, G. S. (in press). When interviewing children: A review and update. In J. Conte & B. Klika (Eds.), *APSAC handbook on child maltreatment*. Newbury Park, CA: Sage.

Sutherland, R., & Hayne, H. (2001). Age-related changes in the misinformation effect. *Journal of Experimental Child Psychology*, *79*, 388-404.

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proceedings of the National Academy of Sciences of the United States of America*, *110,* 6313–6317.

Thomas, A.K., Bulevich, J. B., & Chan, J.C.K. (2010). Testing promotes eyewitness accuracy with a warning: Implications for retrieval enhanced suggestibility. *Journal of Memory & Language*, *63*, 149-157.

Yang, C., Potts, R., & Shanks, D. R. (2018). Enhancing learning and retrieval of new information: a review of the forward testing effect. *Science of Learning*, *3*, 8.

Wheeler, M. A., Ewers, M., & Buonanno, J. F. (2003). Different rates of forgetting following study versus test trials. *Memory*, *11*, 571-580.

Wilford, M. M., Chan J. C. K., & Tuhn S. J. (2014). Retrieval enhances eyewitness suggestibility to misinformation in free and cued recall. *Journal of Experimental Psychology*: *Applied, 20*, 81–93. DOI:10.1037/xap0000001.

Table 1.

*Response frequency as a function of item type and prior interview group in Experiment 1*

|  | Group | N | Mean | SD |
|---|---|---|---|---|
| Misinformation items (chose misinformation) | NICHD | 27 | 2.41 | 1.22 |
|  | No interview | 27 | 1.52 | 1.09 |
| Control items (chose correct) | NICHD | 27 | 2.41 | 0.97 |
|  | No interview | 27 | 2.93 | 1.33 |
| Consistent items (chose correct) | NICHD | 27 | 3.22 | 0.97 |
|  | No interview | 27 | 3.00 | 1.36 |
| Misinformation items (chose correct) | NICHD | 27 | 1.56 | 1.10 |
|  | No interview | 27 | 2.11 | 1.05 |

Table 2.

*Response frequency as a function of item type and prior interview group in Experiment 2*

|  | Group | N | Mean | SD |
|---|---|---|---|---|
| Misinformation items (chose misinformation) | NICHD | 27 | 1.15 | 1.10 |
|  | No interview | 27 | 1.56 | 1.42 |
| Control items (chose correct) | NICHD | 27 | 2.11 | 1.19 |
|  | No interview | 27 | 1.78 | 1.28 |
| Consistent items (chose correct) | NICHD | 27 | 2.82 | 1.15 |
|  | No interview | 27 | 2.37 | 1.25 |
| Misinformation items (chose correct) | NICHD | 27 | 2.00 | 1.30 |
|  | No interview | 27 | 1.26 | 0.81 |