UNIVERSIDADE DE LISBOA

FACULDADE DE PSICOLOGIA

The role of retrieval for the correction of memory errors across development: a comparison between young adults and adolescents

Joana Filipa Reis de Almeida

MESTRADO INTERUNIVERSITÁRIO EM NEUROPSICOLOGIA CLÍNICA E EXPERIMENTAL

2022

UNIVERSIDADE DE LISBOA

FACULDADE DE PSICOLOGIA

The role of retrieval for the correction of memory errors across development: a comparison between young adults and adolescents

Joana Filipa Reis de Almeida

Dissertação orientada pela Professora Doutora Ana Raposo e pela Doutora María J. Maraver

MESTRADO INTERUNIVERSITÁRIO EM NEUROPSICOLOGIA CLÍNICA E EXPERIMENTAL

2022

Acknowledgements

The submission of my dissertation marks the conclusion of my master's degree and becomes a step in my professional course. Personally, however, it is also proof of my commitment and dedication to this project, which was potentiated and supported by the work of very important people I cannot leave unmentioned, without whom I would not be able to share this success.

To my supervisors Professor Ana Raposo, PhD, and María J. Maraver, PhD, for their unwavering support and feedback, that was always so kindly given. For believing in me and allowing me the opportunity to present this project at the *Jornadas de Investigação em Psicologia 2022* event organised by the faculty and to take it to this year's international APPE-SEPEX meeting at the University of Algarve, in the format of a scientific poster. For giving me the chance to develop myself and my professional skills with their guidance.

To Ana Lapa, MSc., for her help and constant involvement in the process of data collection, despite having taken on other simultaneous professional challenges.

To Marta Barros, an MSc student whose dissertation used the same experimental design to focus on the analysis of the Testing Effect, for the companionship and support system we develop while tackling the writings of our thesis.

To all of you who share my success and supported me in these trying times, you have my profound gratitude.

This work was supported by Fundação para a Ciência e a Tecnologia (FCT) through a Project Grant (PTDC/PSI-ESP/28414/2017).

Index

Abstract ……. 4

Resumo Alargado ……... 5

Introduction …… 9

The Formation of False Memories ……...…….. 9

Testing with Pragmatic Inferences ……..……. 13

Correcting False Memories …….. 15

When Errors Benefit Learning …….. 17

Learning and Error Correction in Adolescents …… 20

Our Study ……... 21

Method …….. 23

Participants ……... 23

Experimental Design ……. 23

Materials ……...…… 24

Procedure …….. 25

Encoding Phase ……. 25

Intermediate Phase …… 26

Final Memory Test …… 26

Source Monitoring Test ……. 26

Statistical Analysis ……..... 27

Results …… 28

Response Type ……..... 28

Persistence of Correct Responses and Error Correction Indexes ……..... 30

Discussion …….. 32

Conclusion ……..37

References …….. 38

Appendix A …… 48

Abstract

Pragmatic inference sentences (e.g., "*The baby stayed awake all night"*) are a robust way to study false memories through the dissociation between what people remember and what they know from general semantic knowledge (e.g., "*The baby cried all night*"), by inducing memory errors that can later be corrected after receiving corrective feedback.

Previous research on false memories has shown that adults benefit from active retrieval versus passive reading when memorizing sentences with pragmatic inferences. However, there is little information regarding how adolescents correct memory errors, as cognitive control abilities are still under development.

The present study aims to analyse how adolescents (12-17 years old) and young adults (22-27 years old) correct false memory errors generated by pragmatic inference sentences. For that, we compared performance in active (giving an answer) and passive (reading someone else's answer) recognition tasks followed by corrective feedback. Critically, the proportion of errors in both conditions was matched thanks to the yoked-pair design and the counterbalance of the sentences. In a final cued-recall task, memory of both age groups was compared as a function of their performance in the active vs. passive tasks, and we calculated a measure of persistence of the correct responses and an index of error correction.

Overall, error correction for actively retrieved sentences was higher than for those passively read. This effect was replicated in adolescents, even though it was not modulated by age. Our results demonstrated that adolescents benefit from the incorporation of feedback after retrieval compared to the passive processing of information, as adults do.

Keywords: memory retrieval, error correction, learning, development, adolescents

Resumo Alargado

A nossa memória de eventos passados nem sempre reflete a experiência vivida. Isto deve-se ao facto de a memória ser maleável e vulnerável a influências internas (e.g., similaridades contextuais e semânticas expectáveis) e externas (e.g., introdução de informação errónea ou corretiva). De acordo com Bartlett (1932), esta maleabilidade permite a sua reconstrução através da recuperação de informação. Estes erros de memória dão origem às falsas memórias – quando a nossa memória não coincide com os acontecimentos experienciados – que podem ocorrer no dia-a-dia numa população normativa saudável (Robins, 2019).

Segundo Carneiro e colegas (2020), a distorção da memória ocorre de forma a contemplar uma coerência entre a impressão tida da experiência vivida e o conhecimento e contexto atuais no momento em que recuperamos a informação. A proposta de Finn (2017) defende que a memória é atualizada depois de ser recuperada (MUAR) e que pode ser reconsolidada quando as pistas de recuperação são apresentadas com informação similar, que tanto pode ser errónea como corretiva. Nesta linha, este efeito facilitador da recuperação pode levar à integração de novas informações falsas e à consequente criação de falsas memórias ou à integração de feedback corretivo e atualizar memórias erróneas, quando o feedback é apresentado imediatamente após o erro (Carneiro et al., 2021).

De facto, para proceder à correção de falsas memórias, os erros devem ser primeiramente detetados e compreendidos para que possam ser corrigidos (Mullet & Marsh, 2016). Na literatura, o paradigma DRM (Deese-Roediger-McDermott) é frequentemente usado para gerar falsas memórias de natureza semântica/associativa e estudar a correção do erro (e.g., Carneiro et al., 2021). No entanto, a manipulação experimental apresenta baixa validade ecológica por ser dificilmente relacionável com eventos do dia-a-dia. Por outro lado, o paradigma das Inferências Pragmática contempla o conhecimento geral adquirido como fator relevante no processamento de frases que sugerem informação extra ou implícita para além da expressamente apresentada (Brewer, 1977). A título de exemplo, a frase "*O bebé esteve acordado toda a noite*" poderá sugerir que "*O bebé chorou toda a noite*" embora a frase original não apresente a informação "*chorou*". Este paradigma é mais ecológico para o estudo de falsas memórias, já que envolve componentes episódicos e semânticos, e permitiu a Maraver e colegas (2022) demonstrar o efeito facilitador das tarefas de recuperação (vs tarefas de reestudo passivo) na correção dos erros.

Apesar do crescente corpo de pesquisa que apoia um uso ativo da correção de erros na aprendizagem, alguns autores defendem uma abordagem de aprendizagem sem erros. Estes últimos contestam que os erros devem ser evitados para desencorajar a sua ativação na competição pela resposta correta. Para estes autores, o feedback é visto como uma ferramenta de reforço social positivo para quando as respostas são dadas corretamente e nunca com uma orientação corretiva (Metcalfe, 2017). Em contraste, a abordagem de aprendizagem com erros introduzida por Izawa (1970) encoraja a comissão dos erros quando estes são relacionados com a informação pretendida (Grimaldi & Karpicke, 2012; Huelser & Metcalfe, 2012). A literatura não só tem mostrado que a geração de erros facilita a recuperação de memória (e.g., Carneiro et al., 2021; Maraver et al., 2022; Metcalfe, 2017) como tem sugerido que alunos em processo de aprendizagem exerçam um controle mais ativo sobre as informações recebidas (Markant et al., 2016).

Tendo em conta que os adolescentes se encontram numa situação escolar ativa e numa fase de desenvolvimento cognitivo e neurológico, são uma população interessante para o estudo da correção dos erros de memória. É de salientar que a correção dos erros envolve funções executivas como a memória do trabalho, o processamento inibitório e a monitorização do erro, que se encontram em processo de maturação na adolescência até ao início da fase adulta (Ferguson et al., 2021). A literatura tem verificado que os adolescentes têm um desempenho inferior ao dos adultos em tarefas que envolvem monitorização, inibição, flexibilidade cognitiva e capacidade de decisão (e.g., Bunge & Wright, 2007; Luna et al., 2015), embora haja pouca evidência sobre o impacto do desenvolvimento do controlo cognitivo na recuperação de memória episódica (Ofen et al., 2007). No entanto, isto pode dever-se à escassez de estudos com a população adolescente e, até onde sabemos, ainda nenhum estudo foi feito sobre a correção de erros de memória com esta população.

Adotando uma abordagem de aprendizagem com erros, o nosso estudo pretende estudar o efeito da recuperação ativa ao longo do desenvolvimento com adolescentes (12-17) e jovens adultos (22-27 anos) utilizando as frases adaptadas para Português de Carneiro e colegas (2020) num paradigma de Inferências Pragmáticas. Para isto, os participantes (n=82) realizaram uma tarefa de memória com as três seguintes fases: 1) fase de codificação, quando os materiais eram memorizados; 2) fase intermédia, quando os participantes decidiam se uma determinada frase já tinha sido apresentada anteriormente (reconhecimento ativo) ou lhes era apresentada a decisão de outro

participante (reconhecimento passivo), com feedback corretivo apresentado posteriormente independentemente da tarefa; e 3) teste de memória de evocação por pistas, quando os participantes completavam as frases anteriormente vistas. Para a fase intermédia, utilizámos um design de pares-emparelhados intra-sujeitos para que cada participante das duas faixas etárias fosse exposto às duas condições de reconhecimento (ativo e passivo) e para aumentar o valor ecológico do projeto com a exposição passiva à desinformação na forma de resposta incorreta de outro participante. Simplificando, na condição ativa, os participantes recuperaram ativamente informações previamente apresentadas na fase de codificação e decidiram se uma determinada frase já tinha sido apresentada anteriormente. Na condição passiva, os participantes são meramente expostos à resposta de outro participante, que pode ser tanto correta como incorreta. Com a apresentação de feedback corretivo após cada frase na fase intermédia, espera-se que jovens adultos tenham uma proporção de correção de erros significativamente maior quando as frases foram recuperadas ativamente do que quando foram lidas passivamente, conforme resultados de Maraver et al. (2022). Da mesma forma, para adolescentes, hipotetizamos que a recuperação ativa também poderá ser benéfica para a aprendizagem em comparação com o reconhecimento passivo.

Os resultados partiram da análise das frases apresentadas no seu formato indutor de inferências pragmáticas e demonstraram que tanto os adolescentes como os jovens adultos corrigiram mais erros quando as frases correspondentes tinham sido apresentadas na condição de recuperação ativa na fase intermédia, contrastando com aquelas que tinham sido simplesmente vistas. Não verificámos diferenças significativas entre os dois grupos etários ou efeitos de interação na correção dos erros. Do mesmo modo, um índice de controlo sobre a persistência das respostas corretas da fase intermédia até à fase final mostrou que os adolescentes e jovens adultos têm uma maior persistência das respostas corretas na condição de reconhecimento ativo do que passivo, não havendo diferenças entre os dois grupos etários. Mais interessante ainda, um índice de correção do erro (da fase intermédia para a fase final) revelou uma maior capacidade de correção do erro para frases da condição ativa do que passiva, em ambos os grupos etários.

