The Role of Automatic Memory Retrieval in Deception:

Developing a Binding Model of Deceptive and Truthful Discourse

Dissertation

zur Erlangung des akademischen Grades

doctor philosophiae (Dr. phil.)

---

vorgelegt dem Rat der Fakultät für Sozial- und Verhaltenswissenschaften der

Friedrich-Schiller-Universität Jena

von M.Sc. Franziska Schreckenbach

geboren am 22.01.1990 in Werdau

Gutachter:

1. Prof. Dr. Klaus Rothermund, Friedrich-Schiller-Universität Jena
2. Dr. Carina G. Giesen, Friedrich-Schiller-Universität Jena

Vorsitz der Kommission:

Prof. Dr. Rouwen Cañal-Bruland, Friedrich-Schiller-Universität Jena

Tag der mündlichen Prüfung: 13.10.2021

Table of Contents

Danksagung

Das Schreiben einer Doktorarbeit kann eine einsame Aufgabe sein - noch dazu in einer Zeit, in der Homeoffice und Social Distancing ein Teil der Normalität geworden sind. Umso dankbarer bin ich all jenen, die mich unter diesen Umständen trotzdem nach Kräften unterstützt haben, dieses Projekt erfolgreich zum Abschluss zu bringen.

Mein ganz besonderer Dank gebührt natürlich Klaus Rothermund. Lieber Klaus, ich danke dir für deine unschätzbare fachliche Unterstützung; dafür, dass du jedes meiner Manuskripte kritisch auseinandergenommen und mit mir wieder neu zusammengesetzt hast. Das war manchmal schmerzhaft, aber am Ende immer lehrreich und hat mir unglaublich dabei geholfen, ein besseres Verständnis für die richtige "Story" zu entwickeln. Außerdem danke ich dir dafür, dass du über diese lange Zeit hinweg und trotz diverser Elternzeit-Pausen nicht die Geduld mit mir und meinem Thema verloren hast.

Mein Dank geht außerdem an Carina Giesen, die nicht nur bereitwillig die Zweitbegutachtung für diese Arbeit übernommen hat, sondern als „Fachfrau für alles" meine Fragen zu jeglichen Themen immer mit einem Lächeln auf den Lippen und einem passenden Paper im Gepäck beantwortet hat. Vielen Dank dafür, dass du dein großes Wissen bereitwillig mit mir geteilt hast! Bedanken möchte ich mich auch bei dir, Nico: du hast mich überhaupt erst zu diesem Thema gebracht und mich mit deiner Begeisterung für neue Forschungsfragen immer wieder angesteckt. Vielen Dank für die vielen Gespräche, in denen wir Hypothesen formuliert, Experimente geplant, Daten ausgewertet und neue Folgefragen entwickelt haben.

Außerdem möchte ich mich bei all meinen Kollegen bedanken, die mich über die Jahre hinweg begleitet haben. Florian, du hast mich gelehrt, R zu bändigen, indem du jede noch

so seltsame Fehlermeldung mit mir gelöst hast. Franzi, du hattest immer ein offenes Ohr für mich, egal ob es um die Arbeit oder um private Themen ging. Anne, vielen Dank für die vielen aufmunternden Worte, wenn ich einen Durchhänger hatte, und natürlich auch für die Süßigkeiten, wenn mein Vorrat mal wieder aufgebraucht war! Carola, mit deiner freundlichen Art bist du die Seele der Abteilung. Vielen Dank, dass du uns bei all den formalen Arbeiten unterstützt, die auch zum Geschäft gehören. Mein Dank gebührt natürlich auch den wissenschaftlichen Hilfskräften, die mich bei der Durchführung meiner aufwändigen Studien nach besten Kräften unterstützt haben. Vielen Dank, dass ihr nicht nur Leute an PCs gesetzt, sondern auch Interviews geführt habt; dass ihr euch private Dinge habt erzählen oder auch vorlügen lassen!

Ohne den Rückhalt von guten Freunden und Familie wäre diese Arbeit wahrscheinlich nie zustande gekommen. Deshalb danke ich euch, Trine, Claudia, Jan, Juli, Janine und Antonia, dass ihr mich auf ganz unterschiedliche Art und Weise - sei es mit lieben Worten, durch Babysitting, beim Tischtennis, Klettern, Spiele spielen oder einfach gemütlichen Beisammensitzen - immer wieder mit der Kraft versorgt habt, die es gebraucht hat, um diese Aufgabe zu meistern. Ich danke meinen Kindern, Lara, Felix und Josephine, dass ihr mir jeden Tag aufs Neue zeigt, was wirklich glücklich macht: mit Neugier und Abenteuerlust die Welt entdecken, in Pfützen springen und Taschen voller Sand nach Hause tragen. Luis, vielen Dank, dass du von oben auf uns aufpasst!

Jura, du hast mit mir schon so manchen steilen Berg erklommen. Dieser hier war für mich besonders schwer, aber ich wusste, wenn du mich sicherst, dann kann ich nicht abstürzen. Ich danke dir für deine tatkräftige Unterstützung, dafür dass du immer an mich glaubst, und dass du für uns beide in aufregenden Zeiten die Ruhe behältst. Ich genieße jeden Tag mit dir und freue mich schon unheimlich auf unseren nächsten Gipfel!

Summary

Until today, research on deception primarily focuses on the creation and execution of first-time lies, while cognitive processes occurring after this point are only rarely considered. The aim of the current thesis is to extend the knowledge in this field by establishing a connection to current theories on binding and retrieval in action control. More specifically, it is investigated whether automatic processes help to recognize situations that are similar to a formerly deceptive situation, and therefore support people in successfully repeating their lies whenever it is needed. The empirical works presented within the current thesis are based on a study providing first evidence that reencountering a question that one has answered deceptively before, automatically activates the knowledge of having lied to this specific question (Koranyi et al., 2015). To illustrate this new perspective on lying processes, a binding model of deceptive and truthful discourse is introduced that in its structure reflects the possible perspective of binding and retrieval accounts on the topic. The model contains three core components: the relevant situational features (e.g., the person spoken to or the question answered), abstract as well as concrete information about the given response, and internal control states, like a suppression tendency. Whether and under which conditions binding and retrieval processes can be found for these components, is the superordinate issue that the studies presented in Chapters 2 to 5 are concerned with. In Chapter 2, it is investigated whether the frequency of deception influences the binding and retrieval processes found by Koranyi et al. (2015). In daily life situations, telling the truth predominates over lying (e.g., DePaulo & Kashy, 1998), thereby rendering deception a non default response mode that might be processed distinctively. Results indicate that distinctiveness indeed exerts an influence on binding and retrieval processes, as they disappear when the frequency of deception exceeds that of truth-telling in an experimental

setting. However, increasing the distinctiveness of truth-telling by making it the rare exception does not produce binding and retrieval processes for truths, thereby leading to the conclusion that distinctiveness alone is not a sufficient explanation for the observed effects. Instead, additional properties that differentiate deception from truth-telling seem to contribute to the emergence of binding and retrieval.

Chapters 3 and 4 are concerned with establishing a new situational cue: the person one spoke to. In Chapter 3, it is investigated whether binding and retrieval between this cue and the knowledge of one's former veracity can be observed at all. The obtained results indicate that – in contrast to questions as situational cue – a formerly unknown person becomes bound to the knowledge of having disclosed a truth to this person. When reencountering this person, the knowledge of having told the truth is activated automatically, while no such effect can be observed for the knowledge of having lied before. Based on these findings, Chapter 4 further examines the associative structure of bindings between a question, a person, and the knowledge of having told the truth or having lied before. It is shown that bindings primarily occur independently between the question and knowledge of having lied before, as well as between the person and knowledge of having told the truth before. In line with recent research on the nature of bindings (Giesen & Rothermund, 2014b, 2016; Moeller et al., 2016), it is assumed that binary bindings are the default mode when it comes to deception.

Chapter 5 considers the question whether information about internal control states can also be bound to situational cues. For this purpose, lies of omission are used to test whether a suppression of the truth can be reactivated automatically when encountering a similar situation again. It is shown that this suppression tendency can indeed become activated automatically as a consequence of the presentation of a formerly asked question and its associated secret answer.

Finally, and in the light of the presented empirical findings, the core components of the binding model of deceptive and truthful discourse are reconsidered and refined, as are the relations among them. This integration not only helps to improve the model, but also reveals open research questions and allows for specific predictions with regard to future empirical studies. Finally, the results presented in this dissertation can be used to further develop current theories on binding and retrieval in action control.

Zusammenfassung

Bis heute konzentriert sich die Forschung zum Thema Lügen hauptsächlich auf die Entstehung und Ausführung von erstmaligen Lügen, während kognitive Prozesse, die nach diesem Punkt stattfinden, vernachlässigt werden. Das Ziel der vorliegenden Dissertation ist es, die bisherigen Erkenntnisse in diesem Feld zu erweitern, indem eine Verknüpfung hergestellt wird zu aktuellen Theorien über Bindungs- und Abrufprozesse von Handlungsepisoden. Anhand von wiederholten Lügen soll untersucht werden, ob automatische Prozesse dabei helfen, Situationen zu erkennen, die einer vorherigen Lügensituation in bestimmten Aspekten gleichen und so die erfolgreiche Wiederholung von Lügen vereinfachen. Als Basis für diese Forschung dient eine Studie, in welcher der Nachweis dafür erbracht wurde, dass die Konfrontation mit einer Frage, auf die man zuvor gelogen hat, zum automatischen Abruf des Wissens über die vorherige Lüge führt (Koranyi, et al., 2015). Zur Strukturierung der vorliegenden Forschung wird zunächst ein allgemeines Gedächtnismodell für Lügen vorgestellt, aus welchem sich die Fragestellungen der einzelnen empirischen Arbeiten ableiten lassen. Das Modell umfasst drei Hauptkomponenten: die relevanten Merkmale einer Situation (z.B. die Person, zu der man gesprochen hat, oder die Fragen, die gestellt wurden), konkrete sowie abstrakte Informationen über die gegebene Antwort und interne Kontrollzustände. Ob und unter welchen Bedingungen sich Bindungs- und Abrufprozesse zwischen diesen Komponenten finden lassen, ist übergeordnetes Thema der Studien in Kapitel 2 bis 5.

In Kapitel 2 wird untersucht, ob sich die in der Studie von Koranyi et al. (2015) gefundenen Bindungs- und Abrufprozesse zwischen einer Frage und dem Wissen, auf diese Frage zuvor gelogen zu haben, durch die Häufigkeit von Lügen beeinflussen lassen. Da Lügen im realen Leben seltener vorkommen als Wahrheiten (z.B. DePaulo & Kashy, 1998)

und somit vom Standardverhalten abweichen, kann davon ausgegangen werden, dass diese distinkt verarbeitet werden. Es zeigt sich, dass diese Distinktheit dazu beiträgt, dass Bindung und Abruf stattfinden, da die Prozesse verschwinden, wenn die Häufigkeit der Lügen im experimentellen Kontext diejenige von Wahrheiten übersteigt. Umgekehrt lässt sich jedoch nicht zeigen, dass Bindungs- und Abrufprozesse für Wahrheiten hervorgerufen werden können, wenn diese im experimentellen Kontext nur noch selten geäußert werden sollen. Deshalb müssen neben der Distinktheit weitere Eigenschaften eine Rolle bei der Etablierung von Bindungs- und Abrufprozessen spielen, wie es etwa die größere kognitive Anstrengung von Lügen im Vergleich zu Wahrheiten sein könnte.

In Kapitel 3 und 4 wird neben den Fragen ein weiterer situativer Hinweisreiz etabliert: die Person, mit der man gesprochen hat. In Kapitel 3 wird zunächst untersucht, ob es überhaupt Bindungs- und Abrufprozesse zwischen diesem Hinweisreiz und dem Wissen über den Wahrheitsgehalt der vorherigen Aussage gibt. Es zeigt sich, dass – anders als bei Fragen – eine zuvor unbekannte Person mit dem Wissen verknüpft wird, dass man dieser Person eine Wahrheit offenbart hat. Bei erneuter Präsentation dieser Person wird das Wissen darüber, die Wahrheit gesagt zu haben, dann automatisch abgerufen. Aufbauend auf diesen Ergebnissen ist das Ziel der Studie in Kapitel 4, die assoziative Struktur der Bindungen zu untersuchen, die zwischen der Frage, der Person und dem Wissen darüber, die Wahrheit gesagt oder gelogen zu haben, bestehen. Die Befunde zeigen, dass in den meisten Fällen unabhängige, binäre Bindungen zwischen der Frage und dem Wissen über die Lüge sowie zwischen der Person und dem Wissen über die Wahrheit entstehen. Im Einklang mit weiteren Studien zu diesem Thema (Giesen & Rothermund, 2014b, 2016; Moeller et al., 2016) nehmen wir an, dass binäre Bindungen den Normalfall darstellen, während komplexere Bindungen nur in bestimmten Ausnahmefällen entstehen.

Kapitel 5 beleuchtet schließlich die Frage, ob neben dem abstrakten Wissen darüber, gelogen zu haben, auch interne Kontrollzustände Teil der gebildeten Handlungsepisode werden können. Zu diesem Zweck wird eine Sonderform der Lüge, die *Auslassung* genutzt und getestet, ob die Unterdrückung der Wahrheit in einer ähnlichen Situation automatisch wieder aktiviert werden kann. In den Ergebnissen zeigt sich, dass diese Tendenz zur erneuten Unterdrückung der geheimen Information tatsächlich automatisch aktiviert wird, wenn die dazu passende, zuvor gestellte Frage präsentiert wird.

Im abschließenden Teil der Dissertation werden die Ergebnisse der empirischen Studien genutzt, die Hauptkomponenten des zu Beginn vorgestellten allgemeinen Gedächtnismodells für Lügen zu spezifizieren sowie die Struktur der Beziehungen zwischen den einzelnen Komponenten genauer zu definieren. Gleichzeitig lassen sich unter Einbezug des Modells offene Forschungsfragen identifizieren sowie auf Basis der bereits gewonnenen Erkenntnisse konkrete Vorhersagen für zukünftige Studien formulieren. Zusätzlich liefern die Befunde neue Anknüpfungspunkte für die Weiterentwicklung von Theorien zu Bindungs- und Abrufprozessen von Handlungsepisoden dar.

1 Introduction

There are few reasons for telling the truth,
but for lying the number is infinite.
(Carlos Ruiz Zafón)

In every area of our life, we aim at building or maintaining satisfying interpersonal relationships. Whether we want to create a good impression at the job, spend a pleasant evening with our family, or settle a dispute with a friend, we strive for affiliation. Often, we can reach this aim by just being ourselves, but sometimes it appears necessary to actively manage the image others have of us. In such a situation, we need to decide between different strategies: among other possibilities, we can be open-hearted and disclose some secret truth, or we may prefer telling a lie to ensure that others maintain a positive image of us. Although deception is morally despised in public, research shows that people lie surprisingly frequently in daily life (DePaulo et al., 1996). One reason for this apparent discrepancy might be the fact that although deception is a widely used strategy, we lack the ability to reliably detect liars (Hartwig & Bond, 2011). To overcome this weakness, researchers have been long trying to discover reliable methods to detect liars, especially in the field of criminal investigation, where the detection of deception is of vital importance. Furthermore, during the last decades, researchers started to show an increasing interest in everyday lies, by studying their features (e.g., the frequency of lies, reasons for lying, or types of lies (DePaulo et al., 1996; DePaulo & Kashy, 1998), possible cues to deception (e.g., DePaulo et al., 2003), and the cognitive processes underlying lies (e.g. Walczyk et al., 2014).

With regard to cognitive processes, there exist two conflicting poles affecting the fundamental assumption of whether lies are perceived to be cognitively demanding or not. On the one hand, many researchers stand in the tradition of early theories about deception

where a deceptive statement is assumed to be something extraordinary, which in turn has to be more strenuous than telling the (ordinary) truth. On the other hand, the evidence for deception being a normal part of daily conversation increases, thereby leading some researchers to propose a more differentiated view on lies being not necessarily cognitively demanding. The activation-decision-construction-action theory (ADCAT; Walczyk et al., 2014) represents the attempt to describe cognitive elements underlying deceptive answers to questions, while rooting in the premise that truths are easier than lies. The authors propose a linear process of four steps: First, an automatic *activation* of truthful information takes place in a social context that signals a person to tell the truth. Secondly, a *decision* is made whether to lie or to tell the activated truth. The authors assume that people use quasi-rational decision making when lying, including likelihood estimations of salient consequences of their choices if they have sufficient time to do so. Afterwards, the third step, *construction*, takes place imposing cognitive load on a person depending on the strategy of lying. Omissions will cause only little cognitive load while constructing a false alibi might impose high load. Lastly, *action*, i.e. the delivery of the truthful or deceptive statement takes place, which is accompanied by monitoring in both cases. Thereby, Walczyk et al. introduce a cognitive theory of deception that applies to high-stakes situations in which lies are produced, like e.g., the interrogation of a criminal suspect, a job interview with an unqualified candidate, or an unfaithful partner in a relationship being confronted. Although the core of the theory is cognitive load, the authors also acknowledge that there might be situations where lying is cognitively less demanding than truth-telling (e.g., well-rehearsed lies).

Based on the opposite assumption is the information manipulation theory 2 (IMT2; McCornack et al., 2014), which assumes discourse production to consist of many parallel-distributed processes in numerous modular subsystems. The conceptual understanding of deception follows Grice's four maxims of communication (Grice, 1989), in which the dimensions of shared information are defined: quantity, the amount of relevant information shared; quality, the veracity of shared information; manner, the way in which disclosed information expressed; and relation, the relevance of shared information. Every violation of one of these maxims creates a deceptive message. The most important assumption of the IMT2 is the similarity of deceptive and truthful discourse, as both use the same memory

systems to arise. Deception is not naturally more difficult than truth production, which is reflected, among other things, by the lack of empirical findings for reliable behavioral cues to deception. On the contrary, deception is believed to occur as a problem-solving strategy using the most easily accessible information. Hence, it only occurs when the cognitive system evaluates it to be the most efficient solution to a given problem. This reasoning leads to a reversed order of volition and memory compared to the ADCAT: memory chunks are activated first, leading to the emergence of the intention to deceive. Deception is used because constructing it is easier than constructing a (complex or problematic) truth in a way that no one gets insulted.

Taken together, both accounts emphasize the role of memory in the context of deceptive discourse, while taking different points of view on the question whether there is ease or difficulty in the creation and execution of a lie. However, until today, there exists no comprehensive account that goes beyond this step by asking what happens *after* a lie has been produced for the first time. There are some works containing assumptions about how repeated lies differ from first time lies (e.g., Greene et al., 1985; O'Hair et al., 1981; van Bockstaele et al., 2012; Verschuere et al., 2011; Walczyk et al., 2014). Some authors claim that once a lie was successfully delivered, the probability of using similar communication patterns again in similar situations increases (McCornack et al., 2014). This is explained by a decrease in cognitive load because repeated patterns of behavior are always easier to produce than new ones (e.g., Dell et al., 1999; Goel & Pirolli, 1992). Furthermore, several studies were conducted on the effect of repeating lies on the ability to deceive, with most of them showing a robust effect of training. However, this effect only emerged for the previously trained lies, while no transfer to new lies was observed (van Bockstaele et al., 2012).

While these studies descriptively report observations with regard to the repetition of lies, they do not offer an explanation which cognitive processes might underlie these effects. The current thesis aims to fill this gap by providing a new perspective on the cognition of deception that combines it with recent accounts of binding and retrieval in action control. It is motivated by the assumption that people need to carefully remember, for example, whom they deceived and what they lied about in order to not being detected as a liar in the future. Therefore, it aims to investigate whether specific features of a deceptive

situation become stored in memory together with the knowledge of having lied in this situation, and whether these features can retrieve the information about having lied before automatically when being confronted with them on subsequent occasions. The research presented in this thesis was inspired by Koranyi et al. (2015), who were the first to investigate whether there exists an automatic retrieval mechanism that helps us to remember the lies we told before in an efficient way. However, to gain a profound of their study, some knowledge about theories on binding and retrieval is useful, which is why the next section will serve to first introduce this line of research. Afterwards, I will describe the aforementioned study by Koranyi et al. in greater detail and discuss possible implications and limitations with regard to binding and retrieval accounts. Finally, a binding model of deceptive and truthful discourse will be introduced, from which the research questions investigated within the scope of this thesis are deduced.

1.1 Approaches to Episodic Binding and Retrieval

Our everyday life is full of routines. Actions like brushing one's teeth in the morning, laying the table for dinner, or folding the laundry hardly require cognitive resources or enhanced attention in order to be performed successfully. Furthermore, while we proceed with these activities, we still have the capacities to think about completely unrelated things like a current task at work, a previous dispute with a friend, or the latest news heard on the radio. But has this always been the case? Thinking back to the date when we folded our laundry for the very first time, we probably had to focus heavily on this task in order to get a (halfway) satisfying result. Obviously, something changed between then and today. Besides the acquisition of specific motor skills, some kind of automatization must have happened, thus enabling us to shift our attention away from these routine tasks. In the field of cognitive psychology, there exist different theories about the development of automaticity, one being Logan's Instance Theory of Automatization (Logan, 1988, 1992). Logan assumes automaticity to be memory retrieval, defining performance as being automatic "when it is based on single-step direct access retrieval of past solutions from memory" (Logan, 1988, p. 493). Non-automatic behavior, on the contrary, is supposed to be algorithm-based. When a person performs a task for the first time, she follows an

algorithm to find a suitable solution, which becomes encoded in the episodic memory. Every further encounter of the task leads to a retrieval of the stored episode including the solution, and individuals can decide whether they rely on this memory retrieval or on the algorithm-based solution. At some point, retrieval becomes so reliable and fast that the algorithmic route is abandoned. Thus, automatization of a response to specific stimuli is seen as a change of the underlying processes due to a former learning history. This concept entails the assumption that every single encounter with a problem becomes stored separately in memory as a so-called *instance*, with a lot of instances strengthening the memory access. In fact, memory storage is seen as an unavoidable consequence of attention (paid to a stimulus), as is memory retrieval.

In a similar vein, Hommel (1998, 2004) introduced the concept of *event files*. Building up on findings from the field of perception, he aims at extending the idea of *object files*, which are assumed to represent temporary memory structures integrating different features of a perceived object into one single file. Hommel offers a more comprehensive view by claiming that not only perceptual features but also action-related information might be integrated into one common structure. More precisely, he assumes that a single encounter of a stimulus and an associated response (a so-called *stimulus-response episode*) suffices to bind these features together into an event file, which in turn influences subsequent processes immediately. In order to investigate these bindings, Hommel conducted a series of experiments using a sequential prime-probe design, in which the influence of a prime trial *n-1* on a probe trial *n* was investigated. This paradigm as well as his ideas of stimulus-response binding became very popular and led to the development of a huge amount of research on the topic. During these investigations, it became clear that stimulus-response bindings possess a greater degree of complexity than initially assumed, "including multiple levels of stimulus as well as response representation, bindings between stimulus features and responses, and bindings between multiple stimuli" (Henson et al., 2014, p. 379). With respect to the stimuli, beside the form-specific representation of perceptual features, a generalization toward categorically or perceptually similar instances takes place, as has been shown by switching between different exemplars of the same object (Denkinger & Koutstaal, 2009) or between the classification of pictures and written words (Horner & Henson, 2011). Furthermore, stimulus-response bindings have been shown to entail more

abstract information like semantic knowledge (e.g., the classification of a stimulus as living vs. non-living, Waszak & Hommel, 2007) or diagnostic decisions (Nett et al., 2015, 2016). Finally, even such abstract representations like the goal to inhibit a response were shown to be bound to situational features (Giesen & Rothermund, 2014a; Verbruggen & Logan, 2008), which led to the assumption of bindings to not only occur between (concrete or abstract) stimulus and response features but also between these features and internal cognitive states, like e.g., a control process that helps a person to suppress a specific response (e.g., Dignath et al., 2019; Egner, 2014).

How Long Do These Bindings Last?

When comparing Logan's (1988) and Hommel's (1998, 2004) works, a remarkable difference between their ideas becomes obvious: Logan unambiguously proposes an account of acquiring automatic behavior in a long-term way. Consequently, he assumes instances (as episodic structures) to be stored and collected in (long-term) memory until they provide a solid base for automatic behavior based on single-step memory retrieval. The idea of an automatic storage and retrieval of the knowledge of having lied before nicely matches these reflections, as it assumes bindings to emerge and last for a long period of time. Hommel, in contrast, provides an account of short-term action control, with event files being merely transient networks of bindings between different features of a specific event. These bindings can influence subsequent actions immediately, but are not assumed to have a direct connection to long-term memory processes. Such short-term bindings would not suffice to provide a reliable memory trace for one's former lies, which mostly date back longer than some seconds. Instead, a longer-lasting learning mechanism, as proposed by Logan, would be necessary for a successful retrieval of former lies. Within the literature of binding and retrieval, the amount of publications concerned with this topic increased considerably. Hommel himself started to investigate whether binding and learning interact with each other by conducting four experiments in which a typical binding paradigm was combined with a test of contingency learning (Colzato et al., 2006). The authors predicted that an interaction between the effects of binding and learning would be a hint for both processes to depend on each other. As they did not observe this interaction, they concluded that

binding and learning function independently. However, given the fact that the authors built their conclusions on null effects while using rather small sample sizes one could assume that a problem of insufficient power is at the core of these findings. This assumption becomes even more plausible when looking at the descriptive patterns of results, which closely resemble the predicted interactions. Besides, the assumption of the absence of an interaction being supportive of an independence of binding and learning has been contradicted in a recent publication (Schmidt et al., 2020). Therefore, the results of this study can at best lead to preliminary conclusions.