Os resultados sugerem que tanto para os adolescentes como para os jovens adultos existe um efeito benéfico da recuperação ativa de informação para a correção de erros de falsas memórias. Tendo em conta que ambos os grupos etários se encontram em fases de vida que contemplam o contexto educativo (e.g., escolaridade básica e superior), isto pode

traduzir-se na realização de exercícios de carácter exploratório antes do ensino mais estruturado dos materiais, que deve incluir feedback corretivo sobre os erros gerados. A ausência de diferenças comportamentais entre as populações em estudo revela que os adolescentes são capazes de beneficiar de estratégias ativas de recuperação (relativamente a estratégias passivas). Em estudos futuros será importante perceber quais as funções cognitivas (e.g., monotorização, controlo inibitório) subjacentes a este efeito em cada grupo etário.

Palavras-chave: recuperação de memória, correção do erro, aprendizagem, desenvolvimento, adolescentes

Introduction

Every day, we are bombarded with stimuli we instinctively try to commit to memory to better perform in the future. Nevertheless, we still incorrectly recall information daily, which lays the ground for the conflict between false memories and the original event. It becomes thus critical to understand how false memory errors can be corrected and investigate the benefits of making errors during learning. In educational contexts, it is especially relevant to raise awareness of the facilitative role of errors in the process of consolidating recently learned materials into long-term memory. Moreover, adolescents might particularly benefit due to their societal duty of undergoing basic education.

Therefore, in this introduction, I will start by exploring what false memories are and how they might be induced. Afterwards, I will focus on how they can be corrected and elaborate on the role of active retrieval. Finally, I will explain the errorless vs errorful learning theories debate in learning contexts and expose the necessity of experimenting with false memory correction in the adolescent population for both investigational and practical purposes.

The Formation of False Memories

Memory is in a constant malleable state, which allows its reconstruction upon recall of information (Bartlett, 1932). This is the thesis that has revolutionised the idea that memory is merely "storage" in which perceived experience of a situation (i.e., schemas) is impressed on and can be later accessed from, as postulated by the Trace Theory of Memory (Danziger, 2008). Indeed, this malleability helps explain why we do not remember things exactly as we experienced them. After all, how many times do we find inconsistencies in the retelling of an event from different people? Or misremember where we placed an object? This mismatch in remembering evidences the growing-studied phenomenon of false memories and the distortion of the mental constructions of our experienced reality.

In this vein, it is important to elaborate on what is considered a false memory. It is not the retelling of additional information over another person's version of events or a differently positioned perspective that constitutes a false memory. Rather, it is the

inaccuracy of the memory in comparison to the real events. For instance, in clinical sessions, it is clear that our recall of autobiographical memories can be deceived and not match our past experiences, once revised (Loftus, 1997). In forensic settings, it is common to screen for confabulation (i.e., the involuntary production of erroneous memories) to verify the testimonies' probabilistic validity with the cross-checking of audios, videos and/or physical evidence that better showcase the accuracy of events. However, while confabulation might indicate a memory disorder, misremembering errors in our daily lives are normative in healthy populations (Robins, 2019). The difference between the two types of false memories might stem from the influence, or lack thereof, of the information retained from the past event; with misremembering relying on the distortion of that information and confabulation creating whole inaccuracies (Robins, 2016).

Indeed, memory is often distorted to seek coherence between the gist of the past experience and the current knowledge and context from which we are remembering (Carneiro et at., 2020). For example, when trying to remember where an object was placed, an attitude is set up and directed from our present environment to the previous event, creating a dual environment in which we try to self-reflectively identify distinctions in the routine leading to the object being put down, within the duration of the experience (Barlett, 1932; Wagoner, 2013). Distortions in this process might lead to vividly remembering having placed the object in a verifiably inaccurate place with specific details (Chan & McDermott, 2006; Roediger & McDermott, 1995) and high confidence (Loftus et al., 1978; Sampaio & Brewer, 2009).

Interestingly, Paradis et al. (2004) found that, in some cases, people maintain confidence in the report of erroneous elaborated details even when confronted with evidence, which demonstrates the sustainability of false memories. These lapses in episodic memory might even be further maintained by the self-centred recall of the emotional involvement in the event (Wade et al., 2018) but are not limited to first-hand events (Pillemer, 2009). In fact, studies on flashbulb memories - formed around the circumstances when one is told of a public event by an external source (Brown & Kulik, 1977) -, show that vividness and confidence in recall persist over time even though consistency declines, as expected in regular autobiographical memories (Talarico & Rubin, 2003). Bayen and colleagues (2000) showed in their study that when participants do not remember the circumstances in which they came by the required information, they guess that it was presented by the expected source, basing their decision on prior

knowledge to account for their source monitoring uncertainty. This supports the constructivist memory theory, in that people continuously reconstruct their memory of the event, albeit with the memory distortion that underlies the failures in identifying the sources of the mental experience (Johnson et all., 1993).

Loftus' Misinformation Paradigm for eliciting false memories showed that memories of an original event are more susceptible to misinformation, the more time has passed since it took place (Loftus et. al., 1978). In the study, participants were exposed to misleading information when asked if another car had passed by while a red Datsun was stopped at either a *stop* or *yield* sign, after being shown a series of slides picturing the movement of the red Datsun until its impending stop at an intersection. A final forced-choice recognition test presented participants with slides with either the stop or yield signs and tasked them with deciding which had been originally shown. Results revealed that only 41% of participants (N=1232) responded accurately when the question contained misleading information, which evidenced the majority of participants' inability to detect discrepancies between the original and later presented slide. This memory distortion via explicit external influence demonstrates the Misinformation Effect (Loftus, 2005) and stresses the importance of inducing erroneous memories in the study of memory reconstruction and malleability. The study also supports the Discrepancy Detection Principle (Tousignant et al., 1986), which informs that changes to the original event are more likely to occur should these discrepancies to the post-event not be detected soon after exposure.

According to Chan and collaborators (2009), an immediate recall test facilitates the incorporation of misinformation. In fact, retrieval during an initial test has been shown to enhance suggestibility in final tests of free recall (Wilford et al., 2014), forced-choice discrimination (Thomas et al., 2010), cued-recall (Chan et al., 2009) and source discrimination (Chan et al., 2012). Finn (2017) proposed that memory updates after retrieval (MUAR) through memory reactivation. According to this framework, information is encoded before undergoing the process of consolidation that stores it in memory, where it is later retrieved. Additionally, it can also go through a reconsolidation process when information matched to the retrieval cues is introduced after retrieval. During this process, memory is malleable and may thus be updated in accordance with the verifying or contrasting new information presented that in turn enhances or hinders

learning. The MUAR framework is congruent with the enhanced retrieval suggestibility (ERS) effect.

In this vein, false memories may be induced due to the semantic similarity between the listed and related critical words, which is understood to facilitate both correct and false recognition at retrieval (Gunter et al., 2005). Importantly, the DRM (Deese-Roediger-McDermott) is one of the most frequently used paradigms to study associative/semantic false memories. It does so by having participants study lists of associative words (e.g., fruit, strawberry, apple) that lead them to recall semantically related words that were not present (e.g., pear) (Deese, 1959; Roediger & McDermott, 1995). However, this manipulation is artificial, since the experiment is hardly relatable to daily events and is manipulated to cause participants to respond in particular ways to the stimuli presented (Wagoner, 2013) in search of an "efficient causality" (Harré, 2002). Using the DRM paradigm with this framework, Carneiro and colleagues (2021) demonstrated the effect of ERS and proposed that memory is updated thanks to the integration of the new information into the original event and not due to an overwriting process of the old information. The authors verified that retrieval facilitated the integration process of new related information that could either be false and lead to false memories, or it could be true in the form of corrective feedback and update erroneous memories when presented immediately after each error (Carneiro et al., 2021).

Like other word-association paradigms, the DRM relies on the semantic link between the cue and target words to generate false memories, which implies that memory errors are also vulnerable to internal influences even when erroneous information is not supplied (e.g., Van Damme & Smets, 2014). As Bransford and Franks (1971) demonstrated, people integrate separately presented information into holistic ideas if it is semantically related. Indeed, be it with semantic or phonological similarities to a non-presented word (e.g., Chan, McDermott, Watson, & Gallo, 2005; McDermott & Watson, 2001; Sommers & Lewis, 1999) or the effect of emotional arousal activation to the presented event (Van Damme & Smets, 2014), literature has increasingly evidenced the role of these internal processes in the generation of incongruent information with that presented in encoding phases.

Critically, general world knowledge plays a pivotal role when processing information and can therefore be responsible for distortions during memory reconstruction (Sulin & Dooling, 1974). Most relevant to the present paper, pragmatic

inferences are easily found in everyday events and are an illustrative example of how memory distortions accommodate general semantic knowledge. Notably, these inferences can sometimes be favourable by capturing the speaker's intended meaning when it is poorly conveyed (McDermott & Chan, 2006).

Testing with Pragmatic Inferences

The Pragmatic Inference Paradigm conjointly involves semantic and episodic components during the processing of sentences that suggest additional information than what is explicitly meant or implied to be inferred from the information provided (Brewer, 1977; Carneiro et al., 2020). For instance, from hearing that "*The unskilful skateboarder lost his balance on the skateboard*", one could assume that "*The unskilful skateboarder fell off the skateboard*" and remember this later version as the originally heard (sentences from McDermott & Chan, 2006). This paradigm is more robust and ecologically valid for the reason that it allows for the study of false memories with materials that more closely resemble scenarios in social interactions, in which inferences responsible for memory distortions are frequently assumed. These memory distortions are considered false memories being that the original meaning is changed.

Crucially, pragmatic implications exclude sentences that are semantically synonymous, keep the implied logic, and whose pairs have no relation to each other. Conversely, a 'but not' conjunction is used to verify the aptitude of sentences with pragmatic implications in research due to its separative and contrasting nature, by applying it to the sentence resulting of the addition of both the sentence under consideration and its version with the inference stated: "*The unskilful skateboarder lost his balance on the skateboard* 'but did not' fall off of it".

Notwithstanding, Brewer (1977) demonstrated through a qualitative analysis that sentences with pragmatic implications can elicit ambiguous inferences from participants with cued-recall. For example, in his study, the original sentence "*The absent-minded Professor didn't have his car keys*" elicited inferences like "*forgot* his car keys," "*lost* his car keys," or "*left* his car keys in the car" from different participants. Furthermore, the author understood a higher level of pragmatic inferences recalled (26%) compared to correct retrievals (19%) to demonstrate the strong interaction between long-term

knowledge and episodic memory. These results also highlight the impact of the final test format on the recalling process.

A two-alternative forced-choice recognition task revealed that participants often opt for the pragmatic implication sentence even when presented with the correctly worded sentence (McDermott & Chan, 2006). However, Sampaio and Brewer (2009) showed that pragmatic inferences were retained with greater confidence in a simple recognition task than in a two-alternative forced-choice recognition task, most likely because participants in the latter task were more conscious of the misleading nature of certain items inferentially implicated.

Using a simple recognition task, Chan and McDermott (2006) had participants decide if given sentences in their original (i.e., with pragmatic implications; e.g., "*The new baby stayed awake all night")*, inferred (e.g., "*The new baby cried all night*") or inconsistent (e.g., "*the new baby slept all night*") formats had been previously presented. Critically, sentences at study had also been presented in either their original or inferred versions. The authors found that the proportions of accurate and erroneous recognitions were quite high (.73 and .65, respectively) when sentences were presented in their implication form, suggesting that pragmatic inferences are recognised to a similar degree as exact correct phrases, even though pragmatic implications are not fully determined by the original sentences (Brewer, 1977). In fact, one's long-term knowledge of the world's intricate link to language comprehension and memory processes allows sentence memory tasks to construct new semantic descriptions at recall (Bransford et al., 1972; Brewer, 1977).

In this respect, Carneiro and colleagues (2020) conducted a series of studies, in which recognition and cued-recall tasks were used, to determine the effect of language and culture in the testing materials for pragmatic inferences. The authors translated McDermott and Chan's (2006) English sentences with pragmatic implications to Portuguese and found that the number of inferences acquired in each language for the same sentence did not significantly correlate with one another, as the sentences that elicited the highest and lowest levels of false recall were not the same. Thus, in order to obtain the specific ranking of sentences for a given population, researchers are encouraged to localise their sentences with pragmatic implications in relation to culture/language for the study of false memories with pragmatic inferences.

Correcting False Memories

As mentioned above, false memory errors persist through time with high confident levels (Paradis et al, 2004). Interestingly, error sustainability endures when participants are warned of their likelihood before (McDermott & Roediger, 1998) and after (Greene et al., 1982) the study phase. Moreover, errors persist even when participants signal each false memory for posterior recall (McConnell & Hunt, 2007). As evidenced by the misinformation effect (Loftus et. al., 1978), the incorporation of erroneous information distorts memory, however, post-event corrective information can update erroneous memories (Carneiro et al., 2021).

Indeed, merely pointing out to participants that a response is incorrect does not facilitate its correction (Pashler et al., 2005). Errors must first be detected and understood as such so they can be corrected (Mullet & Marsh, 2016). For feedback to be integrated into memory, it should be well-associated with the goals of the retrieval event and should be overtly included in the correct answer (Pashler et al., 2005; Mullet & Marsh, 2016), as well as paid due attention and understood in order to take a corrective form (Metcalfe, 2017). The literature well documented that corrective feedback optimises error commission and that errors are unlikely to be corrected spontaneously without it (Butler et al., 2008). Conversely, the timing of its presentation is not yet established as a decisive factor, nor is it understood if corrective feedback is only effective during a specific time window (see Mera et al., 2021).

While most studies support the view that corrective feedback should be presented shortly after error production (e.g., Grimaldi & Karpicke, 2012; Hays et al., 2013; Vaughn & Rawson, 2012), some have suggested that for more elaborated and meaningful study material (e.g., trivial questions) feedback can be further delayed and still be effective (e.g., Kornell, 2014), likely due to activation of a richer semantic network that underlies a more complex retrieval (Mera et al., 2021). Crucially, Mullet and Marsh (2016) conducted a study with the Pragmatic Inferences Paradigm and found that immediate feedback motivated a higher reduction in the proportion of errors produced, even though both immediate and delayed feedback were effective.

The confidence with which we make errors also plays a role in memory error correction. The Hypercorrection Effect stands for the finding that high-confidence errors are more easily corrected than low-confidence errors when followed by corrective

feedback (e.g., Butterfield & Metcalfe, 2001; Fazio & Marsh, 2010). Moreover, Butterfield and Metcalfe (2006) showed that correct answers whose previous error's confidence level had been high, were sustained over time. This effect could be influenced by the surprise felt at being confronted with unexpected feedback, especially if familiar with the semantic domain of knowledge (Metcalf, 2017). Interestingly, incorrect guesses are less likely to be corrected on a second test of the material than high-confidence errors (e.g., Butterfield & Metcalfe, 2001; Fazio & Marsh, 2010).

Remarkably, other than confidence levels, the testing design also plays a role in error correction.

The Generation Effect shows that generating information in response to a cue leads to better memory than simply reading it (Slamecka & Graf, 1978), even when opposed to the mere presentation of feedback with no corrective intent (Grimaldi & Karpicke, 2012). This may be due to the activation of elaborate information related to the target word that could lead to an increased probability of retrieval of the correct answer (Carpenter, 2009).

A study by Kornell and collaborators (2009) using the Weak Associative Paradigm to provoke a high number of guessing errors, asked participants through a number of trials, to guess the word (target) that followed a given cue word (i.e., pre-testing). As their goal was to study error correction, the correct target word (e.g., playground) bore a weaker association to the cue (e.g., garden) than participants' frequent answers (e.g., flowers) so as to generate errors. Additionally, the authors tested error correction following guess-making and read-only tasks, followed by corrective feedback. As the Testing Effect (Karpicke & Roediger, 2008; Roediger & Karpicke, 2006) postulates, testing the material in the encoding phase (when participants study the materials) is more advantageous than simply re-studying it, and the authors' findings were congruent with this effect. Indeed, participants presented higher correction rates when the words had been actively guessed with cued-recall than when they had only been re-read, which demonstrates that testing the material before a final test may enhance long-term memory retrieval. Additionally, unsuccessful retrievals (i.e., errors) facilitated error correction in both conditions, which supports the view that error commission may be favourable to memory consolidation and, in another word, learning. In spite of these results, it should be stressed that these errors are not representative of daily memory errors

in healthy populations and that literature on the testing effect has focused mainly on the adult population and predominantly studied delayed tests (Roediger & Karpicke; 2006).

With the Pragmatic Inferences Paradigm, Maraver et al. (2022) demonstrated that retrieval tasks have a more permeable effect on error correction than passive restudy tasks, which corroborated the Testing Effect. Moreover, the authors showed that corrective feedback after retrieval leads to better memory performance and subsequent error reduction compared to passive recognition. Consistent with the MUAR framework above mentioned (Finn, 2017), error correction was facilitated by corrective feedback due to memory's characteristic malleability and vulnerability to the incorporation of this new information presented after errors are retrieved. That is, through a memory reconsolidation process, previous erroneous information is updated with the correct answer provided by the corrective feedback. This framework had already been tested by Carneiro et al. (2021) using the DRM paradigm, whose results evidenced that reconsolidation allows the integration of both false and correct (feedback) information after retrieval, and thus facilitates the formation of false memories and the process of error correction, correspondingly.

The subsequent learning effect of successful/unsuccessful retrieval of previously tested information is called Test Potentiated Learning (Carneiro et al., 2020).

When Errors Benefit Learning

Despite the growing body of research supporting an active use of error correction in learning, there is still a body of research arguing that errors should be avoided. This view stemmed from the Interference Theory (Melton & Irwin, 1940; Postman & Underwood, 1973) that claimed experiencing errors enhanced their distinctiveness when competing with the correct answer at retrieval. In fact, the errorless learning assessment is consistent with the early social learning theories (e.g., Bandura, 1986; Skinner, 1953; Terrace, 1963) and supports the view that an error commission is not only a highly prominent event but that it also costs necessary attentional resources to well-perform the memory task (Houtman & Notebaert, 2013; Notebaert et al., 2009). Adding to this, Ausubel (1968) declared his opposition to the exploratory learning strategy, which involves error generation, claiming that errors encourage incorrect practice and are difficult to be overwritten to correction during the learning process. Moreover, this

errorless learning approach understands feedback as a tool of positive social reinforcement to be used when tasks are correctly executed, and not when errors are committed (Metcalfe, 2017).

Conversely, the errorful learning approach dismisses the avoidance of memory errors as a means to preserve memory with the least distortions possible and welcomes erroneous recall as part of the learning process (e.g., Kornell, 2009; Potts & Shanks, 2014). Specifically, errors must be related to the target (Grimaldi & Karpicke, 2012; Huelser & Metcalfe, 2012) and first be detected and provided with corrective feedback in order to be replaced (e.g., Carneiro et al., 2021; Mullet & Marsh, 2016; Pashler et al., 2005). Additionally, literature has presented that error generation may facilitate better memory recall than read-only tasks (Metcalfe, 2017). Markant and collaborators (2016) suggested that students who exercise active control over the information they are being given and ask questions in class benefit from the mental exercise of deciding what questions to ask as well as from the exchange of information (i.e., feedback) in the learning process.

Errorful Learning was first introduced by Izawa (1970), who contrasted trials with more and less error generation in their procedure and concluded that unsuccessful retrievals facilitated learning of the material in subsequent memory tests. In order to study this phenomenon, experimental designs usually rely on pre-testing the material so as to invite errors. However, it is noteworthy to mention that when guesses are completely uninformed, errors do not benefit learning (Kang et al., 2011). Research with didactic objects of study demonstrated the success of unsuccessful retrievals preceding study sessions with medical (Richland et al., 2009) and mathematical complex materials (Kapur & Bielaczyc, 2012) in learning. Critically, Metcalfe and Huelser (2020) suggested that episodic/explicit memory may underlie the learning process and be responsible for its success, as opposed to semantic/implicit memory.

Aiming to comprehend the benefits of Errorful Learning, four explanatory theories have been proposed, focused on the following two main variables: the level of pre-existing semantic relationship between the study materials, and whether the error is explicitly recovered on the final test (Mera et al., 2021).

The Mediator Effectiveness Hypothesis (Carpenter, 2011; Pyc & Rawson, 2010), as the name implies, views errors with the mediating function between the instruction and

the recovery of the correct answer in the final test. In fact, the error is recovered as the second mediator in the memory test, following the cue itself as the first, and should not only be self-generated but also semantically associated with the study material to prove an advantage to learning (e.g., Grimaldi & Karpicke, 2012; Huelser & Metcalfe, 2012). Moreover, this hypothesis is in line with the Testing Effect by stating that retrieval makes mediators more effective due to the strengthening of links between the cue and the target, also congruent with the Spreading Activation Theory postulates (Collins & Loftus, 1975). However, contradictions to this theory argue that participants whose guesses are more relevant and thus should behave as better mediators, do not show improved performance on the final memory test (Clark, 2016). Indeed, it is still unclear how semantically exclusive the relation between the error and target might be or if it also has an episodic component underlain (Metcalfe, 2017).