Several publications followed up on this topic, some being supportive of the assumption of binding and learning being independent from each other (e.g., Herwig & Waszak, 2012; Hommel & Colzato, 2009). However, the evidence pointing into a different direction also increases (see e.g., Frings et al., 2020; Giesen et al., 2020; Logie et al., 2009; Moeller & Frings, 2017; Schmidt, 2013; Schmidt et al., 2016). Lately, a comprehensive model was introduced that aims at explaining different paradigms in the field of learning, timing, binding, and attentional control with only a small set of assumptions about memory storage and retrieval (the parallel episodic processing model 2.0, PEP 2.0; Schmidt et al., 2016; see also Schmidt, 2013). In general, the account resembles Logan's idea of single events being collected and stored together. The only additional assumption is that the latest encoded event has a bigger influence on retrieval than earlier encoded events. The authors also do not assume a race between retrieval and non-automatic processing, but a proportional activation of different episode nodes based on their frequency and recency. This activation in turn leads to a specific response bias. The big advantage of PEP 2.0 is its parsimony, as it refrains from the assumption of additional higher-order control processes (in contrast to Colzato et al., 2006; Herwig & Waszak, 2012; Hommel & Colzato, 2009). Lately, this model was applied to an account of habit acquisition (Giesen et al., 2020), where binding and retrieval processes are seen as the basis for learning. The authors propose a *law of recency* claiming that habitual behavior results from (1) storing stimulus-response bindings in memory and (2) retrieving the most recent of these bindings when re-encountering this stimulus on a later occasion. They contrast this approach with a *law of exercise*, where contingencies in former behavior determine whether a habit is formed. With a modified contingency learning paradigm they demonstrate that effects of

contingency learning disappear when controlling for response retrieval effects. However, they do not abandon the law of exercise, but assume that it only applies after massive repetitions, thereby producing its own effects independently from the retrieval of recent episodes. In a similar vein, the *unitary view* (Schmidt et al., 2020) assumes binding and contingency learning to function on different timescales while merging into each other. In a re-analysis of former studies on contingency learning, the authors find much support for single bindings influencing performance, but also observe an additional effect of contingency. They conclude that both processes are strongly interrelated, as a large portion of the contingency effect seems to stem from individual bindings. Based on these findings, the present thesis also relies on the assumption of an interplay of binding and learning processes instead of a total independence, as this is a necessary precondition for a successful retrieval of former lies in a specific situation.

1.2 The Implicit Cognition of Deception

Some years ago, a first study was conducted in order to test whether mechanisms of binding and retrieval are relevant in the context of deception too (Koranyi et al., 2015). The authors of this study proposed the existence of an automatic retrieval mechanism that helps us to remember the lies we told before in an efficient way. To test their hypothesis, they developed an experimental setup consisting of two parts: first, participants were orally interrogated in a face-to-face interview where they had to answer questions from two different topics. Their task was to answer all questions from one topic honestly, while lying to all questions from the second topic. Secondly, participants had to perform a simple word classification task by deciding whether a presented probe was the word *honest* or *dishonest*. Before that probe, a prime appeared, consisting of one of the questions that were asked during the interview or a formerly unknown question. Koranyi et al. observed a congruency effect, i.e., a faster classification fo the probe word *dishonest* compared to *honest* after the presentation of a question that participants had lied to before. For honest as well as for unknown questions, no such pattern emerged. The authors explained these findings in terms of stimulus-response binding and retrieval, with the meta-knowledge of having lied

to a specific question being stored in memory and automatically retrieved when re-encountering this same question.

The results of this study have some interesting implications for theories on stimulus response binding and retrieval. Apparently, they support other findings about abstract knowledge being bound to situational cues (see e.g., Nett et al., 2015, 2016)., thereby pointing to the great flexibility of binding mechanisms in general. Secondly, and this might be more remarkable, the authors observed binding between a question and the knowledge of having lied after only one pairing (during the interview) for a particularly long period of time. In fact, the time span between the interview and the beginning of the prime task amounted to approximately five minutes (with the end of the prime task being reached another 15 minutes later). These findings contradict former estimations about bindings to only last for a few seconds (see e.g., Herwig & Waszak, 2012), thereby being completely independent from learning mechanisms, and instead lend support to recent accounts assuming an interplay of learning and binding. It is particularly interesting to apply the law of recency (Giesen et al., 2020) to the observed findings: As Koranyi et al. mainly used questions that participants can be assumed to have answered honestly before (e.g., "What is your favorite course at the university?"), the latest encounter of these questions during the interview must have exerted a bigger influence on retrieval than former situations, as it sufficed to produce a congruency effect in the subsequent prime task. However, this study provided only first evidence for an automatic memory retrieval in the context of lies, and therefore results should not be overrated. But following up on this starting point, the aim of the present dissertation is to verify the findings from Koranyi et al. and to deepen the understanding of binding and retrieval mechanisms during truthful and deceptive discourse.

1.3 Developing a Binding Model of Deceptive and Truthful Discourse

To provide an overview of the research presented in this thesis, I first want to introduce a binding model of deceptive and truthful discourse, from which the concrete research questions investigated in the following chapters can be deduced (see Fig. 1). Based on

previous findings about binding and retrieval processes, this model contains four core components:

(1) a stimulus component, containing relevant situational features that can cause retrieval, like the question one answered deceptively before,

(2) a response component, containing information about the specific response chosen, e.g., knowledge of having lied or told the truth in a given situation[1], or the concrete response content,

(3) internal control processes that help to regulate one's behavior by, e.g., suppressing truthful information, and

(4) the particular context, in which the deception or disclosure of a truth occurs.

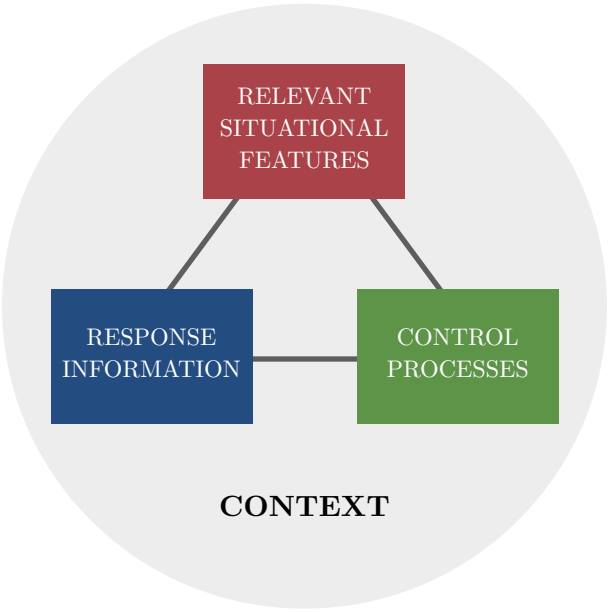

Figure 1: Illustration of the core components of the proposed binding model of deceptive and truthful discourse.

[1] Since a response in the sense of lying or telling the truth again was not involved in the experimental task designed by Koranyi et al. (2015), the knowledge of having lied or told the truth before could alternatively be seen as an abstract situational feature that is bound to and retrieved by other situational features. However, the information that this meta-knowledge contains covaries to a large degree with other response information, as knowledge of having told the truth will supposedly occur together with truthful content, while knowledge of having lied is supposed to occur with deceptive response content. Therefore, and in accordance with former conceptualizations of the response component (see Henson et al., 2014), it is assigned to response information in this model, as it constitutes the classification of the formerly given response.

Importantly, every core component stands for possibly numerous pieces of information about the respective component, i.e., there might be various situational features that are bound to different kinds of response information as well as control processes. Additionally, the given context is supposed to have a moderating influence on these bindings. In the present state, all components of the model are defined only coarsely, making it a skeleton that can help to organize research on the topic. The aim of the present thesis is to add some flesh to its single components, which is why we further defined and manipulated them, thereby investigating which impact they have on binding and retrieval processes in the context of deceptive communication. How this was done specifically, will be summarized in the following sections. Additionally, Figure 2 depicts the research questions investigated in each chapter when described from the perspective of this binding model.

1.3.1 The Context of a Lie

Contextual factors have been shown to play a role in binding and retrieval processes (e.g., Frings et al., 2020), an insight that mainly originates in an increasingly large body of research about the observation that not only features important to the solution of a given task but also task-irrelevant features (distractors) are bound to response information (e.g., Frings et al., 2007; Gibbons & Stahl, 2008; Giesen et al., 2012; Mayr & Buchner, 2006; Moeller et al., 2016; Rothermund et al., 2005). After the occurrence of a specific distractor in a prime trial, the repetition of this distractor in the probe leads to an improved reaction if the required response also repeats, while a change of the required response causes reduced performance. Distractor-response bindings are conceived as being adaptive in everyday life because relevant and irrelevant stimulus features often accompany each other systematically. Thus, seemingly irrelevant features gain predictive power for desired behavior in many situations (cf. Reber, 1989), leading to an enhanced performance when successfully retrieving this behavior again. However, a ubiquitous occurrence of these binding processes is not desirable, as it might lead to an overload of the system while creating useless files. Therefore, some attempts were undertaken to investigate the boundary conditions under which binding and retrieval processes occur. In this regard, it was shown that perceptual grouping as well as gestalt principles moderate the occurrence

of stimulus-response bindings in a way that only those stimuli (or stimulus features) are bound and retrieved that are perceived as (a) belonging together and (b) belonging to the foreground of a scene (Frings & Rothermund, 2011, 2017). Furthermore, stimuli with the same affective valence are more likely to be bound together than stimuli with different affective valence (Giesen & Rothermund, 2011). The implementation of specific task sets as well as feedback given for a former action are also assumed to affect the formation of stimulus-response bindings (Dreisbach & Haider, 2008; Giesen et al., 2017). The same applies to former experience, with established long-term stimulus-response bindings impeding new short-term bindings that contradict their older pendants (Moeller & Frings, 2014).

As these findings demonstrate the influence of contextual factors on binding and retrieval processes, it can be assumed that the context in which a lie occurs also has an impact on binding and retrieval of the knowledge of having lied. To investigate this assumption, we chose to manipulate one specific property of lies, namely their relative frequency compared to truthful statements. It has been shown that the relative frequency of lies in everyday life

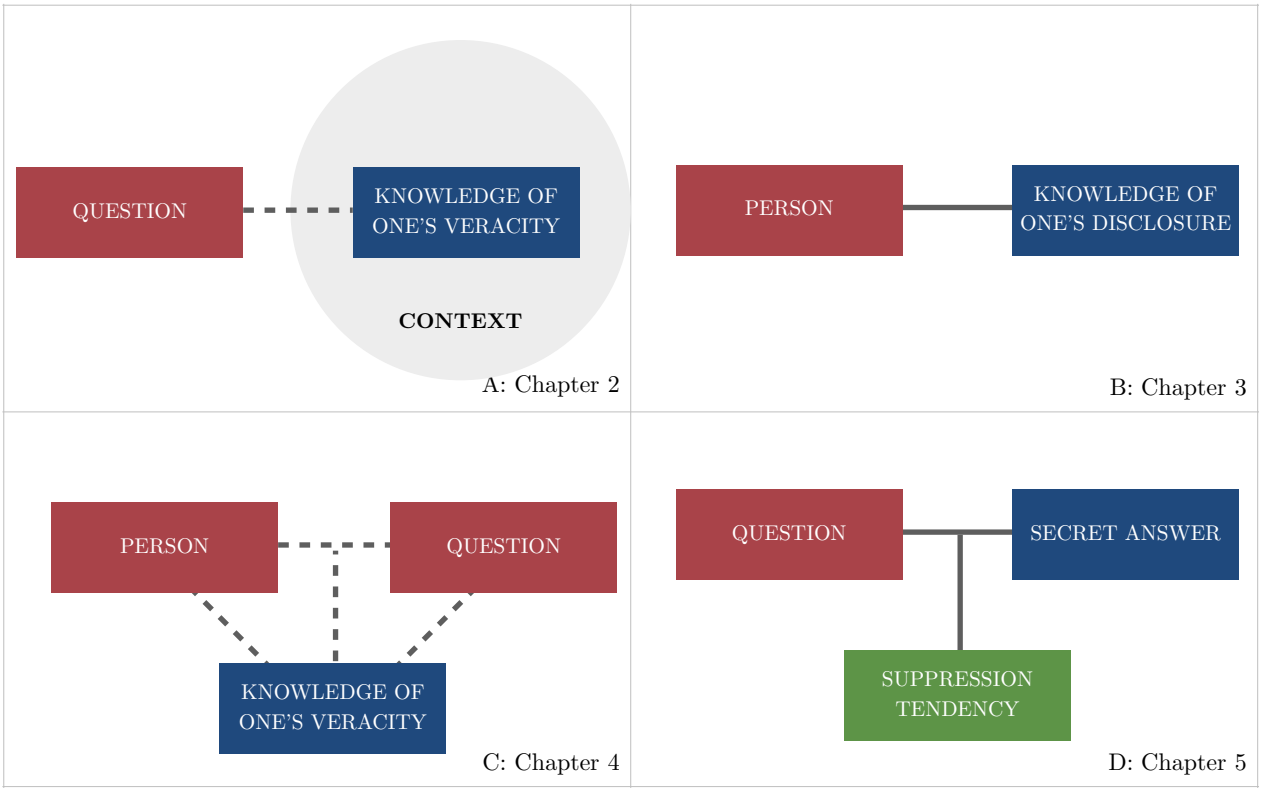

Figure 2: Illustration of the research questions investigated in Chapters 2 to 5 from the perspective of the binding model of deceptive and truthful discourse.

is lower than that of truths (DePaulo & Kashy, 1998), a fact that is supposed to lead to a *distinctive* processing of deception. Empirically, it was shown that the distinctiveness of a perceived episode determines whether it becomes stored in and later on retrieved from memory (so-called *von-Restorff effect*, Hunt, 2006; see also von Restorff, 1933). It is important to acknowledge that distinctiveness is not an inherent property of any object but refers to a specific type of processing an event that is achieved by putting this episode in contrast to other events. This understanding implies that distinctiveness is relative and can be assigned to different objects or events by changing the context in which they appear (e.g., a specific number in an array of letters will be perceived and thus processed distinctively, but it won't be in an array of more numbers). We used this mechanism to manipulate the distinctiveness of deceptive and truthful statements and to compare whether a change herein leads to a change in the observed retrieval effects. Spoken in terms of the binding model of deceptive and truthful discourse, we investigated the moderating influence of context on bindings between the situational feature "question" and the response information of one's former veracity (see Fig. 2a). To accomplish this aim, we used the paradigm developed by Koranyi et al. (2015), but manipulated the proportion of lies and truths being told during the interview in the first part of the experiment. While half of our participants were prompted to lie to 25 % of the questions being asked (*rare lie condition*), the other half had to lie to 75 % of the questions (*rare truth condition*). Given the fact that in our daily life, we tell the truth more often than we lie to someone (DePaulo & Kashy, 1998), we expected to replicate the congruency effect reported by Koranyi et al. in the rare lie condition. However, in the rare truth condition several results were assumed to be possible, depending on which impact the mere change of the relative frequency of lies had: If lies were stored and retrieved from memory because they differ from truths in a specific, qualitative way (e.g., by being cognitively more demanding, as stated in the ADCAT; Walczyk et al., 2014), the congruency effect would be expected to also emerge in this condition. However, if binding and retrieval depended on the (usually low) frequency of lies, a change in their relative frequency should lead to a change in the emerging results, comprising a congruency effect for honestly answered questions in the rare truth condition but no differences for deceptively answered questions. Finally, an influence of both properties, namely quality and quantity, might lead to a mixed result,

with the congruency effect disappearing for frequent lies, but on the other hand not emerging for rare truths. The results of this study are presented in Chapter 2.

1.3.2 Situational Features Causing Retrieval

Binding and retrieval are seen as flexible processes that happen simultaneously for multiple stimuli in a given situation. Applied to the field of deception, one would not only expect the question that one has answered dishonestly before to be stored in and retrieved from memory, but also other relevant contextual factors, like e.g., the person one talked to. However, it might be the case that different kinds of situational features vary in their ability to be bound to other features. In pavlovian conditioning, it is known that the combination of a gustatory stimulus with nausea-inducing toxin is better learned than the combination of an audiovisual stimulus with toxin. In contrast, audiovisual stimuli are easier associated with electric shocks than are gustatory stimuli (Garcia & Koelling, 1966). Obviously, the effectiveness of learning depends on the kind of stimuli that are associated with each other, hence, similar mechanisms might occur in the field of binding and retrieval. Support for this assumption can be found in studies showing that the strength of distractor-response bindings depends on the similarity of illumination between distractor and target (Laub & Frings, 2021), or that stimuli with the same affective valence are more likely to be bound together than stimuli with different affective valence (Giesen & Rothermund, 2011).

Based on these findings, it seems conceivable that questions function differently as a retrieval cue to the knowledge of having lied before than persons, as both cues contain several differences in processing. First, there is a difference in modality, as questions in a conversation are perceived acoustically while persons are perceived visually. Second, perceiving a person as a conversational partner not only elicits perceptual processing but also activates specific relational scripts: whenever we talk to a person, we relate to each other in some way, and this relationship influences our behavior as well as the behavior of our conversational partner. Empirical findings have shown that good friends and family are deceived rather rarely, acquaintances somewhat more often, and unknown people in more than half of the interactions, thereby causing a complete reversal in the distribution of lies and truths (DePaulo & Kashy, 1998). Taking into account that (1) persons should be

stored and retrieved as a cue to former statements too, and (2) usually, we do not tell strangers the whole truth, we conducted several experiments to see whether these factors influence memory processes. Therefore, in Chapter 3, two experiments are presented in which we investigated whether retrieval for persons can be detected at all, and whether this retrieval is more pronounced for honest statements, as we used formerly unknown interrogators. With regard to the binding model of deceptive and truthful discourse, this equals the investigation of a binding between the situational feature "person" and the knowledge of having told the truth before (see Fig. 2b). For this purpose, we adapted the paradigm of Koranyi et al. (2015) in such a way that participants then met two different interrogators during the interview and had to behave honest towards one of them while lying to the other one. In the subsequent classification task, we used facial pictures of the interrogators as primes instead of questions.

Considering that real-life situations always contain a high number of stimuli occurring simultaneously, Chapter 4 serves to integrate the findings about questions and persons and tests a combined account of memory retrieval for honest and dishonest statements. In the domain of distractor-response binding it was already shown that different situational features can be integrated with response codes simultaneously (Colzato et al., 2006; Giesen & Rothermund, 2014b; 2016; Moeller et al., 2016). In most cases, each feature alone was able to retrieve a given response, pointing to an individual (i.e., *elemental*) integration of this response and the different features. However, in some exceptional situations the features were perceived as a holistic body, thereby leading to a joint (i.e., *configural*) integration of all features and the belonging response code. In these cases, only the common repetition of all elements led to a retrieval of the given response (Moeller et al., 2016). While in this study the difficulty of separating the single features from each other was seen as a modulating factor, the authors assume that a flexible adjustment of processing to a current task context is possible. Applying these findings to the field of deception, we investigated whether persons and questions are simultaneously bound to the truth status of a former statement and whether these features lead to retrieval in a configural way (i.e., only the combination of both features retrieves knowledge about the former truth status) or an elemental way (i.e., both features retrieve the truth status on their own). Spoken in terms of the binding model, we investigated the associative nature of

bindings between different situational features and the knowledge of one's former veracity (see Fig. 2c). The results of this study are presented in Chapter 4.

1.3.3 Binding of Internal Control Processes

In Chapters 2 to 4, the observed bindings always contain a situational feature (i.e., a specific person or a question) and the knowledge of having lied or having told the truth before as a response component. Besides, we only investigated one specific kind of deception, that is, outright fabrications, which are assumed to occur comparatively rarely not only from a theoretical perspective (see McCornack et al., 2014), but also based on empirical findings when comparing them with lies of omission (e.g., DeScioli et al., 2011; Levine et al., 2002; Pittarello et al., 2016; Rogers et al., 2017; Schweitzer & Croson, 1999). According to the IMT2, omissions are seen as violations of the maxim of quantity, which often occur without the addition of a quality violation (i.e., the fabrication of a false story) due to their cognitive ease. However, according to the preoccupation model of secrecy (Lane & Wegner, 1995), a model specifically concerned with the cognitive consequences of keeping a secret, there seems to be something remarkable about omissions, as they tend to be easily accessible from memory, with possibly unwanted effects like having intrusive thoughts. Obviously, investigating omissions is important when aiming to gain a comprehensive understanding of memory retrieval processes in the context of deception. Furthermore, investigating omissions enables us to separately look at internal control processes that are assumed to also be associated with situational features and response information in the field of binding and retrieval (Dignath et al., 2019; Egner, 2014; Giesen & Rothermund, 2014a). To test whether such a control process can be bound to situational features in the context of deception too, we investigated omissions by developing a new paradigm that enabled us to measure the automatic retrieval of a suppression response experimentally. In this study, participants were first requested to withhold secret information about a fictitious love-affair. Afterwards, a specific key press response was established as a behavioral indicator of suppression. In the last part of the study, participants performed a Go/No-Go task where the color of a probe word signaled them whether to execute a response or not. Beforehand, a question from the former chat appeared, with the following probe word representing either secret or innocuous

information from the love-affair scenario. With this task, we measured automatic suppression tendencies elicited by the combination of questions and answers. In the case of a question being presented together with matching secret information, a facilitation of the Go response was expected. This experimental procedure enabled us to measure bindings between a situational feature, the secret answer, and an internal control process (in this case, the suppression tendency; see Fig. 2d). The results of this study are presented in Chapter 5.

1.4 Aims and Scope of the Present Thesis

To recapitulate, the present thesis aims at developing a more profound understanding of the role that binding and retrieval processes play in the context of deceptive communication. To reach this aim, the components of a binding model of deceptive and truthful discourse were investigated that are assumed to be part of these processes. Chapter 2 therefore addresses the question whether the frequency of lies as an exemplary context works as a moderator for binding and retrieval. Chapters 3 and 4 focus on the different situational features that can be bound to the knowledge about the truthfulness of one's former statements, and finally, Chapter 5 examines whether previous omissions lead to automatic suppression tendencies when encountering a similar situation again. Finally, the results of these studies will be used to (1) describe their implications for current theories on binding and retrieval, and (2) to evaluate the current status of the binding model of deceptive and truthful discourse by discussing what we know so far and which issues evolve for future investigation.

2 Quantity matters: The Frequency of Deception Influences Automatic Memory Retrieval Effects

No man has a good enough memory
to be a successful liar.
(Abraham Lincoln)

The following article has been published as:

Schreckenbach, F., Rothermund, K., & Koranyi, N. (2020). Quantity matters: The frequency of deception influences automatic memory retrieval effects. *Quarterly Journal of Experimental Psychology, 73*(11), 1774-1783. https://doi.org/10.1177/1747021820924652

The supplemental online material is presented in Appendix A.

Original Article

QJEP

Quarterly Journal of Experimental Psychology
2020, Vol. 73(11) 1774–1783

DOI: 10.1177/1747021820924652
qjep.sagepub.com
SAGE

Quantity matters: The frequency of deception influences automatic memory retrieval effects

Franziska Schreckenbach, Klaus Rothermund and Nicolas Koranyi

Abstract

We investigated automatic retrieval of the knowledge of having lied or having told the truth to a question, depending on (a) the quality of the statement (true vs. false response) and (b) the overall proportion of (dis-)honest responses. We therefore manipulated the proportion of lies and truths being told in an oral interview. Automatic retrieval of this meta-knowledge was assessed with a categorisation task, where the probe words *dishonest* and *honest* had to be classified, while questions from the interview served as task-irrelevant prime stimuli. Results revealed an automatic retrieval of knowledge about having lied to a question only for participants who had told few lies in the interview, but not for those who had told many lies. No retrieval effects were obtained regarding questions that had been answered truthfully. These findings suggest a combined influence of quality and quantity of dishonest statements on automatic memory retrieval, thereby being in accordance with recent accounts of action control.

Keywords

Lying; automatic processes; instance-based learning; distinctiveness

Received: 17 April 2019; revised: 8 April 2020; accepted: 13 April 2020

Recent research provided first insights into a basic mechanism that keeps track of the lies a person has told in response to certain questions. Specifically, as soon as a person encounters a situation that is similar to the one in which the lie was told before knowledge about having lied is automatically retrieved from memory (Koranyi et al., 2015). The approach has its roots in theories of instance-based automatisation of behaviour (Logan, 1988; see also Denkinger & Koutstaal, 2009; Hommel, 1998, 2004; Rothermund et al., 2005), which propose that the execution of an action in a specific situation is stored as an episodic unit in memory together with certain situational cues (alternatively called an *event file*, Hommel, 1998, 2004, *instance*, Logan, 1988, or *stimulus-response episode*, Rothermund et al., 2005). Encountering a similar situation again leads to an automatic activation of this episodic unit, enabling people to react fast and consistently across different situations. This episodic binding and retrieval perspective can explain a vast range of research findings across various paradigms in the field of cognition and action control and automatisation (e.g., Frings, Hommel, et al., 2020; Frings, Koch, et al., 2020; Schmidt et al., 2016). This model can also be applied to the field of deception, where it has been argued that the knowledge about having lied to a specific question is expected to be retrieved automatically from long-term memory when the same question is encountered again (Koranyi et al., 2015).

In the current study, we aim to extend the investigation of an automatic retrieval of knowledge about having lied by also taking into account contextual factors. Specifically, we are interested in whether the overall relative frequency of lies and truthful statements modulates the retrieval of a specific lie (or truthful statement). Such an effect of the contextual frequency of lies versus truthful statements has important theoretical as well as practical implications.

From a theoretical perspective, contexts play an important role for retrieval processes (e.g., Frings, Hommel, et al., 2020; Frings, Koch, et al., 2020): Contexts can either serve as direct retrieval cues themselves (Tulving &

Friedrich-Schiller University Jena, Jena, Germany

Corresponding author:
Franziska Schreckenbach, Institute of Psychology, Department of General Psychology II, Friedrich-Schiller University Jena, Am Steiger 3, D-07743 Jena, Germany.
Email: franziska.schreckenbach@uni-jena.de

Thomson, 1973), or they can modulate retrieval by increasing or decreasing the distinctness of an episode. Research has repeatedly shown that stimuli and episodes that deviate from their surroundings are perceived as being more distinct and tend to attract attention (Duncan & Humphreys, 1989), which in turn facilitates memory encoding and retrieval (so-called "von Restorff effect," Hunt, 2006; von Restorff, 1933). On the contrary, high similarity with the context reduces distinctness and has a negative effect on memory for a specific stimulus or episode. Demonstrating such an influence of the contextual frequency of lies versus truths on related binding and retrieval effects for questions and the truth values of the corresponding answers thus provides another test of episodic binding and retrieval models in the domain of deception and memory for one's own lies. Demonstrating such an influence extends the scope of application of episodic binding accounts to these contexts, which can be considered to reflect evidence for a bottom-up modulation of episodic encoding, binding, and/or retrieval (Frings, Hommel, et al., 2020; Frings, Koch, et al., 2020). Furthermore, it would also provide further evidence for an episodic binding and retrieval account of memory for lies (Koranyi et al., 2015; Schreckenbach et al., 2019).

Variations in the frequency of lies versus truths are also interesting from a more practical or applied perspective. Differences in the default frequencies of lies can reflect personal traits (e.g., relating to honesty/dishonesty), but they may also reflect specific situational settings (e.g., different standards, norms, expectations, or temptations with regard to telling the truth or a lie, or acting in socially desirable vs. authentic ways). If our assumptions are correct, then these differences in the frequency of lying may also influence memory for one's own (and other people's) lies. For instance, frequent liars should have more difficulty in remembering their own lies, and will consequently run into difficulties with regard to keeping up a consistent image of themselves. It may also be more difficult to track one's own (or another person's) lies if this lie occurred in a context that is characterised by a high frequency of lies (e.g., in contexts in which deviations from what is socially desirable are sanctioned).

To test automatic memory for lies, Koranyi et al. (2015) developed a paradigm in which participants took part in an oral interview and were prompted to tell the truth to half of the questions being asked but to lie to the other half. To assess automatic retrieval of knowledge about having lied, a simple categorisation task was used in which the probe words *honest* (in German: *ehrlich*) and *dishonest* (in German: *gelogen*) had to be identified by pressing a corresponding key. In this task, the questions of the interview were presented as task-irrelevant primes to test whether identification of a probe word was facilitated by a corresponding prime question. As expected, the presentation of a question that had been answered dishonestly led to faster identification of the probe word *dishonest* compared with *honest*, suggesting an automatic retrieval of the knowledge of having lied to a question when the question is encountered again. Importantly, no comparable congruency effect was obtained for those questions to which participants had responded honestly during the interview, indicating that an automatic retrieval of information about the truth-value of previous statements was restricted to lies.

The aim of the current article is to extend this approach to the question of whether the frequency of lies has an influence on automatic memory processes. According to current theorising on automatic binding and retrieval processes, these episodic memory processes are influenced by contextual factors reflecting involuntary bottom-up processes (e.g., salience, figure/ground asymmetries; Frings & Rothermund, 2011, 2017). With our present study, we want to investigate whether these boundary conditions also have an influence on memory for lies versus truthful statements. Before describing the specific design and hypotheses of our study, we first introduce two different properties of lies which might influence these processes, namely their quality and their quantity.

Quality of lies

According to models of lie production, the asymmetry in encoding and retrieving lies versus truths could stem from qualitative differences between the two kinds of statements. Various lines of research suggest that giving a lie is qualitatively different from telling the truth. First, it was suggested that lying is more cognitively demanding than telling the truth (e.g., Sporer & Schwandt, 2006, 2007; Vrij et al., 2010; Walczyk et al., 2003, 2014). There are several reasons for this assumption: Because liars have to think of a novel answer, they need to suppress the truth, activate the deceptive answer deliberately (while the truth is often activated automatically), they have to remind themselves to role-play, and they tend to monitor their own behaviour and the targets' reactions to a greater extent than do truth-tellers.

Walczyk et al. (2014; see also Walczyk et al., 2003) go one step further within their activation-decision-construction-action theory (ADCAT), where they specify the particular processes that have to occur before a lie can be successfully delivered. The authors assume that the *default mode of responding* is to answer honestly to a question, with the truth being activated in an automatic fashion, whereas deceiving requires additional controlled processes to take place. These deliberate processes contain the actual decision to lie, its construction, and its execution (the action component) while controlling one's behaviour and monitoring the target's reactions to the lie.

Models that distinguish true and false statements in terms of qualitative features provide a straightforward answer to explain the asymmetry in memory for truths and

lies: Investing effort, deliberation, resources, and/or cognitive control should lead to deeper levels of processing, better encoding, and more efficient retrieval of lies compared with truths.

Quantity of lies

Another account to explain this asymmetry focuses on external rather than internal features: Recent research has invested a lot of effort to gain insights into the frequency of lies. Based on a diary study, DePaulo et al. (1996) concluded that people tell approximately one lie a day. From a moral point of view, this looks like a lot but given the mean amount of social interactions a day (being defined by DePaulo et al. as "any exchange between you and another person that lasts 10 min or more . . . in which the behavior of one person is in response to the behavior of another person.") and the fact that during one of these interactions we typically tell more than one truth, lies are still substantially less frequent than telling the truth. Furthermore, recent findings suggest that the finding of people lying once per day *on average* might be based on a skewed distribution containing a few prolific liars telling several lies a day while most people indeed do not lie every day (see Serota et al., 2010).

Being less frequent renders lies more distinctive, and according to Hunt (2006), distinctiveness is a factor that facilitates memory retrieval (von Restorff effect; von Restorff, 1933). A low relative frequency of occurrence alone thus might already account for an enhanced memory for lies compared with truths, due to distinctive processing. Importantly, this explanation does not draw on specific qualities of lies (e.g., cognitive effort, controlled construction), but explains the memory asymmetry for lies and truths solely on the basis of their relative quantity.

This explanation implies that the asymmetry of retrieving lies versus truths is not an inherent feature of memory, but that it can be reversed by changing the relative proportion of lies and truths: If the low quantity of lies (i.e., their high distinctiveness) is what causes their being stored and retrieved, then it should be possible to reverse the asymmetry by making lies the default and turning true statements into a rare exception.

The current study

To test for the influence of quality and quantity on memory processes for truths and lies, we conducted an experiment that used the same paradigm as Koranyi et al. (2015). In addition to the manipulation of having to give true versus false answers to certain questions, we manipulated the overall quantity of lies and truths in our current study. Some participants had to tell only few (i.e., 25%) lies and many truths in the oral interviews, replicating the standard frequency asymmetry of everyday lies and truths. For another group of participants, however, this asymmetry was reversed (i.e., 75% lies, 25% true answers), which should render true statements more distinct.

Combining the above considerations, three possible predictions are possible: (a) only the quality of lies (i.e., low vs. high effort in producing honest vs. dishonest answers) is responsible for memory processes; (b) only the relative quantity (i.e., high vs. low distinctiveness) of a certain type of statement is what drives memory processes; and (c) a combination of both quality and quantity is necessary for memory encoding and retrieval (i.e., only distinct lies are encoded). Each prediction relates to a specific set of statistical hypotheses (see Figure 1, for a visual illustration of the predicted patterns):

(1) According to the quality account, a lie has an increased chance of being encoded into episodic memory due to the processes that are characteristic for lies (cognitive capacity, resources, depth of processing), independently of the context (see Figure 1a). In statistical terms, this prediction corresponds to a simple two-way interaction of the type of the prime question (answered truthfully vs. dishonestly) and the probe word (honest vs. dishonest). This interaction effect should be driven entirely by questions to which one has lied—no effect should be obtained for questions that were answered truthfully. Furthermore, the pattern of prime question × probe word congruency effects should not differ between the rare lies and the rare truths conditions, resulting in an absence of the three-way interaction.
(2) According to the quantity account, a statement is encoded into memory if it is distinct, that is, if it deviates from the context in which it is given: Accordingly, not only lies, but also truths can be stored and retrieved from memory if they are made in a context in which they are rare (see Figure 1b). Statistically, this pattern would also produce a two-way interaction between prime question and probe word. This interaction should be present for both the rare lies and the rare truths contexts, but these two interaction effects should be driven by different components of the prime factor: Dishonestly answered prime questions should produce faster responses for the probe word *dishonest*, compared with the probe word *honest*, whereas prime questions for which an honest answer had been given during the interview should not affect responding to the probe words. In the rare truths condition, however, a reversed pattern should emerge, with honestly answered prime questions having a facilitative effect on the probe word *honest* compared with the probe word *dishonest*, whereas no such effects should occur for prime questions to which

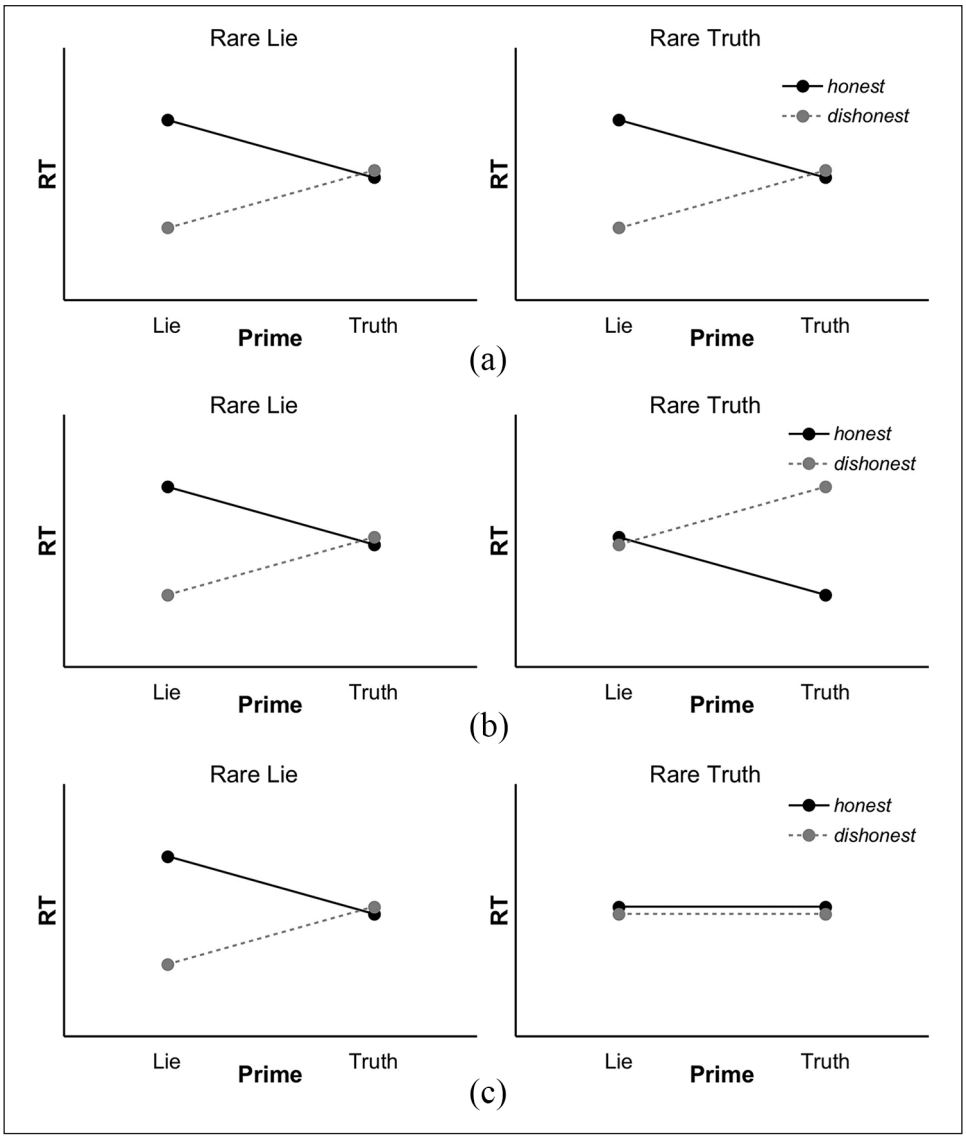

Figure 1. Predicted response time patterns for the qualitative ((a) cognitive effort during lie production causes episodic binding/retrieval), quantitative ((b) distinctness causes episodic binding/retrieval), and combined ((c) lack of resource investment or distinctness prevent episodic binding/retrieval) models as a function of context (rare lie vs. rare truth), prime question (lie vs. truth), and probe word (honest vs. dishonest).

one had lied. In statistical terms, this asymmetry corresponds to an additional interaction effect between probe word and context (in the rare lies context, responding should be faster for the probe word *dishonest*, whereas in the rare truths condition, responses should be faster for the probe word *honest*). Like the quality account, however, the quantity account does not predict a three-way interaction.

(3) According to the combined account, both properties (the quality of having told a lie, and the fact that lies are rare) are needed for episodic memory storage and retrieval. Neither truths (due to their lack of drawing on cognitive resources, resulting in shallow processing) nor lies that are given in a context of many other lies (due to their lack of distinctiveness) are stored and automatically retrieved from memory (see Figure 1c). Statistically, this prediction would result in a three-way interaction between prime question, probe word, and context: The combined account predicts an effect of the prime question × probe word interaction within the rare lies context (which should be driven by the questions to which one had lied). However, there should be no interaction between prime question and probe word within the rare truths condition.

Because the presence of a null effect (as it is predicted for the rare truths condition in Hypothesis 3) requires additional assumptions regarding the strength of a to-be-detected effect

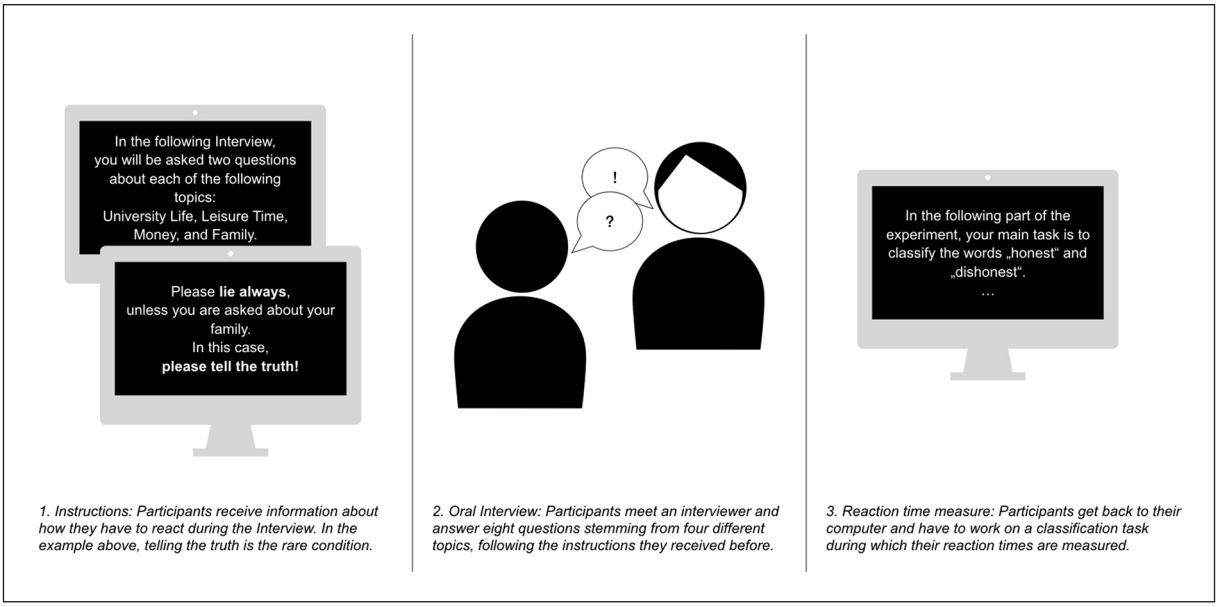

Figure 2. General procedure of the experiment.

when using frequentist statistics, we additionally calculated Bayes factors for all three hypotheses, which also provides information on the likelihood of the null hypothesis in relation to the alternative hypotheses.

Method

Participants and design

The study has been conducted in accordance with ethical standards and was approved by the Ethical Commission of the Faculty of Social and Behavioural Sciences of the University of Jena. Sample size was determined relying on the effect size of $\eta_p^2 = .08$, found by Koranyi et al. (2015) in a similar analysis for a three-way interaction.[1] Accordingly, a power analysis for a two-sided *F*-test was conducted based on this value, setting $\alpha = .05$ and $1 - \beta = .80$, which led to a proposed sample size of 96. To account for possible exclusions, we recruited 99 students from a German University. The data of eight subjects had to be excluded because of extremely high error rates (>30%; average error rates: 3.7% without excluded participants). In addition, two subjects were excluded because of high error rates in control trials (>15%, average error rates: 3.9% without excluded participants) indicating that they did not read the prime questions carefully. The final sample thus consisted of 89 students (34 female) with an average age of $M = 23.5$ years ($SD = 3.56$). All participants gave informed consent by a keypress at the beginning of the experiment. They took part in exchange for €3 and a bar of chocolate. We used a $2 \times 2 \times 2$ factorial design with the within-subjects factors prime question (lie vs. truth) and probe word (dishonest vs. honest) and the between-subjects factor frequency (rare lie vs. rare truth).

Procedure and materials

Upon arrival, participants were seated in front of a computer and received instructions on the screen. First, they were informed that they would participate in an interview, where they would be asked two questions about each of four different topics (university life, leisure time, money, and family; see Supplementary Material for a complete list of questions). Each question referred to an important issue but at the same time should not be too intimate to prevent participants from answering untruthfully even if they had to tell the truth (see Figure 2 for general procedure).

Counterbalanced across participants, frequency context was induced by first telling them that their general task was either to be honest or to lie to all questions during the interview, except for the questions stemming from one of the four topics. When these questions were asked, they should always act against their general rule, that is, tell a lie or the truth, respectively. For example, instructions in the rare lie condition would look like this: "Please answer all questions honestly, except for the questions from the topic university. When these questions are asked, please tell the interviewer a lie." Participants were also prompted not to reveal their dishonesty to the interviewer and to act so as to convince him of the truthfulness of all their statements. To achieve this goal, participants were instructed to always wait for some seconds before providing their answer.

Participants were then guided into a separate room where the interview took place. After a short introduction, the interviewer asked the questions in each topic successively, with the order of topics in the interview being randomised. The interviewers were blind with regard to the frequency condition of participants as well as the distinct

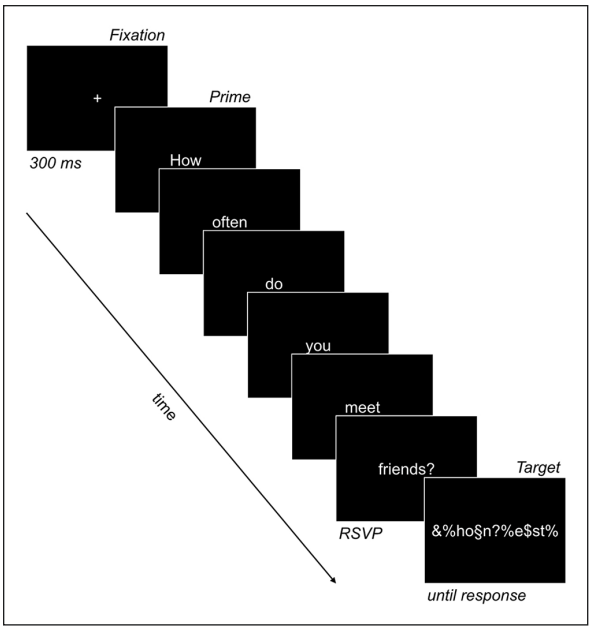

Figure 3. Trial structure of the experiment.

topic. All interviewers were female. The interviews had a duration of approximately 5 min.

After the interview, participants returned to their computers, where they had to perform the same priming task as in Koranyi et al. (2015). As primes, we used the questions from the interviews; six additional questions which had not been posed in the interviews were presented during filler trials to conceal the intention of the study. Participants were asked to decide as fast as possible whether a presented probe word was the word *ehrlich* (English: *honest*) or *gelogen* (English: *dishonest*) by pressing the *J* key for *honest* and the *F* key for *dishonest*. To ensure that probe words were encoded not only on a perceptual but also on a semantic level, they were degraded by inserting special characters between the letters (e.g., *$eh§rl#ic%h* instead of *ehrlich*). The locations of the special characters within the probe word were determined separately in each trial on a random basis. Each trial began with the presentation of a fixation cross (500 ms), followed by a serial word-by-word presentation of the prime question using RSVP (rapid serial visual presentation) with a base duration of 250 ms per word, plus an additional 25 ms per letter. Right after the last word of the sentence, the probe word appeared and remained on screen until the participant's response. The next trial was initiated after an inter-trial interval of 1000 ms (see Figure 3).

The priming task comprised 144 experimental trials with the order of trials randomised individually. To ensure comparable reliability of response time (RT) estimates for both the rare and the frequent conditions, the two prime questions in the distinct condition were presented 24 times each, while the remaining six prime questions of the frequent condition were shown 8 times each, half of the times preceding *dishonest* as probe word and half of the times *honest* as probe word. To ensure semantic processing of the prime questions, 28 experimental trials (randomly chosen out of the 144 trials) comprised an additional memory task that had to be performed directly after classifying the target word (see Wiswede et al., 2013). In the memory task, participants saw a question on the screen which was either the same (50%) or different from the prime question and participants had to answer the question, "Is this the question that you've just seen?" Participants were not informed in advance whether a trial comprised the additional memory task or not, so they had to process the prime questions in each trial.

To ensure that the honest and dishonest keys maintained their full semantic meaning across the experiment, an additional 32 filler trials were randomly intermixed into the experimental trials that required a genuine true/false decision (see Wiswede et al., 2013). In the filler trials, a true (50%) or false (50%) assertion (e.g., "Einstein was a musician") was presented as a prime instead of a question. The assertion was followed by the question, "honest or dishonest?" which was presented as a response cue instead of the probe words. Participants had to respond to this question by pressing the same keys that were also used for the honest/dishonest decision of the experimental trials. The experiment had a duration of approximately 35 min.

Results

Data treatment

All response latencies that were more than one and a half interquartile ranges above the third quartile of an individual's reaction time distribution were categorised as outliers (Tukey, 1977) and discarded (4.8% of all responses). All response latencies below the threshold of 250 ms were discarded (0.1%), as well as erroneous responses (3.5% of all responses). We calculated average response times for each participant and combination of the factorial design (see Figure 4 for the pattern of means).