According to the Search Set Theory, the error committed is one of the various semantically-related candidates activated in the search process to recall the target, which is also activated during the process. As per Grimaldi and Karpicke's (2012) research, feedback enables reinforcement of the target and facilitates its encoding with the partially activated error. Furthermore, Kornell (2009) and Clark (2016) added to the theory the concept of error-suppression, which would be responsible for the weakening of erroneous retrieval routes (Collins & Loftus, 1975), thus facilitating correct recall and dismissing the need to explicitly recalling the unsuccessful retrieval.

The Recursive Reminding Theory (Wahlheim & Jacoby, 2013) postulates that the original unsuccessful recovery is recovered at the retrieval phase along with the correct answer. Rather than declaring necessary semantic associations, this theory suggests that episodic memory is responsible for the successful recovery in the late memory test for the reason that the error and the correct answer supplied via feedback are both encoded together in the same memorable episodic event and then jointly recalled. Notwithstanding, there are studies contradicting this proposition by showing that errors do not need to be recalled during retrieval to correctively recall the target (e.g., Butterfield & Metcalfe, 2001; Metcalfe & Miele, 2014).

Finally, the Error Prediction Theory is consistent with the hypercorrection effect and posits that when unexpected corrective feedback is received and/or a prediction is made once the error is chosen, people are more motivated to correct their answers. This may be explained by the activation of conflict resolution processes steered by the

occurring cognitive discrepancy (Decker et al., 2020), which is encouraged by the expectancy-violating error process that facilitates correct recovery (e.g., Brod et al., 2018). Notably, Metcalfe and Finn (2011) pointed out that high-confidence errors should be learnt with more difficulty than lower-confidence errors due to their greater similarity to the target. Thus, this theory still lacks a clear understanding.

Learning and Error Correction In Adolescents

Most relevant to the present work, the process of memory recognition involves cognitive and voluntary control that supports the inhibition of misleading information (Levy & Anderson, 2002). Executive functions, such as working memory and inhibitory processing, are high-level cognitive processes acquired during infancy whose evolving maturation and refinement crosses adolescence into young adulthood (Ferguson et al., 2021). The growing sophistication of derived cognitive abilities, of which planning abilities and error monitoring are examples, is associated with organic changes occurring in the brain during adolescence (e.g., synaptic pruning, increased myelination) (Huttenlocher, 1990; Wozniak & Lim, 2006). However, the different rates at which distinct cognitive control abilities develop (e.g., Huizinga et al., 2006; Zanolie & Crone, 2018) should be highlighted. Critically, the specialization and organization of these central processes present some vulnerability to impaired development (Luna, 2009).

Concerning memory error correction, maturing top-down executive processes such as response inhibition, working memory and error monitoring, may impact consistent performance in the adolescent population and thus leave them vulnerable to not only committing memory errors (Luna, 2009) but to also lower error correction efficacy.

In children, their vulnerability to making errors is likely due to their still-developing inhibitory control capabilities that increase their susceptibility to interference from distractors (Bjorklund & Harnishfeger, 1990; Luna, 2009). Concerning error correction, following Kornell et al.'s (2009) results, Carneiro and colleagues (2018) demonstrated that children older than five years of age benefit from unsuccessful retrievals. The authors used a weak association paradigm with auditory stimuli, for which children were asked to either guess the associated word (i.e., target) that followed the first word (i.e., cue) heard or just listen to the corresponding target-associated word.

Significantly, unlike Kornell et al.'s (2009) study, participants were exposed to incorrect information and received corrective feedback in both active production and passive study task conditions. Findings not only supported the errorful learning theory but also evidenced the role of active control in error correction from childhood, to which the authors suggested complementing with studies employing episodic retrievals of previous knowledge in the future.

Due to children's characteristic developing prefrontal participation, the consequential development of executive functions is easily observed in the behavioural shift between childhood and adolescence (Luna et al., 2004), when behaviour becomes increasingly prefontally mediated (Fassbender et al., 2004; Fuster, 1997). During adolescence, performance in more complex tasks is improved due to the enhancement of control distraction that results from more efficient and adaptable working memory (Luna, 2009), although the maturation process of the underlying cognitive abilities is not clearly distinguishable in performance with other age groups. Nevertheless, literature has found that adolescents underperform adults in activities requiring monitoring, inhibition, cognitive flexibility and decision-making abilities (e.g., Bunge & Wright, 2007; Luna et al., 2015). Despite this, few studies have offered solid evidence that the still-developing control abilities have an impact on episodic memory retrieval (Ofen et al., 2007). This could be partly due to the little volume of research pertaining to adolescents conducted in comparison to other age groups as well as the various and limiting age intervals that were in fact observed, in spite of it being consensual that adolescence encompasses twelve to seventeen-year-olds (Spear, 2000). Crucially, the striking lack of studies regarding memory recollection in adolescents uncovers a gap in the literature. Moreover, as far as we are aware, there is no study on this age group concerning memory error correction.

Current Study

The present study aims to compare adolescent and young adult performance in memory error correction. To do this, we make use of the Pragmatic Inference Paradigm (Brewer, 1977), with Carneiro and colleagues' (2020) Portuguese adaptation, to incite memory errors more representative of everyday pragmatic lapses during communication and allow for the dissociation between the semantic and episodic components present at recall. Stemming from an errorful learning perspective, we expect to confirm previous

results and verify a significant error correction effect in the memory test in young adults after generating the error and being provided corrective feedback (Maraver et al., 2022). Additionally, due to their maturing executive processes, particularly inhibitory control, working memory and error monitoring, we hypothesize that adolescents might have lower error correction efficacy in comparison to young adults.

To understand the process of error correction, we focused on the role of retrieval and compared the performance of adults and adolescents in active vs. passive recognition tasks. Young adults and adolescents participated in an online study that required performing a memory task with the three following phases: 1) encoding phase, when materials were to be studied; 2) intermediate phase, when participants were to either decide if a given sentence had been previously presented (active recognition) or see another participant's decision (passive recognition), in which case corrective feedback followed; 3) final cued-recall memory test, when participants were to complete the encoded sentences as they retrieved them from memory. For the intermediate phase, we employed a within-subjects yoked-pair design so that each participant of the two age groups was exposed to the two conditions and also to increase the ecological value of the design by including passive exposure to misinformation in the form of another participant's incorrect response in the active condition. To put it simply, in the active condition, participants actively retrieved information previously presented in the encoding phase and made a decision on whether a given sentence has been previously shown. In the passive condition, participants are merely exposed to another participant's response, which might be correct or incorrect. With the presentation of corrective feedback after each trial in this intermediate phase, young adults are expected to have a significantly higher error correction rate when sentences were actively recalled than when they were passively read, as per Maraver et al.'s (2022) results. Similarly, for adolescents, we hypothesize that active retrieval might also be beneficial to learning in comparison to passive recognition.

Method

Participants

A total of 124 participants were recruited for this study, of which 58 were adolescents with their age comprised between 12-17 years old, and 66 were young adults aged from 22 to 27 years old. From this sample, 42 had to be excluded due to several reasons: not speaking European Portuguese (N=4), having internet problems (N=1), simultaneously listening to music (N=1), taking an excessive amount of time to complete the experiment (N=1), having been interrupted (N=7) or having been sent a wrong link (N=25).

The final sample included a total of 82 valid participants: 41 adolescents (20 F; *M*.age = 14.46; *SD*.age = 1.62) and 41 young adults (32 F; *M*.age = 24.1; *SD*.age = 1.61). The sample size was determined using G*Power software (Erdfelder et al., 1996) based on the effect size of error correction observed in Maraver et al. (2022). The estimated sample size was 78 valid participants, which we aimed to increase to guarantee this minimum considering the 8 counterbalance versions of the materials and the limitations of online data collection.

Prior to the start of the experimental session, participants received information about the study and adult participants provided consent for their participation as did adolescent participants' guardians for their charges, in accordance with the Declaration of Helsinki (World Medical Association, 2013). All experimental procedures were approved by the Ethics Committee of the Faculty of Psychology of the University of Lisbon. Participants were rewarded with a €10 voucher to be used in Fnac.

Experimental Design

This study is part of a larger experiment that provided more data than analysed in this paper. This experiment followed a 2 × 2 mixed design with the recognition condition (active vs passive) manipulated within-subjects, and the age group (adolescents vs young adults) manipulated between groups.

Materials

Following Carneiro et al. (2020)'s Portuguese adaptation from McDermott and Chan (2006)'s pool of pragmatic sentences, we used thirty-two of their versioned sentences to elicit the recall of pragmatic inferences (see Appendix A), similar to Maraver et al., (2022). The presentation of the sentences was carefully controlled by generating eight different counterbalance versions of the materials. The format of the sentences was manipulated across subjects, considering the retrieval task (active vs. passive), the format of the sentences presented at encoding (critical/original vs. filler/manipulated), and the format of the sentence presented at the intermediate phase considering the format at encoding (match vs. mismatch). As an example, the critical sentence is one that generates the inference. "*The baby stayed awake all night*" has its correspondent "*The baby cried all night*" which already has the pragmatic implication derived "*cried*" from the pragmatic implication "*stayed awake*", and was used as a filler. The critical sentences are crucial to studying error correction because they induce the generation of the inference and invite the memory error with its pragmatic implication. On the other hand, fillers are manipulated to have the pragmatic reference clearly stated in an effort to preserve the ecological validity of the study. In this vein, there was also a match/mismatch control in the recognition test of the intermediate phase whose format could be the same (match) or different (mismatch) as they were in encoding (see Table 1 for examples).

Table 1

Examples of the Sentences Used in the Encoding and Intermediate Phases

Encoding Phase	**Intermediate Phase**	
	Critical Form *"The baby **stayed awake** all night"*	Manipulated Form *"The baby **cried** all night"*
Critical Form *"The baby **stayed awake** all night"*	Match (16 sentences)	Mismatch (16 sentences)
Manipulated Form *"The baby **cried** all night"*	Mismatch (16 sentences)	Match (16 sentences)

Procedure

The data collection was conducted online and used Qualtrics platform (Qualtrics, Provo, UT) to prepare the experimental design and generate the individual links. Each link was generated manually, in order to create the within-subjects manipulation with a yoked-pair design, and was sent by the person tasked with contacting participants so as to preserve anonymous participation and guarantee the data confidentiality in the analysis.

Young adults were contacted through a pool of emails volunteered for the purpose of participating in studies of the faculty, whereas young adolescents were approached at schools, study centres and social groups or organisations like Escuteiros de Portugal. Critically, the links to the experiment were only sent by email once we had the consent forms and participants had one week to participate and contact us with their chosen identification number. On a side note, the participation links for the adolescent group were sent to the emails indicated by their guardians.