Analyses of variance

To test our assumptions, average response latencies were submitted to a 2 (context: rare lies vs. rare truths) × 2 (prime question: lie vs. truth) × 2 (probe word: dishonest vs. honest) ANOVA with repeated measures on the factors prime question and probe word.[2] Results revealed a main effect of probe word, $F(1,87)=6.74$, $p=.011$, $\eta_p^2=.07$ (90% CI = [.01, .17]), indicating that participants classified the probe word *dishonest* ($M=644$ ms) faster than the probe word *honest* ($M=654$ ms). In addition, an interaction effect of context and prime question emerged, $F(1,87)=8.52$, $p=.004$, $\eta_p^2=.09$ (90 CI = [.02, .19]), which was further qualified by a significant three-way interaction, $F(1,87)=4.50$, $p=.037$, $\eta_p^2=.05$ (90% CI = [.001, .14]), as predicted by the combined model.

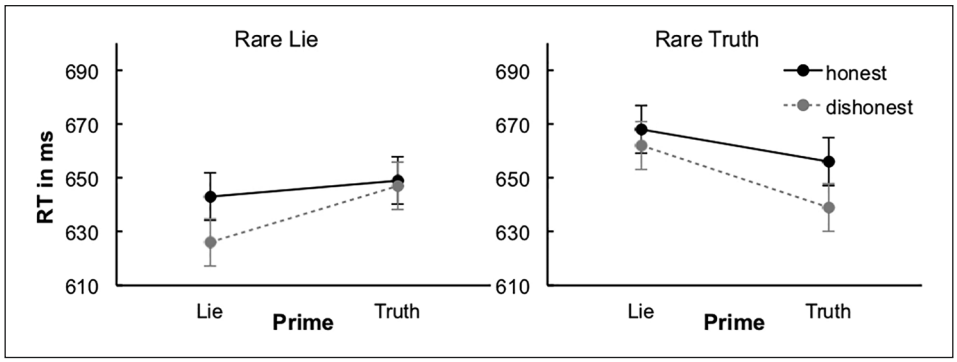

Figure 4. Mean response time (RT) as a function of context (rare lie vs. rare truth), prime question (lie vs. truth), and probe word (honest vs. dishonest). Error bars represent 95% CIs calculated for mixed IM-RM effects as suggested in Jarmasz & Hollands, 2009.

To further disentangle this interaction, we ran separate analyses for both frequency conditions. For the rare lies condition, the results revealed no significant main effects for prime question, $F(1,87)=3.21$, $p=.077$, $\eta_p^2=.04$ (90% CI = [.00, .12]), and probe word, $F(1,87)=3.03$, $p=.085$, $\eta_p^2=.03$ (90% CI = [.00, .12]). Importantly, the predicted interaction between prime question and probe word reached significance, $F(1,87)=2.89$, $p=.047$, $\eta_p^2=.03$ (90% CI = [.00, .11]; one-tailed),[3] with the descriptive pattern pointing in the predicted direction (see Figure 4). Correspondingly, follow-up tests showed that identification of the probe word *dishonest* ($M=626$ ms) was faster compared with the probe word *honest* ($M=643$ ms), $t(44)=3.61$ $p<.001$, $d=.54$ (95% CI = [.22, .85]), after the presentation of a lie prime, whereas no such difference could be found for truth primes ($t<1$). These results resemble the findings of Koranyi et al. (2015), indicating that those participants who regularly had to tell the truth and lied only to two questions during the interview showed the same congruency effect as was found before. Analyses of the rare truths condition only yielded a significant main effect for prime question, $F(1,87)=5.45$, $p=.022$, $\eta_p^2=.06$ (90% CI = [.01, .15]), but no significant interaction effect, $F(1,87)=1.69$, $p=.197$, $\eta_p^2=.02$ (90% CI = [.00, .09]), $BF_{01}=2.28$. Closer inspection of the descriptive pattern reveals two important things: First, there is no difference between probe words after the presentation of a lie prime (which would have been predicted by the quality account), and second, the difference between probe words after the presentation of a truth prime points in the opposite direction to what would have been predicted by the quantity account (i.e., faster identification of the probe word *dishonest* compared with *honest*.[4]

Bayesian statistics

To compare the probabilities for our hypotheses given the data, we also computed Bayes factors, using the R package Bain (Gu, 2016; Gu et al., 2018). Implemented in Bain is the approximate adjusted fractional Bayes factor, which can be used for the evaluation of informative hypotheses (Hoijtink, 2012). The Bayes factors and the posterior probabilities (computed assuming equal prior probabilities) are displayed in Table 1. Constraints for Hypotheses 1 (corresponding to the quality account), 2 (corresponding to the quantity account), and 3 (corresponding to the combined account) were defined in a way that they match the described data patterns in Figure 1. Note that all hypotheses are compared to an H_a, which assumes unconstrained means in all conditions. As can be seen, H_3 is supported more than both H_1 and H_2. A Bayes factor of 18.679 is usually seen as strong evidence in favour of H_3. The Bayesian error probability associated with preferring H_3 equals .281.

Table 1. Bayes factors and posterior probabilities for Hypotheses 1–3, each in comparison to H_a which assumes unconstrained means in all conditions.

Hypothesis	BF.a	Posterior probability
H_1	4.968	.191
H_2	1.325	.051
H_3	18.679	.719
H_a		.039

Discussion

The aim of the present study was to test whether the quantity and/or quality of honest and dishonest responses to questions have an influence on automatic retrieval of the knowledge of having lied to a question before. We compared the predictions of three different models, where only quantity, only quality, or the combination of them determines the storage and automatic retrieval of an episodic unit.

All models predicted the same interaction effect of prime question and probe word in the rare lies condition. This prediction was confirmed in our data, replicating the findings reported by Koranyi et al. (2015). The models differ, however, with regard to their predictions for the rare truths condition, with the quality and quantity account predicting a two-way interaction of prime question and probe word also for this condition, resulting in an absence of a

three-way interaction, whereas the combined model predicts an absence of a congruency effect within the rare truths condition, which should result in a significant three-way interaction.

Our results match with the predictions of the combined model: The three-way interaction of prime question, probe word, and context turned out to be significant, and there was no longer a significant interaction of prime question and probe word within the rare truths condition. Bayesian statistics yield additional support for the combined model. These results speak against both the quality and the quantity account. On one hand, having told a lie is not sufficient for an encoding and retrieval of knowledge about having lied, because this effect is neutralised when lies are the dominant response during the interview. Although telling lies is more cognitively demanding, this feature alone is not sufficient to produce the automatic memory retrieval effect for lies if lies are no longer distinct. On the other hand, having answered truthfully to a question does not produce a significant retrieval effect even when truthful answers are rare, and thus distinct. Thus, this finding speaks against an explanation of the retrieval effects in terms of distinctness.

Apparently, then, encoding and retrieval of knowledge about having lied to a question requires both, investment of cognitive effort (the qualitative feature of lies), and distinctness (the quantitative pre-condition of the effect). None of the two features in isolation is sufficient to produce the effect. Framed differently, both the lack of cognitive effort during response production and the lack of distinctness of the type of response that is produced to a question prevent the storage in and/or retrieval of corresponding question-response episodes from memory.

In the study by Koranyi et al. (2015), inverse efficiency scores (IES) were used as a measure to combine the effects of speed and accuracy. IES have been criticised since then for putting unequal weights on speed and accuracy depending on the level of accuracy (Liesefeld & Janczyk, 2019). This leads to unbalanced comparisons between conditions with different accuracy levels, which is why we decided to consider only RTs in this study. However, additional analyses of IES show no significant effect. This might be due to the described problems of this indicator, but also due to differences between the participants' response behaviour in both studies. When looking at the accuracy rates, it becomes apparent that our sample shows much larger standard deviations (*SD* ranging from 3.5 to 14.2) than it was described in Koranyi et al. (*SD* ranging from 3.2 to 8.9). As it is hard to predict in advance whether effects will mainly map onto reaction time or accuracy data, one promising account in the future might be the use of a different combined measure, like the balanced integration score (BIS), which is recommended by Liesefeld and Janczyk (2019).

One might argue that our findings may be driven by a carry-over effect of the experimental instruction from the interview to the priming task. According to this alternative explanation, when reading a lie question during the experimental task, participants might remember the instruction that they should lie in response to this question because it refers to the topic for which lying is requested, rather than retrieving a specific episode in which they had lied to this specific question. For the following reasons, we think that this is an unlikely explanation for our findings. First, we tested whether merely instructing to lie to a certain question can produce a similar pattern of priming effects in a previous study (Koranyi et al., 2015). In the second experiment of this study, we used additional questions as primes in the RT task which related to the same topics to which the instructions referred but which were never presented during the interview so that participants had not lied or told the truth to these questions during the interview. As these questions did not elicit any congruency effect in the priming task, we concluded that instruction effects are not responsible for our findings. Furthermore, an instruction account cannot explain why (a) priming effects were obtained only for questions to which one had lied but not for questions that were answered truthfully, nor can it explain why (b) priming effects emerged only in the condition in which lies were rare. It thus seems highly unlikely that specific processes like we predicted them for the automatic memory of having lied before resulted from mere instruction effects. For these reasons, we assume that the pattern of findings that we obtained reflects automatic storage and retrieval processes with regard to having told a lie to a specific question in our design and do not think that instruction effects played a significant role here.

To conclude, our study provides a non-trivial explanation for the well-known fact that telling lies regularly typically incurs a high chance of being detected at some point. Our study reveals that at least part of the effect seems to be due to the lack of distinctness of the lies that are told, preventing an efficient memory retrieval of knowledge of having lied, which is a prerequisite for maintaining consistency in the fabric of lies that one has told to others. Our findings thus highlight that although one might get away with an occasional lie, one should definitely avoid becoming a liar.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Franziska Schreckenbach https://orcid.org/0000-0001-7867-533X

Supplementary Material

The Supplementary Material is available at qjep.sagepub.com

Notes

1. The Koranyi et al. (2015) study did not contain the factor "frequency of lies vs. truths" but instead included the factor "old vs. new questions." We thus decided to base our sample size calculations on the size of this three-way interaction as a proxy for the strength of a three-way interaction between the factors prime question, probe word, and frequency.
2. The complete raw data as well as an analysis script for R and a table containing all cell means and standard deviations of the design can be found in the online supplementary attached to this article.
3. Methodologically, the *F*-test for the interaction is equivalent to a *t*-test that tests the difference between congruent and incongruent combinations of prime questions and probe words against zero. Thus, given our specific predictions, a one-tailed test is recommended to increase the power of the test (Maxwell & Delaney, 1990, p. 144).
4. In the study from Koranyi et al. (2015), inverse efficiency scores (IES) were used as a measure to combine effects of speed and accuracy. Although this analysis has the advantage of channelling both indicators of performance into a single DV, IES have also been criticised recently for putting unequal weights on speed and accuracy depending on the level of accuracy (Liesefeld & Janczyk, 2019). To avoid this problem, we decided to analyse only RTs in the present study. Nevertheless, to report findings that are comparable between the studies we also scrutinised IE scores for our data. The 2×2×2 ANOVA resulted in a main effect of probe word, $F(1,87) = 14.42$, $p < .001$, $\eta_p^2 = .14$. The two-way interaction of prime question × probe word ($F[1,87] = 2.79$, $p = .098$) and the three-way interaction ($F[1,87] = 2.49$, $p = .118$) with IES as DV missed significance.

References

Denkinger, B., & Koutstaal, W. (2009). Perceive-decide-act, perceive-decide-act: How abstract is repetition-related decision learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 742–756. https://doi.org/10.1037/a0015263

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, *70*, 979–995. https://doi.org/10.1037/0022-3514.70.5.979

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96(3), 433–458. https://doi.org/10.1037/0033-295X.96.3.433

Frings, C., Hommel, B., Koch, I., Rothermund, K., Dignath, D., Giesen, C., . . .Philipp, A. (2020). Binding and retrieval in action control (BRAC). *Trends in Cognitive Sciences*, *24*, 375–387. https://doi.org/10.1016/j.tics.2020.02.004

Frings, C., Koch, I., Rothermund, K., Dignath, D., Giesen, C., Hommel, B., . . .Philipp, A. (2020). Merkmalsintegration und Abruf als wichtige Prozesse der automatischen Handlungssteuerung—eine Paradigmen-übergreifende Perspektive [Feature integration and retrieval in action control-a perspective across paradigms]. *Psychologische Rundschau*, *71*, 1–14. https://doi.org/10.1026/0033-3042/a000423

Frings, C., & Rothermund, K. (2011). To be or not to be. . . included in an event file: Integration and retrieval of distractors in stimulus-response episodes is influenced by perceptual grouping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1209–1227. https://doi.org/10.1037/a0023915

Frings, C., & Rothermund, K. (2017). How perception guides action: Figure-ground segmentation modulates integration of context features into S-R episodes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 1720–1729. https://doi.org/10.1037/xlm0000403

Gu, X. (2016). *Bayesian evaluation of informative hypotheses* [Doctoral dissertation]. University Utrecht. https://informative-hypotheses.sites.uu.nl/books-and-theses/

Gu, X., Mulder, J., & Hoijtink, H. (2018). Approximated adjusted fractional Bayes factors: A general method for testing informative hypotheses. *British Journal of Mathematical and Statistical Psychology*, *71*, 229–261. http://doi.org/10.1111/bmsp.12110

Hoijtink, H. (2012). *Informative hypotheses: Theory and practice for behavioral and social scientists*. Chapman and Hall/CRC.

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition*, *5*, 183–216. https://doi.org/10.1080/713756773

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, *8*, 494–500. https://doi.org/10.1016/j.tics.2004.08.007

Hunt, R. R. (2006). The concept of distinctiveness in memory research. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and memory* (pp. 3–25). Oxford University Press.

Jarmasz, J., & Hollands, J. G. (2009). Confidence intervals in repeated-measures designs: The number of observations principle. *Canadian Journal of Experimental Psychology*, *63*(2), 124–138. https://doi.org/10.1037/a0014164

Koranyi, N., Schreckenbach, F., & Rothermund, K. (2015). The implicit cognition of lying: Knowledge about having lied to a question is retrieved automatically. *Social Cognition*, *33*, 67–84. https://doi.org/10.1521/soco.2015.33.1.67

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods*, *51*, 40–60. https://doi.org/10.3758/s13428-018-1076-x

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527. https://doi.org/10.1037/0033-295X.95.4.492

Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analyzing data: A model comparison perspective*. Wadsworth.

Rothermund, K., Wentura, D., & De Houwer, J. (2005). Retrieval of incidental stimulus-response associations as a source of negative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 482–495. https://doi.org/10.1037/0278-7393.31.3.482

Schmidt, J. R., De Houwer, J., & Rothermund, K. (2016). The Parallel Episodic Processing (PEP) model 2.0: A single computational model of stimulus-response binding, contingency learning, power curves, and mixing

costs. *Cognitive Psychology*, *91*, 82–108.https://doi.org/10.1016/j.cogpsych.2016.10.004.

Schreckenbach, F., Rothermund, K., & Koranyi, N. (2019). And remember the truth that once was spoken: Knowledge of having disclosed private information to a stranger is retrieved automatically. *Experimental Psychology*, *66*, 12–22. https://doi.org/10.1027/1618-3169/a000427

Serota, K. B., Levine, T. R.:., & Poster, F. J. (2010). The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research*, *36*, 2–25. https://doi.org/10.1111/j.1468-2958.2009.01366.x

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology*, *20*, 421–446. https://doi.org/10.1002/acp.1190

Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception. A meta-analytic synthesis. *Psychology, Public Policy, and Law*, *13*, 1–34. https://doi.org/10.1037/1076-8971.13.1.1

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.

Tulving, E., & Thomson, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373. https://doi.org/10.1037/h0020071

von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld [The effects of field formation in the trace field]. *Psychologische Forschung*, *18*, 299–342. https://doi.org/10.1007/BF02409636

Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological Science in the Public Interest*, *11*, 89–121. https://doi.org/10.1177/1529100610390861

Walczyk, J. J., Harris, L. L., Duck, T. K., & Mulay, D. (2014). A social-cognitive framework for understanding serious lies: Activation-decision-construction-action theory. *New Ideas in Psychology*, *34*, 22–36. https://doi.org/10.1016/j.newideapsych.2014.03.001

Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology*, *17*, 755–774. https://doi.org/10.1002/acp.914

Wiswede, D., Koranyi, N., Müller, F., Langner, O., & Rothermund, K. (2013). Validating the truth of propositions: Behavioral and ERP indicators of truth evaluation processes. *Social Cognitive and Affective Neuroscience*, *8*, 647–653. https://doi.org/10.1093/scan/nss04

3 And Remember the Truth That Once was Spoken: Knowledge of Having Disclosed Private Information to a Stranger is Retrieved Automatically

If you want to be thought a liar,
always tell the truth.
(Logan Pearsall Smith)

The following article has been published as:

Schreckenbach, F., Rothermund, K., & Koranyi, N. (2019). And remember the truth that once was spoken: Knowledge of having disclosed private information to a stranger is retrieved automatically. *Experimental Psychology, 66*(1), 12-22. https://doi.org/10.1027/1618-3169/a000427

The supplemental online material is presented in Appendix B.

Research Article

And Remember the Truth That Once Was Spoken

Knowledge of Having Disclosed Private Information to a Stranger Is Retrieved Automatically

Franziska Schreckenbach, Klaus Rothermund, and Nicolas Koranyi

Department of General Psychology II, Institute of Psychology, Friedrich Schiller University Jena, Germany

Abstract: Whenever individuals reveal personally relevant information to a stranger, they have to remember their self-disclosure for future interactions. Relying on instance-based theories of automaticity, we hypothesized that knowledge about having revealed private information to someone unfamiliar is retrieved automatically whenever this person is encountered again. In two studies, participants were orally interviewed by two different persons and instructed to be honest to one of them and to lie to the other. This instruction was either related to the identity of the interviewers (Experiment 1) or their gender (Experiment 2). Afterward, the target words *honest* and *dishonest* had to be identified in a categorization task in which pictures of the interviewers and of unknown persons served as task-irrelevant prime stimuli. In line with the hypothesis, results revealed congruence effects, indicating faster identification of the target word *honest* following the picture of a person whom one had told the truth.

Keywords: self-disclosure, lying, automatic processes, instance-based learning

Typically, individuals entrust intimate details about the incidents in their lives, their thoughts and feelings only to very few people in their social environment (DePaulo & Kashy, 1998; Dindia, Fitzpatrick, & Kenny, 1997). These confidants are usually close friends or family members, like our partner, our siblings, or our parents. These social bonds are based on feelings of intimacy and a long history of mutual trust and reciprocity (Neyer, Wrzus, Wagner, & Lang, 2011). However, on some occasions, individuals may feel the urge to talk to someone about a personal issue, for instance, because none of their intimates is available or because they want to talk about a secret they want to hide away from these very persons. If such a situation arises, people sometimes confide in a total stranger, which is known as the "stranger on a train" phenomenon (Rubin, 1975). For instance, this might happen when meeting an interesting person at a party and immediately starting an intense discussion about a personally relevant topic, or when simply sitting alone with another person in a train coach while having the need to talk about a major problem.

Whenever individuals have spoken openly and frankly to an unfamiliar person, they are well advised to remember that they had been honest to better regulate future interactions with this person, who has become an insider. For instance, the insider might use the private information to one's disadvantage, and it is therefore important to have an eye on him or her. Alternatively, the insider might also be more open or more loyal to us, due to the fact that we have been as open before. Finally, when meeting the person again and acting according to the default mode (e.g., withholding personally relevant information), the person might be irritated and might even identify the incorrectness of one's statements due to the incongruence with the original conversation.

In the present research, we want to test the idea that remembering one's self-disclosure toward someone unfamiliar is achieved by an automatic mechanism that comes into effect as soon as one has confided in an unknown person. Specifically, we test the hypothesis that a previously unknown conversation partner and the fact that one has been honest are bound together, forming an episodic unit that is stored in memory. When the person is encountered again, the information about having been honest is automatically retrieved from the stored episode.

The broader theoretical framework behind this idea is provided by instance-based or episodic theories of memory

storage and retrieval (Logan, 1988; see also Denkinger & Koutstaal, 2009; Hommel, 1998, 2004; Rothermund, Wentura, & De Houwer, 2005). Within these theories, it is proposed that when an action is performed, information about this action is stored together with situational cues as an episodic unit (alternatively called an *event file*, Hommel, 1998, 2004, *instance*, Logan, 1988, or *stimulus-response episode*, see Rothermund et al., 2005). These episodic units become automatically reactivated when encountering a similar situation again and therefore enable people to react fast and consistently across different situations. Furthermore, recent findings suggest that not only the motoric response toward a specific stimulus itself is encoded as part of an episode, but that relevant meta-knowledge regarding the task context (Horner & Henson, 2009, 2011; Waszak, Hommel, & Allport, 2003, 2005), the semantic meaning (Giesen & Rothermund, 2016; Moutsopoulou, Yang, Desantis, & Waszak, 2015), and the veracity of an action (Koranyi, Schreckenbach, & Rothermund, 2015) is likewise encoded in episodic format and later becomes retrieved when the eliciting situation is encountered again.

A second theoretical idea that led to our research hypothesis relates to the finding that lying is a widespread and sometimes even adaptive phenomenon, with more than 90% of all people using deception at least once in a week and with lies possibly helping to avoid conflicts and to maintain positive relationships (e.g., DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996; Ennis, Vrij, & Chance, 2008; Metts, 1989; Serota, Levine, & Boster, 2010). There is also evidence that lies are told more frequently to strangers than to friends (DePaulo & Kashy, 1998). Furthermore, the widespread assumption that lies are cognitively more demanding than telling the truth has been recently challenged by results from various studies (e.g., Barnes, Schaubroeck, Huth, & Ghumman, 2011; McCornack, Morrison, Palik, Wisner, & Zhu, 2014; Yam, Chen, & Reynold, 2014).

These lines of research provide some evidence for our assumption that telling an unknown person something truthful is not necessarily the default option in every conversation and may even be experienced as being more demanding, especially when personal information is concerned. Given these findings and the fact that the likelihood of encoding distinct or unusual events in memory is higher compared to common events (Cohen & Carr, 1975; see also Brandt, Gardiner, & Macrae, 2006), we assume that knowledge about having disclosed some personally relevant truth to a stranger should become stored in memory because of its distinctiveness and is automatically activated on a later occasion.

We test this idea in two experiments with the same core paradigm, consisting of two parts: First, participants either had to tell the truth or had to lie to different unfamiliar persons in oral interviews. Afterward, facial photographs of their interviewers served as task-irrelevant prime stimuli during a classification task in which the target words *dishonest* and *honest* had to be identified. Based on the assumption that a person becomes associated with the truth status of the responses that were given to this individual, we predicted a congruence effect, that is, faster identification of the target words following a matching picture prime. Specifically, we predicted faster identification of the word *honest* compared to the word *dishonest* after the presentation of a picture of a person to whom one had told the truth before, and faster identification of the word *dishonest* compared to the word *honest* after the presentation of a picture of a person one has lied to before. In this study, we report all measures, manipulations, and exclusions.

Experiment 1

Experiment 1 was a direct test of our hypothesis that knowledge about having told the truth to someone unfamiliar before can be activated automatically in a subsequent similar situation. We also examined whether this retrieval process is specific for truthful situations or deceptive ones by inserting not only pictures of interviewers that participants had to lie to or tell the truth before, but also pictures of unknown persons.

Method

Participants and Design

The sample size was determined relying on a previous experiment by Koranyi et al. (2015), who found an effect size of $\eta_p^2 = .16$ in a similar paradigm. Accordingly, a power analysis was conducted based on this value, $\alpha = .05$ and a power of $1 - \beta = .8$, leading to a sample size of 45. To account for possible exclusions, we recruited fifty students from a German University. The data of two subjects were excluded because of extremely slow reactions, with their mean response time lying more than three interquartile ranges above the third quartile of the mean reaction times of the sample. The final sample then consisted of 48 subjects (33 female) with an average age of $M = 22.0$ years ($SD = 3.16$). They participated in exchange for partial course credit and a chocolate bar. Recruiting was finished after the required number had been reached. We used a 3×2 factorial design with the within-subjects factors prime picture (facial photograph of interviewer one has told the truth vs. lied to vs. not encountered before) and target word (honest vs. dishonest).

Procedure and Materials

Upon arrival, participants were seated in front of a computer and received instructions on the screen. As picture

primes in the priming task, we used facial photographs of the four interrogators on which they wore the same clothes, hairstyle, and accessories as in the actual interviews. All pictures had a size of 400 × 600 px, with each person's eyes placed at the same position in the upper half of the picture (approx. 200 px below the top).

Interviews

First, participants were informed that they would be interviewed twice by different interrogators and that their task was to answer all questions in one interview truthfully, but to lie to all questions in the other one. Additionally, participants were prompted not to reveal their dishonesty to the respective interviewer and to act so as to convince both interviewers of the truthfulness of their statements. In order to achieve this goal, participants were instructed to always wait for some seconds before providing each answer. Afterward, a picture of the first interviewer appeared on the left side of the screen, while on the right side the interviewer's name, the topic that participants would be asked about, and the instruction whether to tell the truth or to lie was presented (e.g., "This is Sascha. Sascha is going to ask you some questions on the subject of your family. Please always tell the truth."). Complementary information was given about the second interviewer (e.g., "This is Sarah. Sarah is going to ask you some questions on the subject of your studies at the university. Please always respond with a lie."). Participants were then guided to separate rooms where the interviews took place. Every participant was interrogated by one of two possible couples of interviewers, with pictures of the other couple serving as a neutral baseline in the subsequent priming task. Each couple consisted of a man and a woman, the assignment of truthfulness and lying to interviewer and topic was counterbalanced across participants, as was the order of interviews. Four questions were related to the topic *family*, the other ones referred to the topic *university* (see Electronic Supplementary Material, ESM 1 for a complete list of all questions). Each question was created to touch a personally important issue but at the same time should not be too intimate in order to prevent participants from answering untruthfully even if they were instructed to tell the truth.

Priming Task

After the interviews, participants performed a priming task on the computer to assess whether the specific persons from the interview automatically trigger retrieval of knowledge about having previously told the truth or a lie, respectively. Each trial had the same temporal structure (see Figure 1): A fixation cross (500 ms) was followed by a prime picture, showing one of the four interviewers. After 300 ms, the target word appeared upon the prime picture, approximately at the position where the eyes in the pictures were located (200 px below the top). Participants were asked to decide as fast as possible whether a presented target was the word *ehrlich* (English: *honest*) or *gelogen* (English: *dishonest*, having lied) by pressing the "J" key for *honest* and the "F" key for *dishonest*. To ensure that target words were encoded and identified not only on a perceptual but also on a semantic level, they were degraded by inserting alphanumeric characters between the letters (e.g., *§$eh&$r§li#c%h* instead of *ehrlich*). The locations of these additional characters within the target words were determined randomly for each trial, ensuring a large degree of variability between the target stimuli. Both stimuli (i.e., the facial photograph of the interviewer and the target word) remained on the screen until the participant responded by pressing either the J or F key on the computer keyboard. The next trial was initiated after an inter-trial-interval of 750 ms. The priming task comprised 192 experimental trials with order of trials randomized individually. All four prime pictures were presented 48 times each, half of the times preceding *dishonest* as target word and half of the times *honest* as target word.

To ensure that the *honest* and *dishonest* keys maintained their semantic meaning across the experiment, 60 additional filler trials were randomly intermixed into the experimental trials that required a genuine true/false decision (see Eder & Rothermund, 2008; Wiswede, Koranyi, Müller, Langner, & Rothermund, 2013). In the filler trials, a true (50%; e.g., "Saturn is a planet") or false assertion (50%; e.g., "Einstein was a musician") was presented word by word in the center of the screen instead of a prime picture. The assertion was followed by the question "honest or dishonest?" that was presented as a response cue instead of a target. Participants had to evaluate the truth of the previously presented sentence by pressing the same keys that were also used for the honest/dishonest classification of the experimental trials. The priming task had a total duration of approximately 20 min.

Results

To combine effects of both speed and accuracy in a single dependent variable, we computed inverse efficiency scores (IES; see Table 1) by dividing RTs by the proportion of correct responses (i.e., 1 − error rate) separately for all combinations of the factorial design.[1] Small values on the IES indicate fast responses and/or low error rates, whereas large values indicate slow responding and/or high error frequencies. Calculating IESs is an established method for

[1] The complete raw data as well as an analysis script for R can be found in the ESMs 2–5 attached to this paper.

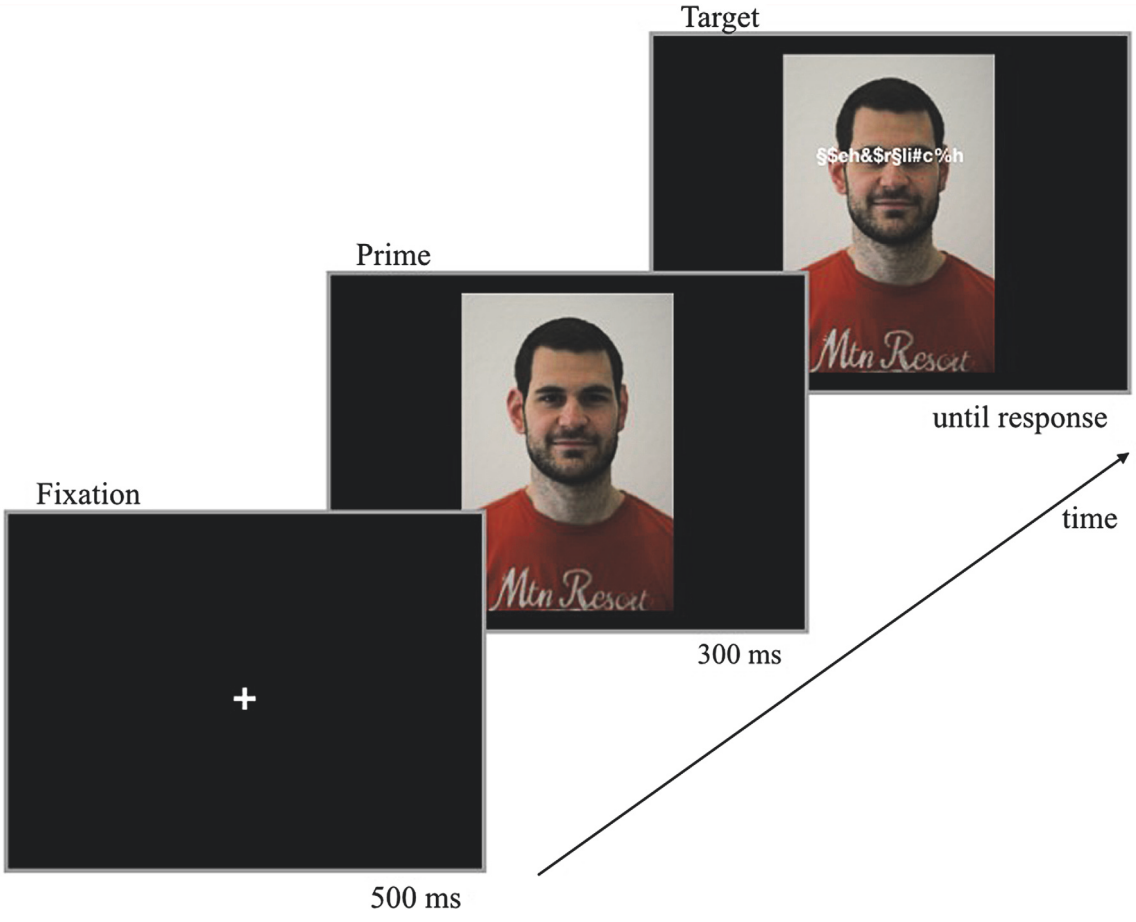

Figure 1. Trial structure of Experiment 1.

Table 1. Mean response time (RT), accuracy (%), and inverse efficiency scores (IES: RT divided by proportion of correct responses) as a function of prime picture (interviewer to whom one had told the truth, a lie, or nothing) and target word (honest vs. dishonest)

	Prime picture condition					
	Truth		Lie		Baseline	
	Target word					
DV	Honest (*SD*)	Dishonest (*SD*)	Honest (*SD*)	Dishonest (*SD*)	Honest (*SD*)	Dishonest (*SD*)
RT	507 (65)	536 (80)	518 (67)	519 (69)	518 (64)	525 (73)
Accuracy (%)	95.8 (4.8)	94.3 (6.5)	94.4 (6.1)	95.6 (5.5)	95.4 (4.1)	95.9 (3.4)
IES	529 (65)	572 (103)	550 (71)	546 (83)	543 (62)	548 (75)

merging the effect variance of response times (RT) and error rates into a single index that also prevents biased results due to differences in speed/accuracy tradeoffs between conditions (Bruyer & Brysbaert, 2011; Koranyi et al., 2015; Townsend & Ashby, 1983). RT in the experimental task that were more than three interquartile ranges above the third quartile of an individual's RT distribution were categorized as far-out values (Tukey, 1977) and therefore discarded (1.4% of all RTs). All RTs below the threshold of 250 ms were discarded as well (0.1%).

The IESs were submitted to a 3 (Prime Picture: interviewer who was told the truth vs. a lie vs. neutral [not encountered]) × 2 (Target Word: honest vs. dishonest) analysis of variance (ANOVA) with repeated measures on both factors. We also specified two a priori contrasts for the factor prime picture to test whether retrieval of information occurred for both truthful and untruthful statements or mainly for one of them. The first contrast was specified to compare truth primes (i.e., pictures of the interviewer who was told the truth) with neutral primes (i.e., pictures of the unknown interviewers), whereas the second contrast compared lie primes with neutral primes.

Results revealed a main effect of target word, $F(1, 47) = 4.64$, $p = .037$, $\eta_p^2 = .09$, indicating that on average,

participants classified the target honest (M = 541) faster than the target dishonest (M = 555), whereas no main effect for prime picture (F < 1) emerged. Additionally, the predicted interaction of both factors was found, $F(2, 46) = 5.83$, $p = .006$, $\eta_p^2 = .20$. Further inspection of planned contrasts revealed that the interaction effect was due to the interaction of target word and the truth versus neutral contrast, $F(1, 47) = 8.31$, $p = .006$, $\eta_p^2 = .15$. In line with the assumption of congruency, IESs were smaller for honest (M = 529) versus dishonest (M = 572) targets in the truth condition, $F(1, 47) = 12.14$, $p = .001$, $\eta_p^2 = .21$, whereas no such difference was found for neutral primes (M = 543 and M = 548, F < 1). No effect was found for the interaction between target word and the lie versus neutral contrast (F < 1).[2]

We also scrutinized reaction times and accuracy separately. Results revealed a significant main effect of target word, $F(1, 47) = 11.79$, $p = .001$, $\eta_p^2 = .20$, for the response time data, as well as a significant interaction between prime picture and target word, $F(2, 46) = 5.20$, $p = .009$, $\eta_p^2 = .18$. For the accuracy data, no significant effect emerged (all F < 2.5), but the descriptive pattern of results stayed the same, with error rates being smaller in the case of congruent primes and targets compared to incongruent primes and targets.

Discussion

The aim of the present experiment was to test the hypothesis that knowledge about having been honest and open to a stranger is activated automatically when encountering this person again on a later occasion. In line with this hypothesis, we found the predicted facilitation effect for the word *honest* when the picture of an interviewer to whom one had revealed personal information was subsequently presented in a word identification task. Apparently, knowledge about having told the truth to a specific person is retrieved automatically when encountering this person again.

There are some limitations of Experiment 1 that we want to consider here further. First, the assignment of response keys to honest/dishonest responses was not counterbalanced across participants, with "honest" always lying on the right side and "dishonest" on the left side of the response keyboard. This might have created an asymmetry in responding that may have made it more difficult to find an effect in the dishonest condition. Although we think that such an influence on the interaction we found is unlikely, a replication with counterbalanced key assignments would be desirable to provide more certainty regarding this point.

A second shortcoming of Experiment 1 is that we cannot rule out the possibility that the congruence effect was driven by a carry-over effect from the instructions. Participants might have remembered the explicit instruction to lie to a specific person during the interview also when seeing the corresponding picture in the priming procedure, which could explain the findings of Experiment 1. We think that this is an unlikely explanation because, if responses were indeed influenced by the interview instructions, one would expect prime-target congruence effects for both conditions, that is, not only for pictures of interviewers to whom one had told the truth but also for those to whom one had lied, which was not the case. Furthermore, we tested the effect of instruction-based carry-over effects for instructed lies and truths in a previous study with a highly similar paradigm and found no evidence for mere instruction-based effects in the absence of actual episodes of lying or truth telling (Koranyi et al., 2015). Still, as the presented data of the first experiment provide no safe evidence against this alternative explanation, we conducted a second experiment that allowed us to test whether the congruence effect was mainly driven by an explicitly instructed association or by episodic encoding and retrieval of the behavior that was actually shown during the interview situation.

Experiment 2

Experiment 2 used the same interview and priming task described in Experiment 1. In order to test carry-over effects from the instructions for the interviews to the priming task, participants now received generic instructions regarding the gender of interviewers to whom they had to lie or tell the truth, respectively, instead of receiving instructions regarding the identity of these interviewers (e.g., "respond with a lie to questions posed by women, respond truthfully to questions posed by men"). In the following priming task, we again presented pictures of the interviewers as well as pictures of unknown persons. If the congruence effect in Experiment 1 was due to instructions, the same response pattern should be found for both pictures of actual interviewers and of unknown persons of the same sex. That is, if participants were instructed to tell the truth

[2] Although we counterbalanced the assignment of truth status to interviewer, it is possible that the congruency effect differs for different combinations of participant gender and interviewer gender. To examine this assumption, we ran additional analyses, including the factors participant sex (male vs. female) and sex of the interviewer who had to be told the truth (male vs. female). An interaction of both factors could be interpreted in terms of an effect of gender match/mismatch on the automatic retrieval effect of honest responses. However, these analyses yielded no significant results besides the interaction of prime and target (all other Fs < 2).

to women but to lie to men, every time a picture of a woman is presented, classification of the target "honest" should be faster compared to the target "dishonest," no matter if this person was the former interviewer or not. In contrast, if automatic binding and retrieval reflects automatic encoding and retrieval of actual episodes of lying and truth telling, then the congruence effect should emerge only after the presentation of pictures of the actual interviewers, but not after the presentation of a person who had not been encountered as an interviewer during the interview.

Method

Participants and Design

The sample size was determined relying on the results of Experiment 1. Specifically, we analyzed the interaction between prime picture (honest vs. dishonest) and target word (honest vs. dishonest) and found an effect size of $\eta_p^2 = .19$. Accordingly, a power analysis was conducted based on this value, $\alpha = .05$ and a power of $1 - \beta = .95$, to give the proposed alternative explanation in terms of instruction-based carry-over effects a fair chance to emerge in the non-encountered condition, but also to be sure that there is no such influence in case of nonsignificant priming effects in this condition. This analysis resulted in a sample size of $n = 59$. To account for possible exclusions, we recruited 68 students from a German University. For their participation, they received 2.50 € and a chocolate bar. Recruiting was finished after the required number had been reached. The data of 8 subjects could not be meaningfully analyzed because of error rates higher than 25% (average 4.9% errors, without excluded participants). Two subjects were excluded because of extremely slow reactions, with their mean response time lying more than three interquartile ranges above the third quartile of the mean reaction times of the sample. Additionally, three participants had to be excluded because they did not remember the instructions to whom to lie correctly. Therefore, the final sample consisted of 55 participants (37 female) with an average age of $M = 23.4$ years ($SD = 3.38$).

We used a 2 × 2 × 2 factorial design with the within-subjects factors prime picture truth status (facial photograph of interviewer gender who had to be told the truth vs. had to be lied to), prime picture encounter status (picture of an interviewer one has encountered vs. not encountered during the interview), and target word (honest vs. dishonest).

Procedure and Materials

All procedures and measures were the same as in Experiment 1, except for the changes reported here. Participants again took part in an oral interview and subsequently performed the priming task. After reading the general instructions about truth telling and lying in the upcoming two interviews, they received the explicit instruction that they should lie to the man and tell the truth to the woman (or vice versa, counterbalanced across participants). The remainder of the interview part was the same as in Experiment 1. Participants were interrogated by one of two possible couples of interviewers (counterbalanced across participants), with pictures of the other couple being used for the not-encountered interviewers condition in the subsequent priming task.

The priming task on the computer was the same as in Experiment 1 with some minor changes: Response keys were D and L and the assignment of target words (honest vs. dishonest) to response keys was counterbalanced across participants. Thirty filler trials that required a genuine true/false decision (Eder & Rothermund, 2008; Wiswede et al., 2013) were intermixed into the experimental trials. In Experiment 1, we were not able to control whether participants followed our instructions, which is why we added some final questions after they had finished the priming procedure. These questions were asked in order to check whether they remembered their instructions to whom to tell the truth and whether they had known one of their interviewers in advance (see ESM 1 for these questions too). The whole experiment had a duration of about 25 min.

Results

We excluded outlier values (2.2%) from the response latency data according to the same criteria as in Experiment 1, and calculated IES for all combinations of the factorial design. The IESs were submitted to a 2 (Prime Picture Truth Status: interviewer gender who had to be told the truth vs. had to be lied to) × 2 (Prime Picture Encounter Status: interviewer encountered vs. not encountered during the interview) × 2 (Target Word: *honest* vs. *dishonest*) ANOVA with repeated measures on all factors. All cell means are depicted in Table 2 (also for RTs and error rates).

Results revealed a main effect of Target Word, $F(1, 54) = 4.79$, $p = .033$, $\eta_p^2 = .08$, indicating that on average, participants classified the target honest ($M = 626$) faster than the target dishonest ($M = 638$). We also found a significant three-way interaction, $F(1, 54) = 2.99$, $p = .045$ (one-tailed),[3] $\eta_p^2 = .05$. To unpack this interaction, we ran

[3] Methodologically, the *F*-test for the interaction is equivalent to a *t*-test that tests the difference between congruency effects for pictures of persons who were encountered and not encountered during the interview against zero. Thus, given our specific predictions, a one-tailed test is recommended in order to increase the power of the test (Maxwell & Delaney, 1990, p. 144).

Table 2. Mean response time (RT), accuracy (%), and inverse efficiency scores (IES: RT divided by proportion of correct responses) as a function of prime picture (interviewer to whom one had to tell the truth or a lie), target word (honest vs. dishonest), and picture familiarity (interviewer whom one had seen during the interview vs. not seen during the interview): Experiment 2

		Prime picture condition			
		Truth		Lie	
		Target word			
DV	Familiarity	Honest (*SD*)	Dishonest (*SD*)	Honest (*SD*)	Dishonest (*SD*)
RT	Seen during the Interview	592 (97)	610 (100)	601 (102)	607 (115)
	New	592 (95)	607 (112)	593 (89)	607 (113)
Accuracy (%)	Seen during the interview	95.9 (6.0)	94.9 (6.8)	94.1 (8.0)	95.5 (6.4)
	New	94.6 (5.5)	95.4 (5.5)	95.8 (5.9)	95.9 (5.8)
IES	Seen during the interview	618 (96)	646 (120)	641 (107)	637 (116)
	New	627 (100)	637 (118)	621 (100)	635 (124)

Note. Standard deviations appear in parentheses.

separate 2 × 2 ANOVAs with the factors prime picture truth status and target word for pictures of the previously encountered and not-encountered interviewers. For known interviewers, a marginally significant main effect of target word emerged, $F(1, 54) = 3.31$, $p = .074$, $\eta_p^2 = .06$, as well as the predicted interaction of Prime Picture × Target Word, $F(1, 54) = 2.92$, $p = .047$ (one-tailed), $\eta_p^2 = .05$. In line with the findings from Experiment 1, for honest interviewers, IES were faster for "honest" ($M = 618$) than for "dishonest" ($M = 646$) targets, $F(1, 54) = 5.73$, $p = .02$, $\eta_p^2 = .10$, whereas for interviewers that had to be lied to, no corresponding effect emerged ($M = 641$ and $M = 637$, respectively; $F < 1$). In contrast, for unknown interviewers, no Prime Picture × Target Word interaction was observed ($F < 1$), indicating that pictures of people who were not encountered as interviewers during the interview did not have an influence on performance during the classification task.[4]

We again scrutinized reaction times and accuracy separately. For the response time data, only a significant main effect of target word was found, $F(1, 54) = 9.70$, $p = .003$, $\eta_p^2 = .15$, but the descriptive pattern of results stayed the same, with reaction times being faster in the case of congruent primes and targets compared to incongruent primes and targets in case of interviewers who were encountered during the interview, while no systematic effect can be seen for pictures of interviewers who did not take part in the interview. For the accuracy data, results revealed a significant interaction between Prime Picture Truth Status × Prime Picture Encountered Status, $F(1, 54) = 4.11$, $p = .048$, $\eta_p^2 = .07$, which was further qualified by a significant three-way interaction, $F(1, 54) = 3.12$, $p < .042$ (one-tailed), $\eta_p^2 = .06$. Congruency effects were obtained for encountered interviewers but not for not-encountered interviewers.

Discussion

The results of Experiment 2 replicate the findings from Experiment 1 by showing that knowledge about having told the truth to an unknown person is retrieved automatically. Furthermore, they support the claim that congruence effects reflect automatic encoding and retrieval of meta-knowledge about the actual interview situation rather than carry-over effects of instructions: Because priming effects only emerged for pictures of actual interviewers but not for pictures of not-encountered interviewers of the same sex, results cannot be explained in terms of instructions, which were framed in generic terms referring to interviewer gender in the second experiment. However, it should be noted that the effect sizes in Experiment 2 were somewhat weaker than in Experiment 1.

General Discussion

In two experiments, we found evidence for the assumption that the knowledge about having been honest to a stranger is retrieved automatically. When primed with pictures of a person to whom one had to tell the truth before, response times were faster and more accurate for "honest" versus "dishonest" targets, indicating that reencountering this

[4] Furthermore, we again looked for effects of gender assignment by adding the factors participant sex (male vs. female) and truth sex (male vs. female). Results revealed a marginally significant interaction of sex and prime, $F(1, 51) = 3.45$, $p = .069$, $\eta_p^2 = .06$, indicating that male participants responded faster after honest primes while female participants responded faster after dishonest primes, as well as a significant interaction of truth sex and prime, $F(1, 51) = 13.98$, $p < .001$, $\eta_p^2 = .22$, which was further qualified by a marginally significant interaction of truth sex, prime and familiarity, $F(1, 51) = 3.28$, $p = .076$, $\eta_p^2 = .06$, which cannot be interpreted in meaningful ways. No further connections with gender could be observed in the present data.

person automatically activated knowledge about having been honest before. No prime-target congruence effect was found for pictures of persons who had not been part of the interview, even if participants should have told them the truth given the instructions (e.g., presentation of a male person as a prime when the instruction was to tell the truth to a male person). These findings suggest that the information about having disclosed intimate or personal information to a person is a diagnostic piece of information that is bound to this person and stored in memory, ready to be automatically activated when this person is encountered again on a later occasion. Such a mechanism of marking previously unfamiliar persons in whom one has confided, and automatically retrieving this information when this person is encountered again helps to behave adequately and consistently in future interactions with this person.

Asymmetry of Associating Strangers With Knowledge About Having Told the Truth or a Lie

Importantly, we did not find evidence of a similar marking of strangers to whose questions one had responded untruthfully, that is, not revealing anything intimate or personal. Storing and retrieving information about persons to whom one has either told a personal truth or has avoided this by having told a lie thus is apparently asymmetric. To explain this asymmetry, we assume that only information that deviates from the default becomes salient enough to warrant episodic storage and retrieval. In accordance with the idea that other people who are unfamiliar are typically not entrusted with personal information, evading the truth, not being honest, or telling a lie in response to personal questions thus seems to be the standard or "unmarked" (Hamilton & Deese, 1971) form of interaction, which is not salient enough to warrant to be bound and stored with the unfamiliar person into episodic memory. Not having revealed anything personal to this person – even having told an outright lie – may thus not leave any long-lasting traces with regard to this person. On the contrary, having an intimate and honest conversation with a previously unfamiliar stranger stands out against this background of superficial and non-committal interactions and thus represents a "marked" form of discourse. In these distinct situations, the person becomes bound to the knowledge of having been honest, and this episodic association is automatically retrieved on later encounters with this person.

It is noteworthy that this asymmetry of storing knowledge about having told either the truth or a lie to a specific person contrasts with what we found with regard to episodic storage and retrieval of truth- and falsehood-related information regarding specific questions to which one has responded truthfully or with a lie (Koranyi et al., 2015). In this previous study, we found that specific questions were associated with the knowledge that one lied to this question but were not bound with the information that one answered truthfully to a question.

When explaining this apparent discrepancy in the pattern of results, a first important point is to emphasize that the findings of the two studies are not incompatible. Instead, the two studies differed with regard to what they tested: Whereas the current study investigated whether persons became associated with knowledge regarding having told the truth or a falsity, the previous study by Koranyi et al. (2015) investigated whether questions acquired such associations. It is perfectly possible that in both studies, both types of associations emerged and were simultaneously stored in memory – with different asymmetries. We cannot test this possibility because only one type of associations was assessed by presenting either pictures of persons (current study) or questions (previous study) as primes in the test task. Had we presented also questions as primes in the current study, we might also have found priming effects for questions to which one had lied (facilitating detection of the target word *dishonest*) but no effects for questions to which one had responded truthfully, as in the previous study by Koranyi et al. (2015), despite the fact that we found an opposite asymmetry with regard to the episodic storage of person-related information. Thus, a first possibility is that for persons and questions, different types of episodes become encoded into memory: Previously unfamiliar persons can become associated with information of having told the truth but not with having told a lie, whereas questions can only become associated with the knowledge that one has lied to this question but not with the information that one has answered truthfully.