Upon opening the link, participants were asked to provide some demographic information (i.e., gender, age, year of birth and first language) and informed to pay attention to the study before starting the experiment. As mentioned above, this study is part of a larger experiment that provided more data than analysed in this report (see Figure 1 for a visual comprehension of what is focused on in this study).

Encoding Phase

Participants were instructed to memorise 32 sentences and solve a math question after each sentence has been presented. A trial run of six other sentences made sure participants understood the instruction. Half of the sentences were presented in their original form while the other 16 were manipulated and shown with their respective pragmatic inference. Presentation order was randomized. Each sentence was shown on the screen for 4.5 seconds.

After the task, participants worked through a five-minute distractor task in which they were asked to count the differences between four blocks of two images.

Intermediate Phase

Subsequent to this, participants underwent a recognition task that featured both active and passive recognition conditions. The instruction consisted of either a) reading a sentence and deciding if it had been presented in the encoding phase (active recognition) or b) reading the sentence and seeing another participant's decision (passive condition) depending on how the sentences were presented. Importantly, these conditions were controlled within age groups through a yoked-pair design, so that participants only saw responses of a previous participant from their age group. Overall, participants were presented with 32 randomized sentences for at least 6.5 seconds, half the sentences per condition, and counterbalanced as upper mentioned in Materials. Importantly, corrective feedback was given after each sentence and viewed for 4.5 seconds, regardless of the recognition condition.

Similarly to the previous task, after this phase participants were again asked to spot the differences between two images. This is to keep participants from restudying the correct answers after being given the corrective feedback.

Final Memory Test

Lastly, participants underwent a cued-recall task where they had to fill in the blanks and complete the sentences with the words or expressions that they recalled being shown at the beginning of the experiment. The blanks stand for the original segments that are suggestive of pragmatic implications in the critical sentences or the pragmatic inferences present in their manipulated versions, depending on the sentence counterbalance. They had one minute to respond to each sentence.

In addition to this, after each sentence was presented, participants answered a confidence rating question, ranging from 'no confidence' to 'some confidence' and the very confident appraisal 'very confident'.

Source Monitoring Test

While dismissed in our analysis for this study, this test was also administered after the memory test to explore the performance of both age groups on their ability to recall if

the 32 sentences presented had been given a response to or had been merely observed in the intermediate phase. The same confidence rating was conducted after presentation of each sentence, as done prior to this task.

Having concluded the experiment, participants evaluated their level of attention to the study and the trustiness of their answers in a final questionnaire before contacting our responsible team member and sharing their participation code. These codes enable us to identify each participant's data and validate the experiment's conclusion so that they can be sent their rewarding voucher thereafter.

Figure 1

Experimental Design With Highlight To The Focus Of This Paper (bordered with continuous lines and coloured in standard black)

Encoding Phase	**Intermediate Phase**		**Final Memory Test**	**Source Monitoring Test**
	Task (within participants)	Feedback		
16 critical sentences + 16 fillers *"The baby stayed awake all night"* + 25 + 4 =	Active recognition *"The baby cried all night"* Correct or Incorrect? Passive recognition *"The baby cried all night"* Correct – Incorrect	*"The baby stayed awake all night"*	Final cued recall *"The baby _____ all night"* Confidence rating *"How certain are you of your response?"* Nothing/Some/Very	Source Monitoring *"The baby cried all night"* I read a response/ I gave the response Confidence rating (as in the Final Memory Test)

Statistical Analysis

The statistical analysis was performed using the 26th version of IBM SPSS Statistics software (IBM Corp., Armonk, NY, USA).

The criteria for judging participants' responses at the final cued-recall test follows Brewer's (1977) scoring procedure. Responses are scored as correct when the participants' response matches the encoded sentence (e.g., for the sentence "*The baby stayed awake all night*", the participant responds "*stayed awake*" in the cued-recall test). Responses can also be judged as pragmatic inference errors when participants recall pragmatic implications (e.g., participants respond with "*cried*"), intrusions when participants recall words that convey a different meaning to the sentence (e.g., "*slept*") or omissions when there is no response given. The proportion of each response type was calculated, resulting in four dependent variables.

In order to study error correction, the analysis focuses on the sentences that were presented in their critical format (i.e., sentences containing pragmatic implications that generate the errors). Particularly, the dependent variable in study is participant's responses in the final cued-recall memory test for the 16 critical sentences that were presented at encoding., while the manipulated variables are the age group (i.e., adolescents and young adults) and the type of retrieval in the intermediate phase (i.e., active retrieval and passive recognition). The latter was analysed for each type of response in both age groups. Moreover, we calculated an index of persistence of correct responses as a measure of control in addition to the index of error correction by comparing the proportion of correct responses at the final cued-recall test as a function of the type of response (hit or error) at the intermediate phase.

Results

An Independent Samples Test revealed no differences in the proportion of errors made at the intermediate phase (active sentences) between adolescents (M = .52, SD = .19) and young adults (M = .55, SD = .16), t(79.945) = -.30, p = .762, d = -.16, in critical sentences. Note that for sentences presented in the passive condition no response was required and hence performance cannot be analysed.

Response Type

For each type of response in the final test (i.e., correct responses, pragmatic inference errors, intrusions and omissions) given to the 16 critical sentences, we

calculated a Repeated Measures ANOVA with the active and passive recognition conditions as the within-subjects variables and the age group as the between-subjects factor. Mauchly's test for sphericity is not relevant since we are only measuring two factors for the within-subject variable and therefore sphericity is assumed for all response types. Each analysis' descriptive means and respective standard deviations for adolescents/young adults and active/passive recognition task conditions can be found in Table 2.

For the correct responses, we found a main effect of the recognition task, considering the active and passive conditions, $F(1,80) = 27.68$, $p < .001$, $\eta 2\ p = .257$, with correct responses in the final memory test when the critical sentences had been presented in the active ($M = .64$; $SD = .24$) compared to the passive ($M = .52$, $SD = .23$) recognition condition at the intermediate phase. Results did not show a main effect of the age groups, $F(1,80) = .10$, $p = .751$, $\eta 2\ p = .001$, nor did they show an interaction effect between the task conditions and the age groups, $F(1,80) = 1.55$, $p = .216$, $\eta 2\ p = .019$. This means that adolescents and young adults did not present significant differences concerning the proportion of correct responses given and that adolescents' correct response rates did not significantly differ in the active and passive task conditions in comparison to young adults' performance.

For the pragmatic inference errors, we also verified a main effect of task, $F(1,80) = 21.93$, $p < .001$, $\eta 2\ p = .215$. This indicates that pragmatic inference errors were significantly lower in the active ($M = .26$, $SD = .19$) than passive ($M = .35$, $SD = .21$) recognition conditions. Pragmatic inference errors did not show a significant effect of age, $F(1,80) = .024$, $p = .878$, $\eta 2\ p = .000$. Similarly, we did not find an interaction effect between the recognition task conditions and the age groups, $F(1,80) = .02$, $p = .880$, $\eta 2\ p = .000$.

Concerning intrusions and omissions, no significant differences were found pertaining to task effects ($p = .330$ and $p = .291$, respectively) or interaction effects with the age group factor ($p = .121$ and $p = 1.000$, correspondingly).

Table 2

Means and Standard Deviations of the Four Dependent Variables (Correct Responses, Pragmatic Inference Errors, Intrusions, and Omissions) Across the Four Experimental Conditions (Active and Passive Recognition in Adolescents and Young Adults)

	Active Recognition		Passive Recognition	
	Adolescents	Young Adults	Adolescents	Young Adults
Correct Responses	.66 (.22)	.62 (.27)	.51 (.25)	.53 (.21)
Pragmatic Inferences	.26 (.19)	.26 (.19)	.35 (.21)	.36 (.21)
Intrusions	.07 (.08)	.11 (.11)	.11 (.11)	.10 (11)
Omissions	.01 (.04)	.01 (.05)	.02 (.07)	.02 (.05)

Persistence of Correct Responses and Error Correction Indexes

We used Repeated-Measure ANOVAs to determine how correct responses persisted from the intermediate phase to the final memory test and how previous errors in the intermediate phase were later corrected in the cued-recall task. Indeed, as our goal is to study error correction, we calculated indexes to check for differences in these aspects pertaining to both age groups and the recognition conditions employed in the intermediate phase.

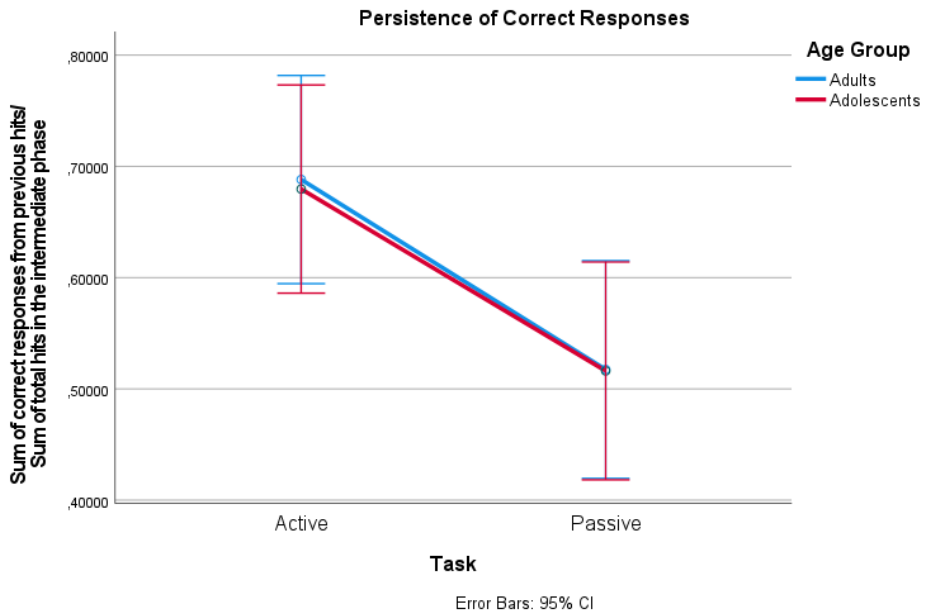
As in Maraver et al. (2022), the Persistence of Correct Responses Index was calculated as a measure of control for error correction and was derived from the sentences that were correctly recognised/presented and persisted so in the final cued-recall. It was computed as the sum of the correct responses in the final memory test that had already been correctly recognised or shown in the intermediate phase divided by the sum of the sentences in the intermediate phase that were correctly recognised or presented for viewing, as per the recognition conditions (see Figure 2, upper panel). Our analysis showed a main effect of recognition task, $F(1,80) = 18.40$, $p < 0.001$, $\eta 2$ p = .187, which revealed that corrected responses given in the active recognition condition (M = .68, *SD*