It is also possible that the interview contexts of the two studies differed, leading to different asymmetries in storing and retrieval of truth-related information. The current study focused on persons as the decisive element that determined whether participants had to respond honestly or dishonestly, regardless of the questions that were posed. In this context, it would be natural to focus one's attention onto one of the two interviewers, and apparently, people remembered the interviewer to whom they had revealed personal information more, linking him or her with this information, compared to the other interviewer to whom they had not told the truth. The previous study, on the contrary, specified the content of the question as the decisive element that determined the type of the response – we did not even vary the interviewer in the previous study but used only one interviewer who posed different questions, some of which had to be answered truthfully whereas other had to be answered with a lie. Focusing on questions might have activated a different salience asymmetry, with truthful answers

being the natural, default response to a question, whereas lies represented the salient alternative to this default. According to this account, the context determines whether persons or questions are attended and will be stored in memory; this focus of attention in turn determines whether honest answers or lies are the salient mode of responding that deviates from the default and will be stored together with the person or question in episodic memory.

Furthermore, even more complex hypotheses are possible as well: For instance, one can easily conceive of an interview in which the type of response is determined by the combination of person and question (e.g., always tell the truth for questions regarding Topic A if they are posed by Person X and for questions regarding Topic B if they are posed by Person Y, but always tell a lie for questions regarding Topic A if they are posed by Person Y and for questions regarding Topic B if they are posed by Person X). In such a situation, combinations of persons and questions are the decisive factor that determines responding, and it could be that in such a situation, these combinations are stored in memory together with their respective response behaviors (cf. related research on the activation of stereotypes by combinations of categories and contexts, Casper, Wentura, & Rothermund, 2010, 2011; Müller & Rothermund, 2012; Wigboldus, Dijksterhuis, & van Knippenberg, 2003).

Specificity of the Binding Effect for Lies and Truths

Given the explanations for the discrepancy between our effects and those from Koranyi et al. (2015), the question may arise whether this kind of binding is specific to the truth status of an interaction. Actually, we believe that it is also possible to store other aspects of a certain situation (e.g., information about the mood in which one was acting toward the other person), as long as this information can help individuals to create future interactions in a more efficient and successful way. Still, this is a question that can only answer based on additional data.

Dependency of the Effect on the Content of the Questions

One difficulty that we are often dealing with is the creation of suitable questions for the interview situations that participants have to go through. As stated above, these questions are chosen in order to create a balance between asking people about relevant topics but at the same time trying to be not too intimate, to make sure that participants do not lie when they are supposed to tell the truth. Unfortunately, we do not yet have any systematic approach on this, which is why it remains unclear which influence the selection that we made has on the current results. To us, it seems plausible that talking to someone about an unimportant subject might not leave any traces whereas sharing private information with a former stranger does. Definitely, additional studies are required to address this topic more comprehensively.

Interpersonal Closeness

Another important question stemming from the current research regards the familiarity or interpersonal closeness of the person to whom one has responded either truthfully or dishonestly. Our study exclusively focused on strangers with whom one had had no personal relation before the experiment. In this situation, telling the truth was shown to be a non-default response that was stored and later retrieved from memory. It could very well be, however, that the opposite asymmetry would emerge for interactions with familiar persons. We would expect that honesty and intimacy are the default for interactions with close others (e.g., spouse/romantic partner, family members, good friends), so that answering truthfully to them would not constitute a remarkable event that would have to be stored in memory. Instead, lying to them would be a highly atypical and salient event that would probably leave strong person-related traces in episodic memory. However, in such a situation, it may not be the person alone that becomes associated with a "dishonesty tag," since this would collide with the mental representation of this person as an important and close relation, implying that the person would still be associated with honesty. Perhaps, this dilemma could be solved by coding the *combination* of this person with the specific question to which one had lied as an exception in memory that would be reactivated only in situations that contain both elements (i.e., person and question).

Furthermore, it would be interesting to investigate the consequences of having told the truth to a previously unfamiliar person on the relationship status that results with regard to this person. Previous research has accumulated some evidence indicating that disclosing personal information toward a stranger generates closeness (e.g., Aron, Melinat, Aron, Vallone, & Bator, 1997; Sprecher, Treger, Wondra, Hilaire, & Wallpe, 2013). Following these lines, one can assume that whenever a person discloses some intimate information to a stranger, a feeling of closeness might emerge toward this person, with implications for the evaluation of this person. Having told a lie to a stranger, however, would probably not produce any closeness and might even lead to negative evaluations. Interpersonal trust and openness might therefore be part of a relationship development process, in which the truth that one has spoken sets the stage for further intimate and truthful future interactions. Establishing a mental link between the

previously unfamiliar person and the knowledge that one has been honest to him or her might provide an important first step in the development of an intimate and trusting relationship.

Conclusion

This study demonstrated that having revealed personal information to a former stranger establishes an association between this person and the knowledge that one has been honest. Reencountering this person again leads to an automatic retrieval of this "honesty tag" that is important in regulating future interactions with this person and that might constitute an important first step in the development of an intimate relation. Future research is needed to more closely investigate the mediating processes underlying this effect and of the conditions that might moderate the specific pattern of episodic storage and retrieval of associating persons with truth and falsity-related information.

Electronic Supplementary Materials

The electronic supplementary material is available with the online version of the article at https://doi.org/10.1027/1618-3169/a000427

ESM 1. Text (.docx)
Supplementary questions.
ESM 2. Data (.txt)
Raw data of Experiment 1.
ESM 3. Data (.txt)
Raw data of Experiment 2.
ESM 4. Data (.txt)
Analysis of Experiment 1.
ESM 5. Data (.txt)
Analysis of Experiment 2.

References

Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and Social Psychology Bulletin, 23*, 363–377. https://doi.org/10.1177/0146167297234003

Barnes, C., Schaubroeck, J., Huth, M., & Ghumman, S. (2011). Lack of sleep and unethical conduct. *Organizational Behavior and Human Decision Processes, 115*, 169–180. https://doi.org/10.1016/j.obhdp.2011.01.009

Brandt, K. R., Gardiner, J. M., & Macrae, C. N. (2006). The distinctiveness effect in forenames: The role of subjective experiences and recognition memory. *British Journal of Psychology, 97*, 269–280. https://doi.org/10.1348/000712605X73685

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica, 51*, 5–13. https://doi.org/10.5334/pb-51-1-5

Casper, C., Rothermund, K., & Wentura, D. (2010). Automatic stereotype activation is context dependent. *Social Psychology, 41*, 131–136. https://doi.org/10.1027/1864-9335/a000019

Casper, C., Rothermund, K., & Wentura, D. (2011). The activation of specific facets of age stereotypes depends on individuating information. *Social Cognition, 29*, 393–414. https://doi.org/10.1521/soco.2011.29.4.393

Cohen, M. E., & Carr, W. J. (1975). Facial recognition and the von Restorff effect. *Bulletin of the Psychonomic Society, 6*, 383–384. https://doi.org/10.3758/BF03333209

Denkinger, B., & Koutstaal, W. (2009). Perceive-decide-act, perceive-decide-act: How abstract is repetition-related decision learning? *Journal of Experimental Psychology: Learning, Memory and Cognition, 35*, 742–756. https://doi.org/10.1037/a0015263

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology, 74*, 63–79. https://doi.org/10.1037/0022-3514.74.1.63

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*, 979–995. https://doi.org/10.1037/0022-3514.70.5.979

Dindia, K., Fitzpatrick, M. A., & Kenny, D. A. (1997). Self-disclosure in spouse and stranger Interaction: A social relations analysis. *Human Communication Research, 23*, 388–412. https://doi.org/10.1111/j.1468-2958.1997.tb00402.x

Eder, A. B., & Rothermund, K. (2008). When do motor behaviors (mis)match affective stimuli? An evaluative coding view of approach and avoidance reactions. *Journal of Experimental Psychology: General, 137*, 262–281. https://doi.org/10.1037/0096-3445.137.2.262

Ennis, E., Vrij, A., & Chance, C. (2008). Individual differences and lying in everyday life. *Journal of Social and Personal Relationships, 25*, 105–118. https://doi.org/10.1177/0265407507086808

Giesen, C., & Rothermund, K. (2016). Multi-level response coding in stimulus-response bindings: Irrelevant distractors retrieve both semantic and motor response codes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 1643–1656. https://doi.org/10.1037/xlm0000264

Hamilton, H. W., & Deese, J. (1971). Does linguistic marking have a psychological correlate? *Journal of Verbal Learning and Verbal Behavior, 10*, 707–714. https://doi.org/10.1016/S0022-5371(71)80079-8

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition, 5*, 183–216. https://doi.org/10.1080/713756773

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8*, 494–500. https://doi.org/10.1016/j.tics.2004.08.007

Horner, A. J., & Henson, R. N. (2009). Bindings between stimuli and multiple response codes dominate long-lag repetition priming in speeded classification tasks. *Journal of Experimental Psychology: Learning, Memory and Cognition, 35*, 757–779. https://doi.org/10.1037/a0015262

Horner, A. J., & Henson, R. N. (2011). Stimulus-response bindings code both abstract and specific representations of stimuli: Evidence from a classification priming design that reverses multiple levels of response representation. *Memory & Cognition, 39*, 1457–1471. https://doi.org/10.3758/s13421-011-0118-8

Koranyi, N., Schreckenbach, F., & Rothermund, K. (2015). The implicit cognition of lying: Knowledge about having lied to a question is retrieved automatically. *Social Cognition, 33*, 67–84. https://doi.org/10.1521/soco.2015.33.1.67

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492–527. https://doi.org/10.1037/0033-295X.95.4.492

Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analyzing data: A model comparison perspective*. Pacific Grove, CA: Wadsworth.

McCornack, S., Morrison, K., Paik, J. E., Wisner, A. M., & Zhu, X. (2014). Information manipulation theory 2: A propositional theory of deceptive discourse production. *Journal of Language and Social Psychology, 33*, 348–377. https://doi.org/10.1177/0261927X14534656

Metts, S. (1989). An exploratory investigation of deception in close relationships. *Journal of Social and Personal Relationships, 6*, 159–179. https://doi.org/10.1177/026540758900600202

Moutsopoulou, K., Yang, Q., Desantis, A., & Waszak, F. (2015). Stimulus–classification and stimulus–action associations: Effects of repetition learning and durability. *The Quarterly Journal of Experimental Psychology, 68*, 1744–1757. https://doi.org/10.1080/17470218.2014.984232

Müller, F., & Rothermund, K. (2012). Talking loudly but lazing at work – behavioral effects of stereotypes are context dependent. *European Journal of Social Psychology, 42*, 557–563. https://doi.org/10.1002/ejsp.1869

Neyer, F. J., Wrzus, C., Wagner, J., & Lang, F. R. (2011). Principles of relationship differentiation. *European Psychologist, 16*, 267–277. https://doi.org/10.1027/1016-9040/a000055

Rothermund, K., Wentura, D., & De Houwer, J. (2005). Retrieval of incidental stimulus-response associations as a source of negative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 482–495. https://doi.org/10.1037/0278-7393.31.3.482

Rubin, Z. (1975). Disclosing oneself to a stranger: Reciprocity and its limits. *Journal of Experimental Social Psychology, 11*, 233–260. https://doi.org/10.1016/S0022-1031(75)80025-4

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research, 36*, 2–25. https://doi.org/10.1111/j.1468-2958.2009.01366.x

Sprecher, S., Treger, S., Wondra, J. D., Hilaire, N., & Wallpe, K. (2013). Taking turns: Reciprocal self-disclosure promotes liking in initial interactions. *Journal of Experimental Social Psychology, 49*, 860–866. https://doi.org/10.1016/j.jesp.2013.03.017

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge, UK: Cambridge University Press.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology, 46*, 361–413. https://doi.org/10.1016/S0010-0285(02)00520-0

Waszak, F., Hommel, B., & Allport, A. (2005). Interaction of task readiness and automatic retrieval in task switching: Negative priming and competitor priming. *Memory & Cognition, 33*, 595–610. https://doi.org/10.3758/BF03195327

Wigboldus, D. H. J., Dijksterhuis, A., & van Knippenberg, A. (2003). When stereotypes get in the way: Stereotypes obstruct stereotype-inconsistent trait inferences. *Journal of Personality and Social Psychology, 84*, 470–484.

Wiswede, D., Koranyi, N., Müller, F., Langner, O., & Rothermund, K. (2013). Validating the truth of propositions: Behavioral and ERP indicators of truth evaluation processes. *Social Cognitive and Affective Neuroscience, 8*, 647–653. https://doi.org/10.1093/scan/nss042

Yam, K. C., Chen, X., & Reynolds, S. J. (2014). Ego depletion and its paradoxical effects on ethical decision making. *Organizational Behavior and Human Decision Processes, 124*, 204–214. https://doi.org/10.1016/j.obhdp.2014.03.008

History
Received September 24, 2017
Revision received July 6, 2018
Accepted July 23, 2018
Published online February 19, 2019

Open Data
Raw data and analysis of the Experiments are available in the Electronic Supplementary Materials, ESM 1–5.

ORCID
Franziska Annett Schreckenbach
https://ordic.org/0000-0001-7867-533X

Franziska Schreckenbach
Department of General Psychology II
Institute of Psychology
Friedrich Schiller University Jena
Am Steiger 3
07743 Jena
Germany
franziska.schreckenbach@uni-jena.de

4 Feature Specific Binding of the Knowledge of Having Lied Before: Evidence for Elemental Integration

Ask me no questions,
and I tell you no lies.
(Oliver Goldsmith)

The following article is submitted for publication to the *Quarterly Journal of Experimental Psychology.* (May 11th, 2021)

Schreckenbach, F. & Rothermund, K. (2021). *Feature specific binding of the knowledge of having lied before: Evidence for elemental integration.* MS submitted for publication.

The supplemental online material is presented in Appendix C.

Feature specific binding of the knowledge of having lied before: Evidence for elemental integration

Franziska Schreckenbach and Klaus Rothermund

Friedrich Schiller University Jena, Germany

Word Count: 5.784

Correspondence concerning this article should be addressed to Franziska Schreckenbach, Friedrich Schiller University Jena, Institute of Psychology - Department of General Psychology II, Am Steiger 3, D- 07743 Jena, Germany

(Email: franziska.schreckenbach@uni-jena.de; Telephone: 0049 (0) 3641 - 9 45124; Fax: 0049 (0) 3641 - 9 45122)

Abstract

Previous research on feature specific bindings has shown that by default, bindings are binary and elemental, that is, individual objects or single features of these objects can retrieve responses separately and independently. In our study, we applied these findings to the automatic retrieval of former deceptions. Specifically, we investigated whether the person or the question to which one has answered deceptively can retrieve this knowledge independently, or whether there is also evidence for configural bindings comprising a combination of person and question information to retrieve the truth status of the former episode. We found evidence for retrieval based on single cues (i.e., person or question), supporting that the binary binding principle also holds in the context of retrieving knowledge about former lies.

Keywords: lying, episodic binding and retrieval processes, instance-based learning

When we are in trouble, we sometimes search for a way out of the situation by telling a lie. If this works fine, we are successful in deceiving someone for the moment, but we still run the risk of being detected on a future occasion. Therefore, we should remember at least some details about our lie, for instance, what we were asked about, or to whom we have lied. Recent research provided first evidence for an automatic mechanism that helps us to remember the lies we have told in response to certain questions: When reencountering a question that one has lied to before, the knowledge about having lied is automatically retrieved from memory (Koranyi et al., 2015; Schreckenbach et al., 2020). Similarly, automatic memory retrieval has been found for the conversational partner that one has lied to on a former occasion (Schreckenbach et al., 2019). The aim of the present research was to find out whether these findings can be replicated in a more complex surrounding where different cues are simultaneously available for storage and retrieval.

Specifically, we investigated whether the combination of the question and the person to whom one had lied boosts retrieval effects over and above what question and person information can retrieve when they are presented in isolation. This question is relevant for both theoretical and practical purposes. On a theoretical level, these studies will inform us whether binding and retrieval processes in the context of lying are organized in a binary, elemental way, allowing independent retrieval for single features or single objects that were part of the respective episode, or whether binding and retrieval can also be organized in a more complex, configural fashion that requires combinations of features or objects for efficient episodic retrieval.

On a more practical level, our study will inform us whether retrieval of knowledge about having lied will generalize across different persons (for a certain question) or across different questions (for a certain person), as would be predicted by an elemental, binary perspective on retrieval, or whether and when retrieval of this knowledge is confined to the specific combination of question and person to whom one has lied.

Analyzing retrieval of knowledge about lies from an episodic retrieval perspective

The main assumptions of the present approach to memory storage and retrieval of lies are derived from theories of instance-based automatization of behavior (Logan, 1988; see also Denkinger & Koutstaal, 2009; Hommel, 1998, 2004; Rothermund et al., 2005). Within these theories, it is assumed that information about the execution of an action together with certain situational cues becomes stored as an episodic unit in memory (alternatively called an *event file*, Hommel, 1998, 2004, *instance*, Logan, 1988, or *stimulus-response episode*, Rothermund et al., 2005; see also Mayr & Buchner, 2006). These episodes are retrieved from memory and become automatically re-activated when one encounters a similar situation again, which then enables the person to act fast and consistently across different situations. Only recently, a broad framework has been provided that aims at explaining a vast range of research findings in the field of action control and automatization from an episodic binding and retrieval perspective (Frings et al., 2020).

Applying these approaches to the field of deception, Koranyi et al. (2015) developed a paradigm in which participants first took part in an oral interview and afterwards were tested for automatic memory retrieval of knowledge about having lied in a priming task. Participants received instructions to tell the truth to half of the questions asked during the interview but to lie to the other half. Afterwards, they had to perform a simple classification task where the probe words *honest* (in German: *ehrlich*) and *dishonest* (in German: *gelogen*) had to be identified by pressing a corresponding key. Immediately before each probe, a question of the interview was presented as a task-irrelevant prime to test whether identification of a probe word was facilitated by a corresponding prime question. In line with the predictions, the probe word *dishonest* was identified faster than the probe word *honest* when preceded by a question that had been answered dishonestly during the interview. This result was interpreted as automatic retrieval of the knowledge of having lied to a question when the same question is encountered again. Similar retrieval mechanisms were observed not only for questions but

also for the person to whom one has told a lie vs. the truth (Schreckenbach et al., 2019). In the corresponding study participants performed the same classification task after an interview while being primed with facial pictures of their former interrogators. Again, results revealed a congruency effect indicating that person cues automatically retrieve knowledge about the truth status of one's former behavior in former interactions with this person.

While the results of these studies can be taken as first evidence for an automatic memory retrieval of knowledge about former lies, the exact nature of the configuration of the presumed bindings is still unclear. Do people bind the truth status of their statements to specific situational cues or features, with lies being bound to questions and persons separately and independently, or do they form more complex episodes consisting of combinations of these cues (i.e., a specific person asking a certain question) with their knowledge of having lied to this person with regard to this question?

Drawing on the large literature on episodic binding and retrieval, there is overwhelming evidence that "bindings are binary", that is, the situational elements of an episode (objects or features of objects) are typically bound with the co-occurring response in binary S-R episodes, enabling each individual object or feature to retrieve the given response separately and on its own (e.g., Giesen & Rothermund, 2014, 2016). Evidence for more complex, configural binding and retrieval processes that require combinations of features or objects for retrieval is rare, and is confined to situations in which the elements of the initial episode are either not identified in an automatic and independent fashion (Moeller et al., 2016) or when they form contextual backgrounds (Mayr et al., 2018).

The present paper aims to transfer these findings into the field of deception, thereby providing a more applied perspective on these binding and memory retrieval mechanisms. More specifically, we want to investigate whether different features of the episode in which one has lied (i.e., information regarding the question and the person who asks the question) can retrieve the episode independently, reflecting elemental or binary bindings and retrieval,

or whether the combination adds something unique to the retrieval process that is not captured by the separate elements, reflecting configural binding and retrieval.

Our former studies suggest that episodic retrieval of lies also follows the principle of binary bindings and retrieval processes: Although participants always encountered combinations of relevant cues during the interview situation (e.g., a specific person asking specific questions), single cues in the subsequent priming task were sufficient to produce retrieval effects (Koranyi et al., 2015; Schreckenbach et al., 2019). The ability of single cues to retrieve episodic information supports the claim that each feature of an episode is separately bound to the response, and can retrieve the episode independently and on its own. Therefore, the assumption of a configural integration of these cues seems unnecessary, which is why we assume these bindings to be binary by default. However, we have not yet investigated this topic in a systematic fashion in previous studies. Most importantly, we have not compared retrieval effects for single and combined cues, which is why we cannot rule out the possibility that the matching combination of cues might still have a super-additive effect on retrieval, indicating the existence of configural binding/retrieval.

From an applied perspective, recalling one's former lies across different persons and different questions is typically required in order to remain fully consistent with one's lie across different future situations. Often, it is necessary to deceive not just one person with respect to a specific issue; to remain consistent one should tell this lie to a range of different people (e.g., when lying about an exam one failed, this needs to be done consistently to people who know each other [one's friends, parents, siblings] in order not to be detected). A similar generalization is also required when it comes to future interactions with the person whom one has told a specific lie: Remembering the lie one told before to a specific question will make it easier to respond consistently when interacting again with the same person even when these future interactions center on a different topic. Answers to other questions may have some thematic overlap with the question to which one lied before, which requires that

one adapts the answers to these questions to stay consistent with the original lie. An independent and elemental binding of both person and question information to the knowledge of having lied or told the truth guarantees such generalization effects, since each of the cues (person or question) on its own suffices to retrieve the knowledge about having lied before, which seems desirable from a functional perspective. Therefore, we predict elemental bindings when it comes to the knowledge of one's former statements, as this knowledge usually is of general relevance, either with regard to other people or other questions/topics.

A first pilot study that we conducted to address this question, however, yielded somewhat surprising results (see Supplementary Material for a full description of the methods and results of this experiment): We found the typical retrieval effects for question cues to which one had told the truth or lied when they were combined with a picture of the interviewer who had posed these questions during the previous interview. These effects were eliminated, however, when the questions were combined with a picture of an unknown person. The latter finding may indicate that combinations of the person and the question were stored together with the truth status of the response in a configural way, necessitating a joint presentation of the combination in order to retrieve the truth status of the former response. This conclusion, however, would be in conflict with our former findings where we found reliable retrieval effects for single cues. Alternatively, the findings might be explained if one assumes that combining questions with an unknown face can interrupt retrieval.

The current study

We conducted another experiment to investigate the question of elemental vs. configural binding in the context of truths and lies more systematically, and under more ecologically valid and meaningful conditions. Participants took part in oral interviews and afterwards performed a classification task in which they had to indicate via key press whether a presented probe was the word *honest* or the word *dishonest*. As primes, we presented a combination of the different cues that were used in former studies on automatic retrieval mechanisms of deception (i.e., questions, Koranyi et al.,

2015, Schreckenbach et al., 2020, and persons, Schreckenbach et al., 2019). With this manipulation, we also increased the complexity of this paradigm, thereby mirroring more lifelike conditions where multiple social interactions with different person take place in close temporal succession (one estimate being 12 interactions per day; see Zhaoyang et al., 2018), and similar or the same questions can be posed by different people. Combining question and person cues during the test also resembles everyday interactions where multiple cues (i.e., person and question cues) compete for the retrieval of previous episodes.

The experiment was designed to test whether different cues are used for memory storage and retrieval of the knowledge about having lied and whether these cues are bound and stored with this knowledge in combination (configural binding) or individually (elemental binding). We used a similar experimental design as in Schreckenbach et al. (2020) where participants met two different interrogators and had to lie to one of them while telling the truth to the second one. However, this time, participants were asked the same set of questions about the two different topics by each of the interrogators. Their task was then to always be honest to one of the interrogators while lying about one of the two topics to the other one. Thereby, we increased the complexity of the interview situation, while at the same time establishing a more realistic proportion of lies and truths with the truth being predominant and lies being told only rarely (cf. DePaulo & Kashy, 1998; DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996).