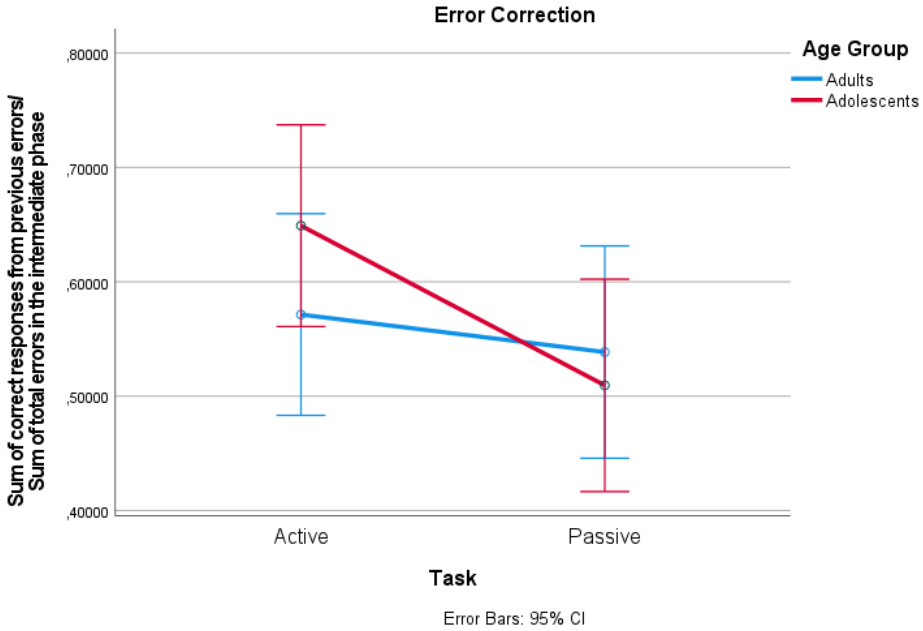
= .30) persisted more to the final test than correct responses viewed in the passive condition (*M* = .52, *SD* = .31). No age group effect was verified, *F*(1,80) = .01, *p* = .930, η2 p = .000, nor did we find an interaction effect between task and age groups, *F*(1,80) = .01, *p* = .926, η2 p = .000.

Next, we derived an Index of Error Correction based on the sentences that had previously generated errors in the intermediate phase and were later correctly recalled in the final memory test. The calculus for this index followed the sum of correct responses in the final test that had been formerly been incorrectly recognised or presented in the intermediate phase divided by the sum of total errors made in the intermediate phase in either recognition condition (see Figure 2, bottom panel). Our results showed a main effect of task, *F*(1,80) = 5.71, *p* = .019, η2 p = .067, which indicated that participants corrected errors in the final test significantly better when the corresponding critical sentences had been actively recalled in the intermediate phase than when they were passively recognised. We did not find a main effect of age, *F*(1,80) = .21, *p* = .650, η2 p = .027, nor an interaction effect, *F*(1,80) = 2.19, *p* = .143, η2 p = .027.

Figure 2

Mean Persistence of Correct Responses and Mean Error Correction in the Active and Passive Recognition for Adolescents and Young Adults





Discussion

Being part of a larger experiment, the present study aimed at analysing how adolescents and young adults correct false memory errors generated by pragmatic inference sentences. Our results showed that both age groups benefit from active retrieval to correct false memory errors. Both displayed a higher proportion of correct responses and a lower rate of pragmatic inferences following the active vs. passive recognition conditions, replicating Maraver et al. (2022). Our findings thus replicate previous results for an adult population (Maraver et al., 2022) and extend them to adolescents. Furthermore, the persistence of correct responses index evidenced a robust testing effect for both age groups.

The calculated Index of Error Correction evidenced a significant main effect of the task, revealing that active retrieval of information (compared to passive retrieval) improves memory accuracy from the intermediate phase to the final cued-recall test. This means that even when participants incorporate pragmatic errors into the information encoded in the intermediate phase they are able subsequently update said information to retrieve the correct response in the final cued-recall memory test. This goes in line with the Testing Effect and, more specifically, with Finn's (2017) proposal that Memory Updates After Retrieval (MUAR). When corrective feedback is provided, active retrieval

promoted error correction in contrast to passive recognition, suggesting that during retrieval memory becomes more amenable to change, facilitating memory update and the incorporation of correct information. More generally, these findings evidence the role of retrieval in memory malleability and vulnerability to new information. Critically, these cognitive processes seem to be similarly engaged by young adults and adolescents.

In order to control the correct answers in the final memory test whose sentences had already been correctly recognised in the intermediate phase, we calculated an Index of Persistence of the Correct Responses. Our main findings indicated that both age groups had a higher proportion of persisting correct responses in the final test when they had been previously presented in the active recognition condition (rather than in the passive condition). This evidences a robust testing effect for both age groups. Interestingly, Maraver et al.'s (2022) used a similar experimental design aiming at studying the effect of feedback and retrieval on the correction of memory errors stemming from pragmatic inferences. The authors calculated an Index of Persistence of Pragmatic Inference Errors and demonstrated that error retrieval followed by corrective feedback does not make errors more prominent, persistent and difficult to correct, as would be expected by the Errorless Learning perspective, and an Error Correction Index corroborated the main of effect of task we found.

Indeed, our study results meet our hypothesis and corroborate Maraver et al.'s (2022) findings by verifying a significant error correction effect in the memory test of young adults. However, in spite of showing adolescents' similar favouring effect of active recognition when contrasted with passive recognition, we did not find evidence for error correction to be modulated by our chosen age groups. Indeed, the absence of behavioural differences between the populations at study was predicted by Ofen et al. (2007), who warned that few studies in the literature have presented clear evidence for the impact of still-developing control abilities on episodic memory retrieval. Still, these findings only reinforce the need to further comprehend the maturation process of the cognitive functions underlying the process of error correction in different age groups. Moreover, not only is the developmental factor in need of further investigation, but it is also crucial to determine the cognitive functions (e.g., monitoring, inhibitory control) implied in the process of memory distortion and memory reconsolidation.

Remarkably, consistent with the continuous biological change and characteristic functional maturation of adolescence (Ferguson et al., 2021), a longitudinal study entitled

the Adolescent Brain Cognitive Development study (ABCD) is currently underway in the United States, having recruited 10000 youths from the starting ages of 9-10 years old. It is led by a consortium of developmental cognitive neuroscientists and, among its goals, is to define the factors of cognitive development related to the variability in behavioural trajectories (see Jernigan et al., 2018). This will contribute to the current research on developmental cognitive changes in adolescence.

Crucially, our analysis demonstrated a robust Testing Effect, which is congruent with the Errorful Learning approach to learning, by evidencing the facilitative role of unsuccessful retrieval in the process of error correction, when corrective feedback is provided. In educative contexts, Errorful Learning is often translated to active learning methodologies, which may include false starts induction and the use of flashcards. The first is typically used in Japan, where errors are encouraged by asking students to first try to solve and understand materials on their own before teachers direct a discussion of their attempts that includes obligatory corrective feedback and allows for errors to be ultimately corrected (Metcalfe, 2017). The second practice allows for memory errors to be updated thanks to the forced recall of material capable of generating retrieval errors and the immediate corrective feedback available on the verse of the study card. According to our findings, students should be encouraged to make errors while actively retrieving answers, before discussing the learning material and being supplied with the corrective feedback that enables error correction, through the update of the erroneous memory.

This is in line with Markant and colleagues' (2016) suggestion that students should exert more active control over the lectured information so that they better incorporate it into memory. The authors defend that formulating questions in class, for instance, involves cognitive control over what questions to ask, how to organize the question and how to confront the information received with previous doubts. In fact, for Markant et al. (2016), active control of information leads to the forming of distinct sensorimotor associations, elaborate encoding as a result of goal-directed exploration, improved coordination of selective attention and encoding, adaptive material selection based on existing memory, and metacognitive monitoring that improves memory. In spite of this, it should be noted that learning methodologies for educational contexts would be better explored in classroom settings than laboratory environments so as to increase ecological validity.

In this optic, even though the Pragmatic Inference Paradigm has greater ecological validity than most paradigms for error correction (e.g., Weak Association Paradigm, DRM, other word-associated paradigms) due to its consideration of general world knowledge, the use of day-to-day sentences and the episodic memory component underlying retrieval, it is still run in a laboratory or remote condition. Furthermore, as Metcalfe and Huelser (2020) suggested, episodic/explicit memory may be responsible for the success of the learning process, as opposed to semantic/implicit memory. However, since semantic development lasts well beyond childhood until adulthood (Luna, 2009) and the pragmatic inference paradigm engages both episodic and semantic components at retrieval, it could be of interest to further study these components along development.

The present study design was carefully controlled. The yoked-pair design employed in the intermediate phase allowed us to analyse the recognition conditions (active and passive) with a within-subjects design and have all participants engaged in an equal number of sentences per counterbalance condition. This was controlled so that participants were only matched with participants of their age group with personalised and individual links. Moreover, the format of the sentences was also pseudorandomised across subjects to assure that each counterbalance condition had the same number of sentences presented to each participant. This counterbalance controlled the retrieval task (active vs. passive), the format of the sentences presented at encoding (critical/original vs. filler/manipulated), and the format of the sentence presented at the intermediate phase considering the format at encoding (match vs. mismatch).

Regarding the use of distractor tasks, the math operations following each sentence presentation at the encoding phase aimed at forcing participants to break mental study of the sentences and pay attention to each sentence individually as it was presented, by shifting their cognitive resources and requesting the use of different cognitive abilities. Similarly, the distractor task asking to spot differences between two images followed the same idea to avoid restudy between the three study phases (i.e., encoding, intermediate and final memory test). These tasks also served to capture attention to the study, a measure that could exclude valid participants should self-report low attention levels at the end of the experiment. Additionally, the use of previously tested translated material was also a great advantage of our study and made it critical for us to control participants' spoken language, even in regard to distinguishing Portuguese from Portugal and Portuguese from other countries as a limiting factor for valid participation.

Critically, we included in our study the occasional misinformation, provided when a participant's correspondent yoked-pair had given the inaccurate response at sentence recognition. Importantly, regardless of the sentence presented in the passive condition, corrective feedback followed after each sentence of this task had underlined the portion of the sentence tested, which directed attention to the cause of the error and did not simply negate incorrect recognition responses (Seifert, 2002), and thus limited memory conformity. In this vein, following Mullet and Marsh's (2016) notion that errors must be detected and understood so they can be replaced in memory, additional analysis may be able to investigate how this misinformation might have facilitated error detection and posterior error correction, given the within-subject recognition task and the immediate corrective feedback provided.

A limitation of the study was the limited demographic information collected from participants at the beginning of the experiment. Given that adolescents participate in mandatory educational activities and a great percentage of young adults attend superior education institutions, the number of years of schooling should be inquired for expertise control. Particularly, since the majority of young adults that participate in psychological studies are contacted by their host universities to participate in their ongoing studies, leading to bias in the population and poorer generalization control. Moreover, students of humanities courses (e.g., translation, language courses) may be more practised in the nuances of semantic interpretation than students graduating in engineering and thus better distinguish critical sentences from sentences containing the pragmatic implication. The same should be controlled for adolescents due to the rising enrolment in professional courses. These effects should be controlled not only for the study of error generation but also for error correction.

In addition to the suggestions for future studies made throughout this section, it could be interesting to replicate the study with the sentences' presentation at the study phase as auditory stimuli. This alteration could lead to more errors generated and would more closely resemble pragmatic inference errors committed in our everyday communications with other people. Critically, it would likely present differences to the results observed with visual stimuli and along development.

On a final note, to the best of our knowledge, our study was the first to explore false memory error correction stemming from pragmatic inferences in adolescents. Framed by the Errorful Learning perspective and congruent with the MUAR framework

(Finn, 2017), we hope to have contributed to the existing gap in the literature concerning adolescents and encouraged future studies to pursue investigation on the underlying cognitive functions of error correction across healthy development.

Conclusion

The results of the present study lie in the scope of the errorful learning perspective. Indeed, our analysis suggested that active retrieval of encoded information facilitates error correction when the previous unsuccessful retrieval was followed by corrective feedback. This was true for both adolescents and young adults and was not modulated by the age groups at study. Furthermore, the results were congruent with the Memory Updated After Retrieval framework (Finn, 2017) by verifying that memory is vulnerable to new information and can update false memory errors when the information provided is in the form of corrective feedback. The lack of observable behavioural differences between adolescents and young adults indicates that the underlying cognitive functions supporting error correction need to be further investigated. As such, future studies should focus on determining the mnemonic and executive functions responsible for error correction across development.

References

Ausubel, D. P. (1968). *Educational Psychology: A Cognitive View*. Rinehart & Winston

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.

Bayen, U. J., Nakamura, G. V., Dupuis, S. E., & Yang, C. L. (2000). The use of schematic knowledge about sources in source monitoring. *Memory & Cognition, 28*, 480-500. https://doi.org/10.3758/BF03198562

Bjorklund, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review, 10*(1), 48–71. https://doi.org/10.1016/0273-2297(90)90004-N

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive psychology, 3*(2), 193-209. https://doi.org/10.1016/0010-0285(72)90003-5

Brewer, W.F. (1977). Memory for the pragmatic implications of sentences. Memory & Cognition, 5(6), 673-678. https://doi.org/10.3758/BF03197414

Brod, G., Hasselhorn, M., & Bunge, S. A. (2018). When generating a prediction boosts learning: The element of surprise. *Learning and Instruction, 55*, 22–31. https://doi.org/10.1016/j.learninstruc.2018.01.013

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5*(1), 73-99. https://doi.org/10.1016/0010-0277(77)90018-X

Bunge, S. A., & Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Current Opinion in Neurobiology, 17*, 243–250. https://doi.org/10.1016/j.conb.2007.02.005

Butler, A. C., Fazio, L. K., & Marsh, E. J. (2011). The hypercorrection effect persists over a week, but high-confidence errors return. *Psychonomic Bulletin & Review, 18*, 1238-1244. https://doi.org/10.3758/s13423-011-0173-y

Butterfeld, B., & Metcalfe, J. (2001). Errors Committed with High Confidence Are Hypercorrected. *Journal of Experimental Psychology: Learning Memory and Cognition, 27*(6), 1491–1494. https://doi.org/10.1037/0278-7393.27.6.1491

Butterfeld, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning, 1*(1), 69–84. https://doi.org/10.1007/s11409-006-6894-z

Carneiro, P., Lapa, A., & Finn, B. (2018). The effect of unsuccessful retrieval on children's subsequent learning. *Journal of Experimental Child Psychology, 166*, 400-420. https://doi.org/10.1016/j.jecp.2017.09.010

Carneiro, P., Lapa, A., Reis, J., & Ramos, T. (2020). Testing pragmatic inferences: The impact of language and culture. *Psicológica, 41*(1), 1-20. https://doi.org/10.2478/psicolj-2020-0001

Carneiro, P., Lapa, A., & Finn, B. (2021). Memory updating after retrieval: when new information is false or correct. *Memory, 29*(9), 1156-1175. https://doi.org/10.1080/09658211.2021.1968438

Carpenter, S.K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(6), 1563-1569. https://doi.org/10.1037/a0017021

Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(6), 1547-1552. https://doi.org/10.1037/a0024140

Chan, J. C. K., & McDermott, K. B. (2006). Remembering pragmatic inferences. *Applied Cognitive Psychology, 20*(5), 633-639. https://doi.org/10.1002/acp.1215

Chan, J. C. K., McDermott, K. B., Watson, J.M.& Gallo, D.A. (2005). The importance of material-processing interactions in inducing false memories. *Memory & Cognition, 33*, 389–395. https://doi.org/10.3758/BF03193057

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*(1), 66–73. https://doi.org/10.1111/j.1467-9280.2008.02245.x

Chan, J. C. K., Wilford, M. M., & Hughes, K. L. (2012). Retrieval can increase or decrease suggestibility depending on how memory is tested: The importance of source complexity. *Journal of Memory and Language, 67*(1), 78–85. https://doi.org/10.1016/j.jml.2012.02.006

Clark, C. M. (2016). *When and Why does Learning Profit from the Introduction of Errors?* [Doctoral Dissertation, University of California Los Angeles]. eScholarship Publishing. https://escholarship.org/uc/item/6zv5867p

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407–428. https://doi.org/10.1037//0033-295X.82.6.407

Danziger, K. (2008). *Marking the mind: A history of memory.* Cambridge University Press.

Decker, A., Finn, A., & Duncan, K. (2020). Errors lead to transient impairments in memory formation. *Cognition, 204*, 104338. https://doi.org/10.1016/j.cognition.2020.104338

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*(1), 17-22. https://doi.org/10.1037/h0046671

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28*(1), 1–11. https://doi.org/10.3758/BF03203630

Fassbender, C., Murphy, K., Foxe, J. J., Wylie, G. R., Javitt, D. C., Robertson, I. H., & Garavan, H. (2004). A topography of executive functions and their interactions revealed by functional magnetic resonance imaging. *Brain Research. Cognitive brain research, 20*(2), 132–143. https://doi.org/10.1016/j.cogbrainres.2004.02.007

Fazio, L. K., & Marsh, E. J. (2010). Correcting false memories. *Psychological science, 21*(6), 801-803. https://doi.org/10.1177/0956797610371341

Ferguson, H. J., Brunsdon, V. E. A., & Bradford, E. E. F. (2021). The developmental trajectories of executive function from adolescence to old age. *Scientific Reports, 11*(1), 1382. https://doi.org/10.1038/s41598-020-80866-1

Finn, B. (2017). A framework of episodic updating: An account of memory updating after retrieval. *Psychology of Learning and Motivation, 67*, 173-211. https://doi.org/10.1016/bs.plm.2017.03.006

Fuster, J. M. (1997). *The prefrontal cortex* (3 ed.). Raven Press

Greene, E., Flynn, M. S., & Loftus, E. F. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning and Verbal Behavior, 21*(2), 207-219. https://doi.org/10.1016/S0022-5371(82)90571-0

Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition, 40*(4), 505-513. https://doi.org/10.3758/s13421-011-0174-0

Gunter, R. W., Ivanko, S. L., & Bodner, G. E. (2005). Can test list context manipulations improve recognition accuracy in the DRM paradigm? *Memory, 13*(8), 862-873. https://doi.org/10.1080/09658210444000458

Harré, R. (2002). *Cognitive science: A philosophical introduction*. Sage.

Hays, M. J., Kornell, N., & Bjork, R. A. (2013). When and why a failed test potentiates the effectiveness of subsequent study. *Journal of Experimental Psychology: Learning Memory and Cognition, 39*(1), 290–296. https://doi.org/10.1037/a0028468

Houtman, F., & Notebaert, W. (2013). Blinded by an error. *Cognition, 128*(2), 228–236. https://doi.org/10.1016/j.cognition.2013.04.003

Huelser, B. J., & Metcalfe, J. (2012). Making related errors facilitates learning, but learners do not know it. *Memory and Cognition, 40*(4), 514–527. https://doi.org/10.3758/s13421-011-0167-z

Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia, 44*, 2017–2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010

Huttenlocher, P. R. (1990). Morphometric study of human cerebral cortex development. *Neuropsychologia, 28*(6), 517–527. https://doi.org/10.1016/0028-3932(90)90031-I

Izawa, C. (1970). Optimal potentiating effects and forgetting-prevention effects of tests in paired-associate learning. *Journal of Experimental Psychology, 83*(2), 340–344. https://doi.org/10.1037/h0028541

Jernigan, T. L., Brown, S. A., & Dowling, G. J. (2018). The Adolescent Brain Cognitive Development Study. *Journal of Research on Adolescence, 28*(1), 154-156. https://dio.org/10.1111/jora.12374.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*(1), 3-28. https://doi.org/10.1037/0033-2909.114.1.3

Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences, 6*, 299–305. https://doi.org/10.1016/S1364-6613(02)01923-X

Loftus, E. F. (1997). Memory for a past that never was. *Current Directions in Psychological Science, 6*(3), 60-65. https://doi.org/10.1111/1467-8721.ep11512654

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*(4), 361-366. https://doi.org/10.1101/lm.94705

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human, Learning and Memory, 4*(1), 19-31. https://doi.org/10.1037//0278-7393.4.1.19

Luna, B. (2009). Developmental Changes in Cognitive Control through Adolescence. *Advances in Child Development and Behavior*, 233–278. https://doi.org/10.1016/s0065-2407(09)03706-9

Luna, B., Marek, S., Larsen, B., Tervo-Clemmens, B., & Chahal, R. (2015). An integrative model of the maturation of cognitive control. *Annual Review of Neuroscience, 38*, 151–170. https://doi.org/10.1146/annurev-neuro-071714-034054

Kang, S. H. K., Pashler, H., Cepeda, N. J., Rohrer, D., & Carpenter, S. K. (2011). Does incorrect guessing impair fact learning? *Journal of Educational Psychology, 103*(1), 48-59. https://doi.org/10.1037/a0021977

Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences, 21*(1), 45-83. https://doi.org/10.1080/10508406.2011.591717

Karpicke, J. D., & Roediger III, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*(5865), 966-968. https://doi.org/doi:10.1126/science.1152408

Kornell, N. (2014). Attempting to answer a meaningful question enhances subsequent learning even when feedback is delayed. *Journal of Experimental Psychology: Learning Memory and Cognition, 40*(1), 106–114. https://doi.org/10.1037/a0033699

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(4), 989-998. https://doi.org/10.1037/a0015729

Maraver, M. J., Lapa, A., Garcia-Marques, L., Carneiro, P., & Raposo, A. (2022). Can we learn from errors? Retrieval facilitates the correction of false memories for pragmatic inferences. *PLOS ONE, 17*(8), e0272427. https://doi.org/10.1371/journal.pone.0272427