During the subsequent classification task, participants were presented with pictures of either one of their actual interrogators or an unknown person. On top of this picture, a question from the interview appeared, thereby leading to several different combinations of primes: the picture of a person whom participants always told the truth could be combined with (a) a question that they also always answered truthfully or (b) a question that they sometimes had answered deceptively. Similarly, the picture of the deceived person could be combined with (c) a question that was always answered truthfully or (d) the question to which

the person had lied. Finally, we also used unfamiliar faces as primes in combination with questions that were either answered (e) truthfully or (f) sometimes deceptively to see whether the recognition of the questions alone led to an observable retrieval effect. In the following, we will describe the two different hypothesized outcomes corresponding to either elemental or configural binding (see Figure 1 for a visual presentation of the predicted patterns of results):

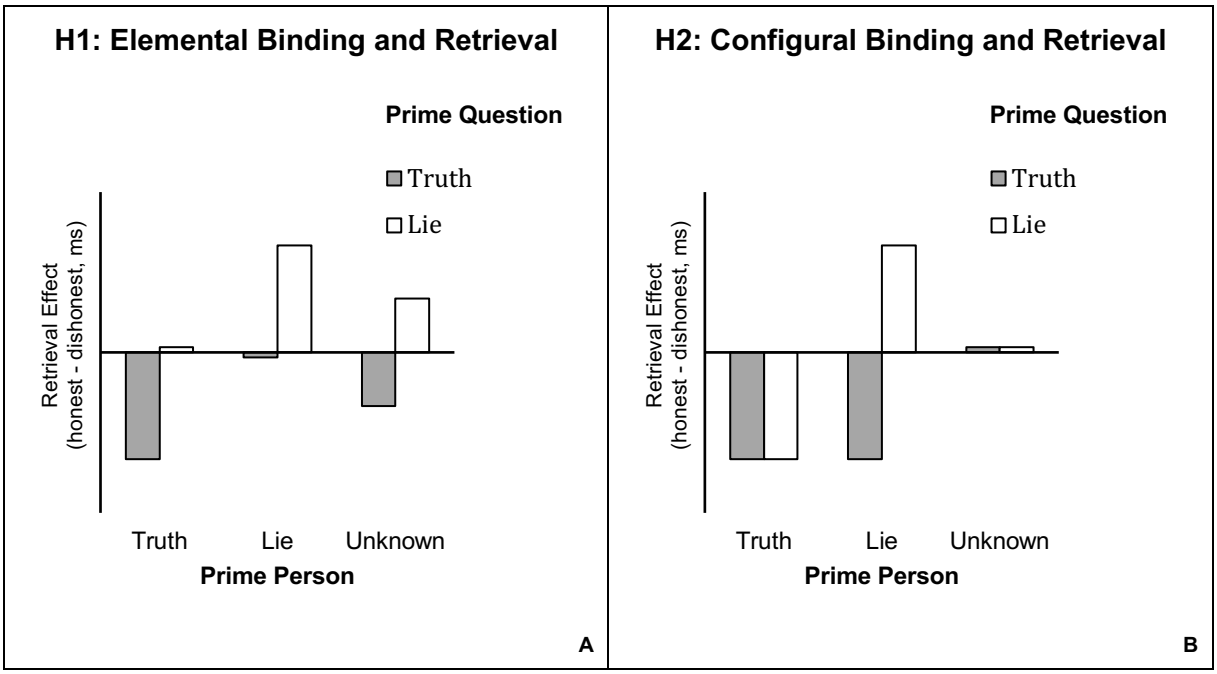

Figure 1. Predicted retrieval effects (positive values reflect faster identification of the probe word dishonest, thus indicating retrieval of the knowledge of having lied) for elemental (H1) and configural (H2) binding and retrieval as a function of prime picture (truth vs. lie vs. unknown) and prime question (truth vs. lie).

(H1) *Elemental integration:* According to the elemental account, both kinds of cues are integrated with the knowledge about the truth status of the response in separate binary bindings, and retrieve this knowledge independently of each other. Due to the higher salience and rarity of lies, having lied to a person or question only once suffices to connect this person or question with knowledge about having lied (Koranyi et al., 2015; Schreckenbach et al., 2020), whereas consistent truth-telling will connect the person or

question with a “truth”-mark. The elemental integration account thus predicts congruency effects of both person and question cues that sum up – or neutralize each other – in an additive fashion (Figure 1a). If person and question primes retrieve the same type of knowledge, strong congruency effects favoring responses to either true (probe word *honest*) or false (probe word *dishonest*) responses are thus expected if both person and question retrieve knowledge about having spoken the truth or having lied, respectively. If person and question primes retrieve opposite information, however, the resulting congruency effect should be close to zero. If the person prime is neutral (unknown person), the question prime should still retrieve information facilitating either true or false responses for questions to which one has told the truth or lied, respectively, but these effects should be weaker than those in which both person and question retrieve the same information. Statistically, this pattern corresponds to two independent main effects of person and question primes on the resulting congruency effects (i.e., the difference in responding to honest and dishonest primes).

(H2) *Configural integration:* According to a configural account, it is always the specific combination of person and question that is bound to the response and that retrieves the knowledge about the truth status of the answer that one gave in response to this specific combination (see Figure 1b). In this case, we expect a facilitation of the probe word *honest* whenever the picture of the person to whom one always told the truth is presented. The effect should not differ between the two types of question primes because all questions were answered truthfully in the presence of this person. That is, even when the person is paired with a question to which one has lied before, but to a different interviewer, the specific combination should retrieve information about the truth of the answer that was given to this specific combination of person and question. For the prime picture of the person to whom one has lied (to some questions), we expect congruency effects that depend on the type of the question that is presented as a prime.

A facilitation of the probe word *honest* (compared to the probe word *dishonest*) is expected after the presentation of the question to which participants told the truth, whereas a facilitation of the probe word *dishonest* (compared to the probe word *honest*) is expected after the presentation of the question that participants lied to during the interview with this specific interviewer. Finally, no congruency effects are expected to emerge whenever a prime picture of an unknown person is shown, regardless of whether the questions where always answered truthfully or where lied to during the interview. The reason for this absence of effects is that the specific combinations of person and question had not been encountered before and thus are unable to retrieve any information from memory. Statistically, this hypothesis corresponds to an interaction of person and question primes on the resulting congruency effects, with different types of questions pushing the congruency effect into opposite directions when combined with person primes to which one has (sometimes) lied, but having no effect on the resulting congruency effects when combined with persons who are either unknown or to whom one has always told the truth.

Based on the clear dominance of elemental bindings in the SR binding literature, and also based on our previous findings of reliable retrieval effects of lies for single cues, we favor the elemental binding account. Given that the findings of our pilot study can also be interpreted to suggest configural bindings, we do not want to rule out the possibility that configural bindings may emerge under certain conditions. In order to provide optimal conditions for the occurrence of configural binding and retrieval processes, we let the two experimenters ask the same set of questions in the current study. With these instructions, it is important not to confuse to which interviewer one has lied and told the truth to a particular question. Our experiment was thus designed to provide a strong test for configural bindings – if they exist, they should show up under these conditions.

Method

Participants and design. The study has been conducted in accordance with ethical standards and was approved by the Ethical Commission of the Faculty of Social and Behavioural Sciences of the University of Jena (FSV 19/18). The sample size was determined relying on previous experiments by Koranyi et al. (2015) and Schreckenbach et al. (2019), who found effect sizes between $\eta_p^2 = .05$ and .20 in similar paradigms. Due to the great differences, we decided to take the mean effect size of $\eta_p^2 = .13$ as an anchor and conducted a power analysis based on this value, $\alpha = .05$ and a power of $1 - \beta = .8$, leading to a proposed sample size of 57. To account for possible exclusions, we recruited 64 students from a German University. The data of one subject could not be meaningfully analyzed due to excessive error rates (> 30 %) in the priming task (average 4.3 % errors, without excluded participant). Furthermore, the data of one additional subject were excluded due to an error rate of > 20 % which led to missing values in several cells of the design. The final sample then consisted of 62 subjects (43 female) with an average age of $M = 24.1$ years ($SD = 3.64$). All participants gave informed consent via keypress at the beginning of the experiment. For their participation they received 4 € and a chocolate bar. We used a 3 x 2 x 2 factorial design with the within-subjects factors prime person (facial photo of the person one has met during the interview and told the truth vs. the person one has met and lied to about one topic vs. an unfamiliar person), prime question (question that one has always told the truth vs. sometimes lied to during the interview) and probe word (*honest* vs. *dishonest*).

Materials. The material for the interview comprised eight questions, four of which related to one of two different topics (university and friendship; see Appendix for a complete list). Each question was created to touch a personally important issue but at the same time should not be too intimate in order to prevent participants from answering untruthfully even if they were instructed to tell the truth. The same set of eight questions was asked by both interviewers during the interview, but the assignment of topic to response instruction was counterbalanced across participants, as was the order of topics during the interview.

As prime pictures we used facial photos of four interrogators on which they wore the same clothes, hairstyle, and accessories as in the actual interviews. All pictures had a size of 400 x 600 px, with each person's eyes placed at the same position in the upper half of the picture (approx. 200 px below the top). Participants were assigned to one of several couples of interrogators, with pictures of this couple serving as familiar and pictures of a different couple serving as unfamiliar prime picture stimuli during the priming task. To avoid confounds, only female interrogators were used.

Procedure. Upon arrival, participants were seated in front of a computer and received instructions on the screen. First, they were informed that they would participate in two successive oral interviews. Then, the picture of a female interrogator was presented on the left side of the screen, while on the right side the interviewer's name, the topics that participants would be asked about, and the instruction whether to always tell the truth or to lie to one given topic was presented (e.g. "This is Clara. Clara is going to ask you some questions about the topics 'friendship' and 'university'. Please tell her the truth always.") After reading and memorizing these instructions, participants received complementary instructions about the second female interrogator (e.g., "This is Sophie. Sophie is going to ask you some questions about the topics 'friendship' and 'university'. Please tell the truth to all questions about friendship but lie to all questions about university.") Additionally, participants were prompted not to reveal their dishonesty to the interrogator and to act so as to convince her of the truthfulness of all of their statements. In order to achieve this goal, participants were instructed to always wait for some seconds before providing each answer. The order of instructions as well as the assignment of topic or interrogator to response instruction was counterbalanced across participants, as was the order of topics during the subsequent interviews. After the instructions, participants were guided to a separate room where they met the first interrogator. After the first interview, the interrogator left the room and the second

interrogator entered to conduct the second interview. During the two interviews, the same set of eight questions was asked by the two interviewers.

Priming task. After the interviews, participants performed a priming task in which primes consisting of a combination of a person and a question were used as retrieval cues for the knowledge about one's former statements. Each trial had the same temporal structure (see Figure 2): A fixation cross (500 ms) was followed by a prime picture, showing either an interrogator or an unknown person. After 300 ms the prime question was presented upon the prime picture, approximately at the position where the eyes in the pictures were located (200 px below the top). The question was presented word by word using RSVP (rapid serial visual presentation) with a base duration of 250 ms per word, plus an additional 25 ms per letter. Right after the last word of the prime question, the probe word appeared at the same position of the screen. Participants were asked to decide as fast as possible whether a presented target was the word *ehrlich* (English: *honest*) or *gelogen* (English: *dishonest*, having lied) by pressing either the *D* or the *K* key on the keyboard. The assignment of response keys to probe word was counterbalanced across participants. To ensure that probe words were encoded and identified not only on a perceptual but also on a semantic level, they were degraded by inserting alphanumeric characters between the letters (e.g., *§$eh&$r§li#c%h* instead of *ehrlich*). The locations of these additional characters within the probe words were determined randomly for each trial, ensuring a large degree of variability between the probe stimuli. Both stimuli (i.e., the facial photo of the interviewer and the probe word) remained on the screen until the participant responded by pressing one of the assigned keys on the computer keyboard. The next trial was initiated after an inter-trial-interval of 750 ms. The priming task comprised 240 experimental trials with order of trials randomized individually. The prime pictures of the interrogators were presented 80 times each, while two different unfamiliar pictures were presented 40 times each. Familiar pictures 10 times preceded each of the 8 questions from the interview while unfamiliar prime pictures did so 5 times. Every

combination of primes was half of the times followed by the probe word *dishonest* and half of the times by the probe word *honest*. To ensure semantic processing of the prime questions, 24 experimental trials (randomly chosen out of the 240 trials) comprised an additional memory task that had to be performed directly after classifying the probe word (see Wiswede et al., 2013). In the memory task, participants saw a question on the screen which was either the same (50%) or different from the prime question and participants had to answer the question "Is this the question that you've just seen?" Participants were not informed in advance whether a trial comprised the additional memory task or not, so they had to process the prime questions in each trial.

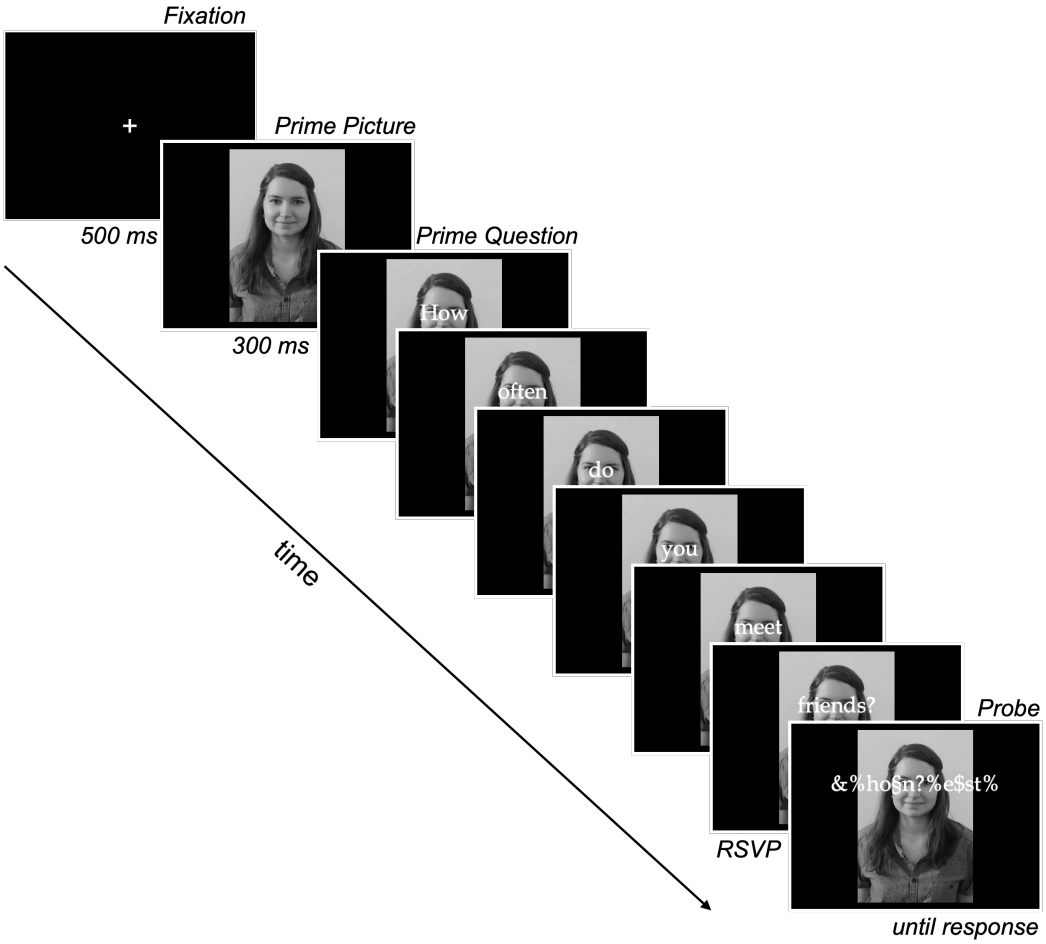

Figure 2. Trial Structure of the Experiment.

To ensure that the *honest* and *dishonest* keys maintained their semantic meaning across the experiment, 30 additional filler trials were randomly intermixed into the experimental trials that required a genuine true/false decision (see Eder & Rothermund, 2008; Wiswede et al., 2013). In the filler trials, a true (50%; e.g. "Saturn is a planet") or false assertion (50%, e.g., "Einstein was a musician") was presented word by word in the center of the screen instead of a prime picture. The assertion was followed by the question "honest or dishonest?" that was presented as a response cue instead of a probe word. Participants had to evaluate the truth of the previously presented sentence by pressing the same keys that were also used for the honest/dishonest classification of the experimental trials. The whole experiment had a total duration of approximately 45 minutes.

Results

All response latencies that were more than one and a half interquartile ranges above the third quartile of an individual's reaction time distribution were categorized as outliers (Tukey, 1977) and discarded (5.8% of all responses). All response latencies below the threshold of 250 ms were discarded (0.1%), as well as erroneous responses (4.0% of all responses). For each participant, we calculated the difference in response times for *honest* and *dishonest* probe words as an indicator of retrieval effects ($RT_{honest} - RT_{dishonest}$; positive values reflect faster identification of the probe word dishonest, thus indicating retrieval of the knowledge of having lied), separately for each combination prime person and prime question in the factorial design (see Figure 3 for the pattern of means).

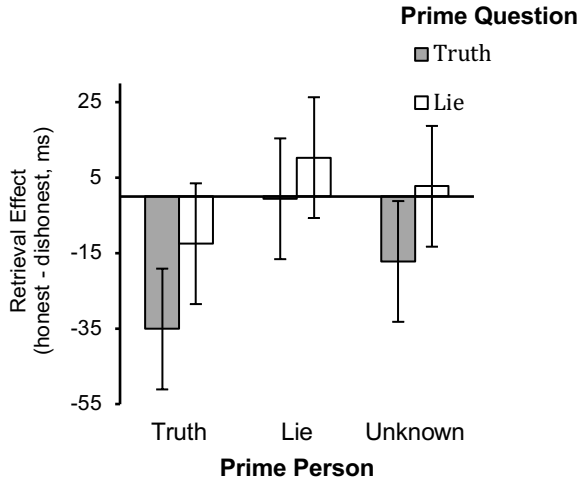

Figure 3. Average retrieval effects (positive values reflect faster identification of the probe word dishonest, thus indicating retrieval of the knowledge of having lied) as a function of prime picture (truth vs. lie vs. unknown) and prime question (truth vs. lie). Error bars represent 95 % CIs calculated for RM interaction effects as suggested in Jarmasz & Hollands (2009).

Analyses of variance. To test our assumptions, average retrieval effects were submitted to a 3 (prime picture: truth vs. lie vs. unknown) x 2 (prime question: truth vs. lie) ANOVA with repeated measures on both factors. We also specified two a priori contrasts for the factor prime picture to test whether retrieval of information occurred for both truthful and untruthful statements or mainly for one of them. The first contrast was specified to compare truth prime pictures with unknown prime pictures, whereas the second contrast compared lie prime pictures with unknown prime pictures.[1]

Results revealed significant main effects for both factors. The prime picture significantly influenced retrieval effects, $F(2, 60) = 6.52$, $p = .003$, $\eta_p^2 = .18$ (90% CI [.04, .30]). Splitting up this main effect into planned contrasts revealed a significant difference in retrieval effects for the truth vs. unknown contrast, $F(1, 61) = 6.58$, $p = .013$, $\eta_p^2 = .10$ (90% CI [.01, .22]). In line with former findings (Schreckenbach et al., 2019), retrieval effects

[1] The complete raw data as well as an analysis script for R and a table containing all cell means and standard deviations of the design can be found under: https://osf.io/g72hj/?view_only=06e4485f7faf4828a575a35c133b6253

were negative indicating facilitation for the probe word *honest* after the presentation of a person prime to which the person had always spoken the truth, $M_{retrieval}$ = -24 ms, $F(1, 61) = 13.03$, $p = .001$, $\eta_p^2 = .18$ (90% CI [.05, .31]), whereas no significant retrieval effect was obtained for unknown prime pictures ($M_{retrieval}$ = -7 ms, $F < 1$). The contrast between person primes to whom one had lied and unknown person prime pictures was not significant either, $F(1, 61) = 1.89$, $p = .18$, $\eta_p^2 = .03$ (90% CI [.00, .13]). Although retrieval effects for the person to whom one had lied indicated a facilitation of the probe word dishonest descriptively, this effect was not significantly different from zero ($F < 1$). We also found a significant effect for the type of the prime question, $F(1,61) = 4.91$, $p = .031$, $\eta_p^2 = .08$ (90% CI [.01, .19]), reflecting a significant retrieval effect indicating faster responding to the probe word *honest* for prime questions to which one had always responded truthfully, $t(61) = -2.63$, $p = .011$, $d = .33$ (95% CI [.08, .59]), whereas no difference in responding to honest and dishonest probes was found after prime questions to which one had lied before ($|t| < 1$). Importantly, we did not obtain an interaction of person and question primes ($F < 1$). Results thus support the notion of independent and additive retrieval effects for person and question primes, in line with an elemental integration of both kinds of cues, thereby supporting H1 more than H2.

Bayesian statistics. In order to compare probabilities for our hypotheses given the data, we also computed Bayes Factors, using the R package Bain (Gu, Mulder, & Hoijtink, 2018; Hoijtink, Mulder, van Lissa, & Gu, 2019). Implemented in Bain is the approximate adjusted fractional Bayes factor which can be used for the evaluation of informative hypotheses (Hoijtink, 2012). The Bayes factors and the posterior probabilities (computed assuming equal prior probabilities) are displayed in Table 1. Constraints for hypotheses 1 and 2 were defined in a way that they match the described data patterns in Figure 1 as well as our descriptions of the hypotheses. As can be seen, H_1 is supported more than H_2. A Bayes Factor

of 56.921 is usually seen as very strong evidence in favor of the H_1. The Bayesian error probability associated with preferring H_1 equals .064.

Table 1. *Bayes factors and posterior probabilities for Hypotheses 1 and 2, each in comparison to H_a which assumes unconstrained means in all conditions.*

Hypothesis	BF.a	posterior probability
H_1	56.921	.936
H_2	2.921	.048
H_a		.016

Discussion

The aim of our study was to test whether knowledge about having lied to a particular person about a specific question is integrated with these cues in a holistic fashion (thereby leading to configural binding and retrieval), or whether both kinds of cues (i.e., the person and the question) can trigger the knowledge about one's former statements individually (corresponding to elemental binding and retrieval processes). Both types of retrieval have been shown to take place in former studies on SR binding and retrieval, with elemental binding and retrieval processes being the default. In line with these findings, we observed a pattern of results that closely matches the assumption of an individual retrieval. Specifically, we found two independent retrieval effects for person primes and question primes supporting the conclusion that both types of cues can individually trigger the knowledge about having lied before. Another feature of our results that supports this view is the fact that question primes influenced retrieval even in the neutral condition, indicating that they can trigger retrieval of knowledge about having told the truth or a lie even when they were combined with unknown person primes with which they had never been shown before. Importantly, no interaction emerged between the two factors, indicating that the specific combination of person and question cues did not have an additional influence on retrieval effects, which speaks against a configural integration of the cues. Our current findings are in line with the SR binding literature which also shows a strong predominance of elemental over configural bindings (e.g., Moeller et al., 2016), and with our

previous studies on the retrieval of knowledge about having lied in which we found evidence for reliable retrieval of single cues (Koranyi et al., 2015; Schreckenbach et al., 2019, 2020). Although we designed our study in a way that should favor configural bindings for combinations of person and question cues, since the same question had to be answered differently during the interview depending on the interviewer who posed that question, we still did not find any evidence for configural bindings under these circumstances. Apparently, elemental binary processing is the default in binding stimuli and responses that also dominates in the processing of person and question cues with information about having lied or told the truth.

Against this background, the finding of our pilot study still appears to be somewhat enigmatic, since we did not find evidence for effects of question primes on the retrieval of truth status when these questions were combined with pictures of unknown persons. Given that these effects were non-significant also in the current study, but descriptively pointed in the expected direction in both studies, we think it is most likely that the lack of these effects in the pilot study reflects a power problem, rather than being indicative of configural bindings. We cannot rule out, however, that configural bindings in the integration of person and question information might emerge under very special circumstances, and that our pilot study was an instance of those conditions.

Conclusion. The present study extends the knowledge about automatic memory retrieval of the knowledge about having lied before by demonstrating that different cues (person and question information) to former lies are typically integrated with knowledge about having lied in a binary, elemental way, which replicates former findings from the SR binding literature (Giesen & Rothermund, 2014, 2016; Moeller et al., 2016) and also of our previous studies on automatic retrieval of knowledge about having lied (Koranyi et al., 2015; Schreckenbach et al., 2019, 2020). By default, these binary bindings guarantee that knowledge about having lied generalizes across different persons and questions. The presence

of just one cue that has been associated to the knowledge of having lied to the current person and/or the current question suffices to retrieve this knowledge from memory, and allows a flexible configuration of current response behavior in order not to be detected as a liar.

References

Brandt, K. R., Gardiner, J. M., & Macrae, C. N. (2006). The distinctiveness effect in forenames: The role of subjective experiences and recognition memory. *British Journal of Psychology, 97*, 269-280. https://doi.org/10.1348/000712605X73685

Cohen, M. E., & Carr, W. J. (1975). Facial recognition and the von Restorff effect. Bulletin of the Psychonomic Society, 6, 383–384. https://doi.org/10.3758/BF03333209

Denkinger, B., & Koutstaal, W. (2009). Perceive-decide-act, perceive-decide-act: How abstract is repetition-related decision learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 742-756. https://doi.org/10.1037/a0015263

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology, 74*, 63-79. https://doi.org/10.1037/0022-3514.74.1.63

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*, 979-995. https://doi.org/10.1037/0022-3514.70.5.979

Eder, A. B., & Rothermund, K. (2008). When do motor behaviors (mis)match affective stimuli? An evaluative coding view of approach and avoidance reactions. *Journal of Experimental Psychology: General, 137*, 262-281.