Markant, D. B., Ruggeri, A., Gureckis, T. M., & Xu, F. (2016). Enhanced Memory as a Common Effect of Active Learning. *Mind, Brain, and Education, 10*(3), 142–152. https://doi.org/10.1111/mbe.12117

McConnell, M.D., & Hunt, R.R. (2007). Can false memories be corrected by feedback in the DRM paradigm? *Memory & Cognition, 35*(5), 999–1006. https://doi.org/10.3758/BF03193472

McDermott, K. B., & Chan, J. C. K. (2006). Effects of repetition on memory for pragmatic inferences. *Memory & Cognition, 34*(6), 1273–1284. https://doi.org/10.3758/bf03193271

McDermott, K.B., & Roediger, H.L., III. (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory and Language, 39*(3), 508–520. https://doi.org/10.1006/jmla.1998.2582

McDermott, K. B., & Watson, J. M. (2001). The Rise and Fall of False Recall: The Impact of Presentation Duration. *Journal of Memory and Language, 45(1), 160–176.* https://doi.org/10.1006/jmla.2000.2771

Melton, A. W., & Irwin, J. M. (1940). The Influence of Degree of Interpolated Learning on Retroactive Inhibition and the Overt Transfer of Specific Responses. *The American Journal of Psychology, 53*(2), 173–203. https://doi.org/10.2307/1417415

Mera, Y., Rodríguez, G., & Marin-Garcia, E. (2021). Unraveling the benefits of experiencing errors during learning: Definition, modulating factors, and explanatory theories. *Psychonomic Bulletin & Review, 29*, 753–765. https://doi.org/10.3758/s13423-021-02022-8

Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology, 68*, 465-489. https://doi.org/10.1146/annurev-psych-010416-044022

Metcalfe, J., & Finn, B. (2011). People's hypercorrection of high confidence errors: Did they know it all along? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(2), 437-448. https://doi.org/10.1037/a0021962

Metcalfe, J., & Huelser, B. J. (2020). Learning from errors is attributed to episodic recollection rather than semantic mediation. Neuropsychologia, 138, 1-8. https://doi.org/10.1016/j.neuropsychologia.2019.107296

Metcalfe, J., & Miele, D. B. (2014). Hypercorrection of high confidence errors: Prior testing both enhances delayed performance and blocks the return of the errors. *Journal of Applied Research in Memory and Cognition, 3*(3), 189–197. https://doi.org/10.1016/j.jarmac.2014.04.001

Mullet, H. G., & Marsh, E. J. (2016). Correcting false memories: Errors must be noticed and replaced. *Memory & Cognition, 44*(3), 403-412. https://doi.org/10.3758/s13421-015-0571-x

Notebaert, W., Houtman, F., Opstal, F. Van, Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: An orienting account. *Cognition, 111*(2), 275–279. https://doi.org/10.1016/j.cognition.2009.02.002

Ofen, N., Kao, Y. C., Sokol-Hessner, P., Kim, H., Whitfield-Gabrieli, S., & Gabrieli, J. D. (2007). Development of the declarative memory system in the human brain. *Nature Neuroscience, 10*, 1198–1205. https://doi.org/10.1038/nn1950

Paradis, C. M., Florer, F., Solomon, L. Z., & Thompson, T. (2004). Flashbulb Memories of Personal Events of 9/11 and the Day after for a Sample of New York City Residents. *Psychological Reports*, *95*(1), 304–310. https://doi.org/10.2466/pr0.95.1.304-310

Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 3-8. https://doi.org/10.1037/0278-7393.31.1.3

Pillemer, D. B. (2009). *Momentous events, vivid memories*. Harvard University Press.

Postman, L., & Underwood, B. J. (1973). Critical issues in interference theory. *Memory & Cognition, 1*(1), 19–40. https://doi.org/10.3758/BF03198064

Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General, 143*(2), 644–667. https://doi.org/10.1037/a0033194

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science, 330*(6002), 335-335. https://doi.org/10.1126/science.1191465

Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied, 15*(3), 243-257. https://doi.org/10.1037/a0016496

Robins, S. K. (2016). Misremembering. *Philosophical Psychology, 29(*3), 432–447. https://doi.org/10.1080/09515089.2015.1113245

Robins, S. K. (2019). Confabulation and constructive memory. *Synthese, 196,* 2135–2151. https://doi.org/10.1007/s11229-017-1315-1

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249-255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803-814. https://doi.org/10.1037/0278-7393.21.4.803

Sampaio, C., & Brewer, W. F. (2009). The role of unconscious memory errors in judgements of confidence for sentence recognition. *Memory & Cognition, 37*(2), 158-163. https://doi.org/10.3758/MC.37.2.158

Skinner, B. F. (1953). *Science and human behaviour*. Macmillan

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory, 4*(6), 592–604. https://doi.org/10.1037/ 0278-7393.4.6.592

Sommers, M. S., & Lewis, B. P. (1999). Who really lives next door: Creating false memories with phonological neighbors. *Journal of Memory and Language, 40*(1), 83-108. https://doi.org/10.1006/jmla.1998.2614

Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience and Biobehavioral Reviews*, 24, 417–463. https://doi.org/10.1016/S0149-7634(00)00014-2

Sulin, R. A., & Dooling, D. J. (1974). Intrusion of a thematic idea in retention of prose. *Journal of Experimental Psychology, 103*(2), 255–262. https://doi.org/10.1037/h0036846

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science, 14*(5), 455-461. https://doi.org/10.1111/1467-9280.02453

Terrace, H. S. (1963). Discrimination learning with and without "errors". *Journal of the experimental analysis of behavior, 6*(1), 1-27. https://doi.org/10.1901/jeab.1963.6-1

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 and Language, 63*(2), 149–157. https://doi.org/10.1016/j.jml.2010.04. 004

Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory & Cognition, 14*(4), 329-338. https://doi.org/10.3758/BF03202511

Van Damme I., & Smets K. (2014). The power of emotion versus the power of suggestion: Memory for emotional events in the misinformation paradigm. *Emotion, 14*(2), 310–320. https://doi.org/10.1037/a0034629

Vaughn, K. E., & Rawson, K. A. (2012). When is guessing incorrectly better than studying for enhancing memory? *Psychonomic Bulletin and Review, 19*(5), 899–905. https://doi.org/10.3758/ s13423-012-0276-0

Wade, K. A., Garry, M., & Pezdek, K. (2018). Deconstructing rich false memories of committing crime: Commentary on Shaw and Porter (2015). *Psychological Science, 29*(3), 471-476. https://doi.org/10.1177/0956797617703667

Wagoner, B. (2013). Bartlett's concept of schema in reconstruction. *Theory & Psychology, 23*(5), 553–575. https://doi.org/10.1177/0959354313500166

Wahlheim, C. N., & Jacoby, L. L. (2013). Remembering change: The critical role of recursive remindings in proactive effects of memory. *Memory and Cognition, 41*(1), 1–15. https://doi.org/10.3758/s13421-012-0246-9

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*(1), 81–93. https://doi.org/10.1037/xap0000001

World Medical Association (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA, 310*(20), 2191-4. http://doi.org/ 10.1001/jama.2013.281053.

Wozniak, J. R., & Lim, K. O. (2006). Advances in white matter imaging: A review of in vivo magnetic resonance methodologies and their applicability to the study of development and aging. *Neuroscience & Biobehavioral Reviews, 30*(6), 762–774. https://doi.org/10.1016/j.neubiorev.2006.06.003

Zanolie, K., & Crone, E. A. (2018). Development of cognitive control across childhood and adolescence. In J. T. Wixted (Ed.). *Stevens' handbook of experimental psychology and cognitive neuroscience* (Vol. 4, 4th ed., pp. 1–24). John Wiley.

Appendix A

Sentences Presented in the Experiment

The present appendix comprises the 32 sentences first presented in the Encoding Phase, and then again in the Intermediate Phase. The bolded words implicate the pragmatic inferences responsible for generating the errors, while the words in brackets complete the original sentences and are the correct answers. Finally, the underlined words identify the space left blank for the final cued-recall test.

1. O ágil gato **alcançou** (apanhou) o peixe com as patas.
2. A meio da noite o marido sonolento foi buscar o jornal e **acertou** no (matou o) mosquito.
3. Depois de deixar os filhos na escola, a mãe foi **buscar** (comprar) pão.
4. O hipnotizador **juntou** (estalou) os dedos e acordou o cliente.
5. A corrida começou quando o árbitro **pressionou o gatilho** (apitou).
6. O barulhento cão de guarda **rosnou** (ladrou) ao homem que passava na rua.
7. O distraído professor **não tinha as** (esqueceu-se das) chaves do carro.
8. O cristão fecha os olhos e **fica uns minutos em silêncio** (reza) antes de cada refeição.
9. O campeão de karaté **bateu** no (partiu o) bloco de cimento.
10. Ela **pegou no** (calçou o) seu par de sapatos preferido e saiu.
11. O esquilo esfomeado estava a **roer** (comer) a relva.
12. Ele **pegou no** (atendeu o) telefone, ao ver que era a mãe quem lhe estava a ligar.
13. A jiboia esfomeada **apanhou** (comeu) o rato.
14. Tudo o que o Ricardo queria era poder **descansar** (dormir) uma noite inteira.
15. O ladrão **tirou** (acendeu) o isqueiro para fazer detonar o explosivo.
16. O vândalo **fez pontaria com** (atirou) a pedra à janela.
17. No dia da entrevista, ela **escolheu** (vestiu) a sua camisa da sorte.
18. O bebé **esteve acordado** (chorou) toda a noite.
19. A criança **picou** (rebentou) o balão com um alfinete.
20. O guarda-noturno **tirou** (bebeu) café do seu termo.
21. O King Kong **esteve no alto do** (subiu o) Empire State Building.
22. O príncipe encantado, gentilmente, **baixou-se em direção ao** (beijou o) rosto da Branca de Neve.

23. Dennis, o Pimentinha, **sentou-se na cadeira** (no colo) do Pai Natal e pediu um elefante.
24. Aquela estação de rádio só **gostava de** (passa) música hard rock.
25. O ladrão entrou numa loja e **pegou num** (roubou um) chocolate.
26. Depois de perseguir o ladrão por três quarteirões, o polícia finalmente **alcançou** (apanhou) o ladrão.
27. O boneco de neve **desapareceu** (derreteu) quando a temperatura atingiu os 26ºC.
28. O simpático empregado recebeu **umas moedas** (gorjeta) do cliente.
29. Assim que chegou à praia, a Marta estendeu a toalha, pôs **o chapéu de sol** (protetor solar) e deitou-se na areia.
30. O Capitão do submarino **disse** (gritou/ordenou): - "Submergir submarino!"
31. Ela **tropeçou** (caiu) ao descer as escadas.
32. O rato foi **atraído** (apanhado) pela ratoeira.