Frings, C., Hommel, B., Koch, I., Rothermund, K., Dignath, D., Giesen, C. G. …, Philipp, A. (2020). Binding and retrieval in action control (BRAC). *Trends in Cognitive Sciences, 24*(5). 375-387. https://doi.org/10.1016/j.tics.2020.02.004

Frings, C. & Rothermund, K. (2011). To be or not to be … included in an event file: Integration and retrieval of distractors in stimulus-response episodes is influenced by perceptual grouping. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(5), 1209-1227. https://doi.org/10.1037/a0023915

Frings, C. & Rothermund, K. (2017). How perception guides action: Figure-ground segmentation modulates integration of context features into s-r episodes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(11), 1720-1729. https://doi.org/10.1037/xlm0000403

Giesen, C. & Rothermund, K. (2014). Distractor repetitions retrieve previous responses and previous targets: Experimental dissociations of distractor-response and distractor-target bindings. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(3), 645-659. https://doi.org/10.1037/a0035278

Giesen, C. & Rothermund, K. (2016). Multi-level response coding in stimulus-response bindings: Irrelevant distractors retrieve both semantic and motor response codes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(10), 1643-1656. https://doi.org/10.1037/xlm0000264

Gu, X., Mulder, J., and Hoijtink, H. (2018). Approximated adjusted fractional Bayes factors: A general method for testing informative hypotheses. *British Journal of Mathematical and Statistical Psychology, 71*(2). https://doi.org/10.1111/bmsp.12110

Hoijtink, H. (2012). *Informative hypotheses. Theory and practice for behavioral and social scientists.* Boca Raton: Chapman and Hall/CRC.

Hoijtink, H., Mulder, J., van Lissa, C., and Gu, X. (2019). A tutorial on testing hypotheses using the Bayes factor. *Psychological Methods, 24*, 539-556. https://doi.org/10.1037/met0000201

Hommel, B. (1994). Spontaneous decay of response-code activation. *Psychological Research, 56*(X), 261-268. https://doi.org/10.1007/BF00419656

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition, 5*, 183-216. https://doi.org/10.1080/713756773

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8*, 494-500. https://doi.org/10.1016/j.tics.2004.08.007

Jarmasz, J. & Hollands, J. G. (2009). Confidence intervals in repeated-measures designs: The number of observations principle. *Canadian Journal of Experimental Psychology, 63*(2), 124-138. https://doi.org/10.1037/a0014164

Koranyi, N., Schreckenbach, F., & Rothermund, K. (2015). The implicit cognition of lying: Knowledge about having lied to a question is retrieved automatically. *Social Cognition, 33*, 67-84. https://doi.org/10.1521/soco.2015.33.1.67

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492-527. https://doi.org/10.1037/0033-295X.95.4.492

Logan, G. D. & Zbrodroff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a stroop-like task. *Memory & Cognition, 7*(3), 166-174. https://doi.org/10.3758/BF03197535

Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analyzing data: A model comparison perspective*. Pacific Grove, CA: Wadsworth.

Mayr, S., & Buchner, A. (2006). Evidence for episodic retrieval of inadequate prime responses in auditory negative priming. *Journal of Experimental Psychology: Human Perception and Performance, 32*(4), 932-943.

Mayr, S., Möller, M., & Buchner, A. (2018). Contextual modulation of prime response retrieval processes: Evidence from auditory negative priming. *Attention, Perception, & Psychophysics, 80*(8), 1918-1931. doi:10.3758/s13414-018-1574-z

Moeller, B., Frings, C., & Pfister, R. (2016). The structure of distractor-response bindings: Conditions for configural and elemental integration. *Journal of Experimental Psychology: Human Perception and Performance, 42*(4), 474-479. https://doi.org/10.1037/xhp0000158

Rothermund, K., Wentura, D., & De Houwer, J. (2005). Retrieval of incidental stimulus-response associations as a source of negative priming. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 31, 482-495. https://doi.org/10.1037/0278-7393.31.3.482

Schreckenbach, F., Rothermund, K., & Koranyi, N. (2019). And remember the truth that once was spoken: Knowledge of having disclosed private information to a stranger is retrieved automatically. *Experimental Psychology, 66*(1), 12-22. https://doi.org/10.1027/1618-3169/a000427

Schreckenbach, F., Rothermund, K., & Koranyi, N. (2020). Quantity matters: The frequency of deception influences automatic memory retrieval effects. *Quarterly Journal of Experimental Psychology, 73*(11), 1774-1783. https://doi.org/10.1177/1747021820924652

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addision-Wesley.

Wiswede, D., Koranyi, N., Müller, F., Langner, O., & Rothermund, K. (2013). Validating the truth of propositions: Behavioral and ERP indicators of truth evaluation processes. *Social Cognitive and Affective Neuroscience, 8*, 647-653. https://doi.org/10.1093/scan/nss042

Zhaoyang, R., Sliwinski, M. J., Lynn, M. M., & Smyth, J. M. (2018). Age differences in adults' daily social interactions: An ecological momentary assessment study. *Psychology and Aging, 33*(4), 607-618. https://doi.org/10.1037/pag0000242

Appendix

English Version of the Questions Used in the Experiment

The questions of one topic were always asked in the depicted order. The assignment of topic to truth condition was counterbalanced, as was the order of topics during the interview. Both interviewers asked the same questions in the same order.

Topic 1: University Life

1. Why did you choose your subject of study?
2. What do you like most about your subject of study?
3. Did you ever fail an examination?
4. What profession do you want to pursue after you have finished your studies?

Topic 2: Friendship

1. What role do friends play in your life?
2. What is the most important aspect in a friendship for you?
3. How often do you meet friends?
4. What do you typically do with your friends?

5 How to Remember Something You Didn't Say: Lies of Omission Can Be Stored and Retrieved From Memory

A good liar knows that the most efficient lie
is always a truth that has a key piece removed from it.
(Carlos Ruiz Zafon)

The following article has published as:

Schreckenbach, F., Sprengholz, P., Rothermund, K., & Koranyi, N. (2020). How to remember something you didn't say: Lies of omission can be stored and retrieved from memory. *Experimental Psychology, 67*(6), 364-372. https://doi.org/10.1027/1618-3169/a000504

The supplemental online material is presented in Appendix D.

Research Article

How to Remember Something You Didn't Say

Lies of Omission Can Be Stored and Retrieved from Memory

Franziska Schreckenbach, Philipp Sprengholz, Klaus Rothermund, and Nicolas Koranyi

Institute of Psychology - Department of General Psychology II, Friedrich Schiller University Jena, Germany

Abstract. When individuals suppress secret information, they should keep this omission in mind to not let this information slip out in future situations. Following recent findings about automatic memory retrieval of outright lies, we hypothesized that suppression tendencies are also automatically retrieved from memory when being confronted with a question to which one has previously omitted secret information. In an online study, participants first had to withhold information about a fictitious love affair during a simulated chat with their relationship partner. To assess automatic suppression tendencies, we developed an indirect response time measure wherein a key that had previously been established to indicate suppression now had to be pressed in response to word stimuli that were presented in a specific color. We found implicit suppression tendencies for words that had been withheld during the interview if they were presented following the prime that involved the question which the secret answer referred to. The question primes or the secret information alone did not elicit a suppression tendency, indicating that suppression responses were automatically retrieved from memory after re-encountering the combination of the question and the critical answer. The results are discussed regarding the theoretical implications for automatic memory processes.

Keywords: lies of omission, automatic processes, instance-based learning

Systematically omitting some details of an event while being honest about the rest has shown to be a widely used conversational strategy. In fact, in various studies, it was shown that participants who could choose between outright lies and more subtle omissions were more inclined to choose the latter (e.g., DeScioli et al., 2011; Levine et al., 2002; Pittarello et al., 2016; Rogers et al., 2017; Schweitzer & Croson, 1999). Omissions are not only cognitively less demanding compared to fabrications (McCornack et al., 2014), but they also have the advantage that people often feel less guilty for omitting relevant information than for committing an outright lie – although their victims do judge them as immoral, a phenomenon called *Omission Bias* or *Omission Effect* (Spranca et al., 1991, but see Willemsen & Reuter, 2016, for a more differentiated view). However, omissions can also become cognitively more accessible than intended by the person who uses them. Within the Preoccupation Model of Secrecy (Lane & Wegner, 1995), it is assumed that keeping a secret leads to an effortful circle of processes: The secret owner tries to suppress thoughts concerning the content of the secret, which, rather than being successful, induces intrusive thoughts of the secrecy. These intrusions lead to renewed efforts of thought suppression, thereby causing a status of cognitive "hyperaccessibility" of the suppressed content. Lane and Wegner (1995) conceived of this accessibility as being a disadvantage for secret keepers, but there might be a functional aspect to this process, as it can be important to remember one's omissions in the future to avoid being discovered.

The core focus of the current study was to further investigate automatic memory retrieval of omissions. For outright lies, we already identified a mechanism that helps liars to remember their lies in an automatic fashion in previous studies (Koranyi et al., 2015; Schreckenbach et al., 2020). Memory of previous lies is triggered by automatically retrieving episodic knowledge about having lied to a question when re-encountering the same question again in a later situation. On the one hand, contrary to outright lies, omissions consist in *not* telling something, and according to the omission bias, senders of omissions often feel less guilty than outright liars. Both of these characteristics may undermine memory retrieval of previous omissions due to a lack of a distinct behavioral memory trace that identifies the omission and/or due to a lack of marking omissions as something that is morally questionable that needs to be remembered on future

occasions. Moreover, true answers have been shown to be activated automatically when being confronted with specific cues (e.g., Duran et al., 2010; Hadar et al., 2012; Walczyk et al., 2003), which might make it even harder for people to withhold this information and to act in accordance with their previous omissions.

On the other hand, withholding relevant information is evaluated negatively by other people, making it important not to be found out after having held back important information. Additionally, based on the assumptions of the Preoccupation Model of Secrecy and the findings of an enhanced cognitive accessibility of secret thoughts (Lane & Wegner, 1995), it seems plausible to assume that under specific circumstances, omissions can be retrieved automatically from memory, which might make it easier for the secret keeper to behave consistently by repeating the previous omission. On a related note, it has been shown that binding processes can also occur between situational cues and the act of stopping or not executing an activated response (Giesen & Rothermund, 2014). This is why we expect that knowledge about omissions is also stored in memory and available for retrieval later. However, we assume that the functional properties of this "omission-retrieval mechanism" differ from the retrieval of outright lies. Specifically, we assume that in case of omissions, stimulus–response episodes are stored in memory that connect a specific question with the *suppression* of previously withheld information. Thus, re-encountering the question again on a later occasion will automatically elicit a tendency to suppress the previously omitted information again.

To measure this automatic retrieval of suppression tendencies for omissions, we developed a new experimental paradigm. For this purpose, participants first had to imagine being involved in a fictitious love-affair scenario in which they cheated on their partner. After reading this scenario, we simulated an interaction with their relationship partner in which participants had to omit all information associated with the affair while being honest about innocuous information. Participants then learned the connection between pressing a specific key and withholding information by establishing this key press response as a behavioral indicator of suppression. In a final part of the study, participants had to perform a Go/No-Go task where probe words were presented in different colors, one of which was a signal to execute the Go response, which corresponded to the "suppress" key of the previous part of the experiment. Before the presentation of these words, questions appeared on the screen as prime stimuli. This task is an indirect measure of suppression tendencies that are automatically elicited by the combination of questions and answers, which should facilitate executing the Go response (with the same key that was formerly established to indicate suppression).

This procedure allowed us to investigate the automatic activation of suppression tendencies for omitted information when being confronted with critical questions. Specifically, we predicted faster execution of Go responses with the previous "suppress" key in trials in which words reflecting secret information (i.e., information associated with the affair) were presented as probes after the matching question had been shown as a prime. Statistically, our prediction corresponds to an interaction between prime sentence (matching vs. nonmatching with the probe) and probe word (secret vs. innocuous content). We predict that Go responses should be faster for secret information compared to innocuous information if the matching question is presented as a prime. In terms of stimulus–response binding and retrieval accounts (e.g., Frings et al., 2020), re-encountering a question that was previously asked by the partner should lead to an automatic retrieval of episodes in which the secret information had been withheld from the partner in response to this very question. This activation should therefore facilitate a renewed suppression response (the formerly established key press response). On the contrary, no facilitation should occur during trials where the prime question does not match the content of the probe word, since no retrieval is triggered by the question before the probe is presented, thereby leading to equal response times (RTs) for secret and innocuous probes.

Method

Participants and Design

The study was conducted in accordance with ethical standards and was approved by the Ethical Commission of the Faculty of Social and Behavioural Sciences of the University of Jena (FSV 19/44). To make the fictional love affair in the experiment as realistic as possible, we only recruited participants who were in a permanent relationship at that time. A priori power calculations (G*Power 3; Faul et al., 2007) revealed that a sample size of $N = 37$ participants would suffice to detect a medium-sized effect ($f = 0.25$) with sufficient power ($1 - \beta = .8$), for a test that corresponds to our hypothesis ($\alpha = .05$, one-tailed), assuming a moderate degree of correlation among our dependent measures ($r = .3$). We succeeded in recruiting $N = 35$ participants during a period of two weeks. One of the participants had to be excluded because she gave inconsistent responses during the introductory part (she indicated that she was homosexual but then entered a male name as the name of her partner). Furthermore, three participants had to be excluded because they did not

behave as instructed during the second part of the experiment.[1] The final sample therefore consisted of 31 heterosexual participants (24 female) aged 1850 years (M = 26.04, SD = 7.94). Participants were recruited through social networks. They received an Amazon voucher worth €5 and were offered partial course credit in exchange for their participation.

We used a 2 × 2 factorial design with the within-subjects factors being criticality of probe (secret vs. innocuous) and matching of prime and probe (matching vs. nonmatching).

Procedure and Materials

Data were collected online using jsPsych (de Leeuw, 2015). Subjects could participate from any place but were asked to ensure a silent environment without distractions. They received the information that the study was conducted to investigate hiding behavior that people might show when they want to keep a secret affair from their partner. They were informed about the anonymization of data and that they could abort the experiment at any time. After this introduction, participants had to indicate their gender and sexual orientation as well as the names of their partner and a good friend who has a different sex from the partner. The good friend was established as a counterpart to the fictitious lover, which we used to establish the nonsecret details in the story about the weekend trip. Afterward, participants had to choose a potential secret lover from one of three images, and they assigned a name to this lover (henceforth, we will always refer to the person that participants had their fictional affair with as the lover, thereby differentiating them from their partner as well as from their friend). Male heterosexual participants should name their (female) partner and a real-life male friend and select a fictional female lover, while female heterosexual participants were instructed to name their (male) partner and a female real-life friend and to choose one of three fictional male lovers (see Electronic Supplementary Materials, ESM 1 for pictures).

Learning the Critical Episode

Afterward, participants read a story about cheating on their partner during a weekend. In this story, they were asked to imagine visiting their friend in one city for 2 days and how they spent the second day with their lover in another city. The story was made up in a way that three important facts must be remembered about what participants had done either with their friend or with their lover, respectively (e.g., to which city they went and what they ate for lunch). To make this task as easy as possible, we chose activities that are rather typical for the two cities (e.g., to eat fish in Hamburg and to eat currywurst in Berlin). Participants were then prompted to only tell their partners about the things they had done with their friend, but to withhold the activities they had done with their lover (see ESM 1 for the whole story):

> "Probably [*name of the partner]* will chat with you in a moment. Of course, he will ask about your weekend and how the trip was. You should not lie to him, but you better omit information about Berlin.
>
> Consequently, you should suppress that you were driving a Trabi, ate Currywurst and visited the Bundestag. You better tell him about Hamburg, that you travelled by Boat, ate Fish and visited the Reeperbahn.
>
> Memorize which information you can share with [*name of the partner*] and which must be omitted."

After reading the story, participants solved cloze tests (by filling in missing words in a text about their activities during the weekend) to check whether they remembered all important facts. If this was not the case, they had to reread the story. After answering the test questions correctly, participants received information about the upcoming chat with their partner in which they were asked to withhold the secret information:

> "[*Name of the partner*] is writing you! Answer her questions in the chat, preferably briefly, with only one word. Be honest, but withhold all information that [*name of the partner*] must not know."

Below the instructions, a text message application appeared in which (seemingly) the partner asked participants three questions about what they had done during the weekend (see Figure 1). In fact, prepared text modules were presented to start the conversation and to ask the questions, as well as some short responses which matched the answers that we expected from participants. Participants could reply freely via the keyboard, but the

[1] During the second part of the experiment, which served to establish an association between omitting critical information and pressing the space bar, these three participants showed high error rates (> 50%). Therefore, we have to assume that they did not establish the intended association, which is why we excluded them from all further analyses.

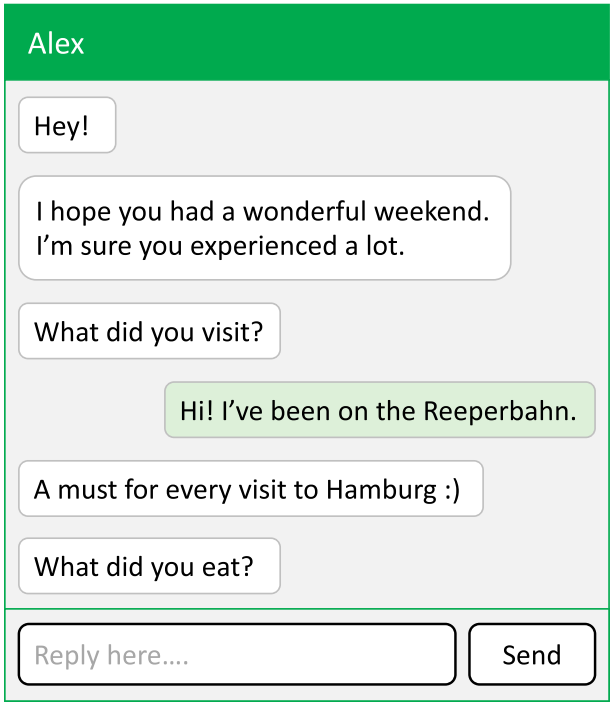

Figure 1. Display of the text message application for the chat with the participant's partner.

answer was not allowed to be sent if it did not include the innocuous answer or if it included secret information. For instance, if the partner asked "What did you eat?", participants' answers had to include "fish" but not "currywurst." During this chat, participants told three facts about their activities in Hamburg while they omitted three facts about what they did in Berlin.

Establishing a Suppression Response

During the following second part of the experiment, an association needed to be established between the omission of information and a specific, measurable behavior. For this purpose, participants engaged in another virtual conversation with their partner. In multiple trials, a silhouette of the participant's partner (female or male, respectively) was shown and a speech bubble with a question appeared (e.g., "Do you love me?"). Then, a potential answer (e.g., "No") was shown at the bottom of the screen that moved toward the silhouette. Participants were instructed to press the space bar as fast as possible if they did not want to utter the answer. In this case, the answer disappeared, and the next question was presented. If participants did not react, the answer reached the silhouette of the partner after three seconds. In total, seven questions were used, and each was shown twice in combination with different answers of which one was expected to trigger a key press (see ESM 1 for the whole list of questions and answers). All questions used in this part of the experiment were different from the questions used in the subsequent Go/No-Go task. However, some of the questions also addressed the topic of cheating on one's partner (e.g., "Do you cheat on me?"). This procedure was carefully chosen to achieve two goals: First, we needed a specific behavioral response for omissions that enabled us to measure suppression tendencies during the following Go/No-Go task. This was reached by establishing the space bar as a form of behavioral response to withhold information during this part of the experiment. Second, we wanted to avoid bindings between the key stroke and the secret information to be established in advance. This is why we chose *different* questions from those used during the chat as well as during the Go/No-Go task in the final part of the experiment.

Assessment of Suppression Tendencies for Critical Omissions

Finally, we asked participants to perform a Go/No-Go task to assess whether the specific questions from the text message conversation automatically trigger the retrieval of knowledge about having previously omitted information. Each trial had the same temporal structure (see Figure 2). A fixation cross (400 ms) was followed by the already known partner silhouette (400 ms). On top of the silhouette, the name of the partner appeared in combination with the verb "asks" (500 ms). Afterward, a prime question was presented word by word using rapid serial visual presentation with a base duration of 250 ms per word, plus an additional 25 ms per letter. This question was always one of the three questions the partner had asked during the chat. Right after the last word of the prime, a gray rectangle containing a single word appeared on the screen. After 400 ms, this probe changed its color or disappeared. If the probe turned red, participants had to press the space bar as fast as possible and their reaction time was measured (Go condition). If the probe turned green, no reaction was needed (No-Go condition). The trial ended as soon as participants pressed the space bar or after 1,000 ms had passed. The probe word was always related to information that had been uttered or omitted during the chat, thus being secret or innocuous. Since each probe word could appear after each prime, they were either matching (e.g., the prime was "What did you eat" and the probe was "currywurst") or nonmatching (e.g., the prime was "What did you visit?" and the probe was "currywurst"). The Go/No-Go task comprised 192 experimental trials, half of which were Go and half of which were No-Go trials. The order of trials was randomized individually. The three prime questions were presented 64 times each, half of the times preceding a matching probe and half of the times preceding a nonmatching probe. Furthermore, this probe comprised

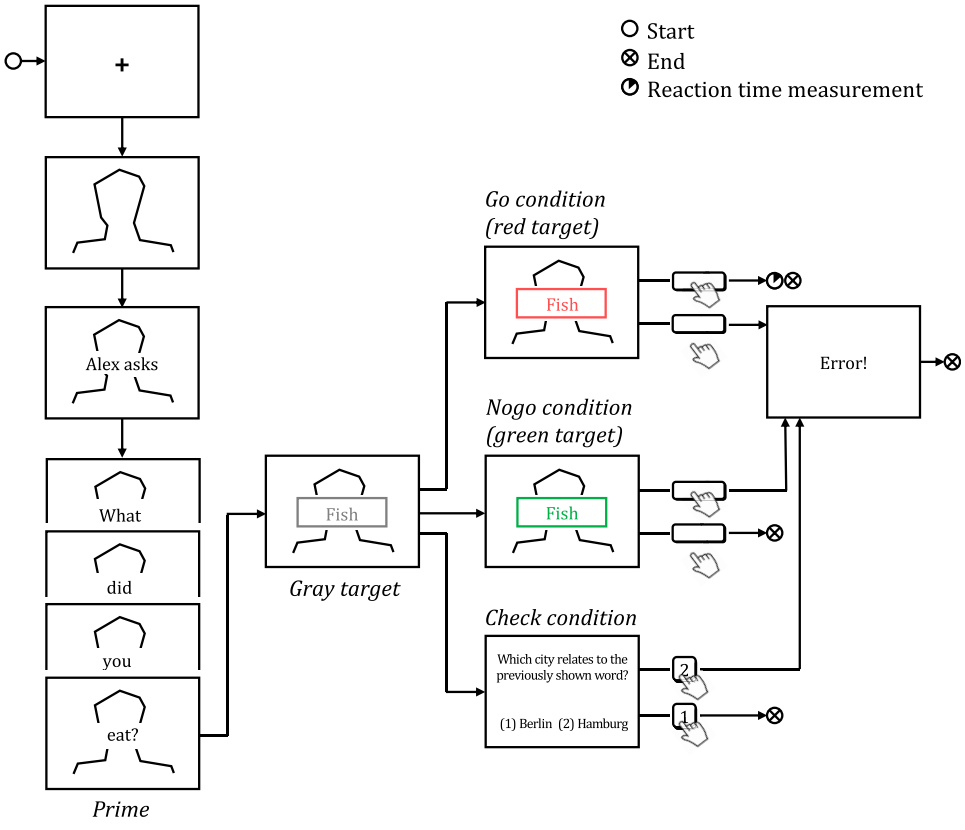

Figure 2. Trial structure of the Go/No-Go task.

secret information half of the times and half of the times innocuous information.

To ensure processing of the prime question, 10% of the trials comprised an additional attention check asking for the question that was presented as a prime stimulus in this trial. To also check participants' attention for probe words, an additional 12 trials were added where the probe did not change its color but disappeared and participants had to answer a question about the probe word (e.g., "Which city relates to the previously shown word: Berlin or Hamburg?"). At the end of the experiment, answers and reaction times were stored in an online database.

Results

RTs in the experimental task that were more than three interquartile ranges above the third quartile of an individual's RT distribution were categorized as far-out values (Tukey, 1977) and therefore discarded (2.4% of all RTs). All RTs below the threshold of 250 ms were discarded (0.2%), as well as erroneous responses (0.5% of all responses).

We calculated average RTs for each participant and combination of the factorial design (see Figure 3 for the pattern of means). To test our hypothesis that reacting to a probe should be accelerated if it relates to secret information and fits to the previously shown prime, average response latencies were submitted to a 2 (criticality of probe: secret vs. innocuous) × 2 (matching of prime and probe: matching vs. nonmatching) ANOVA with repeated measures on both factors. (The complete raw

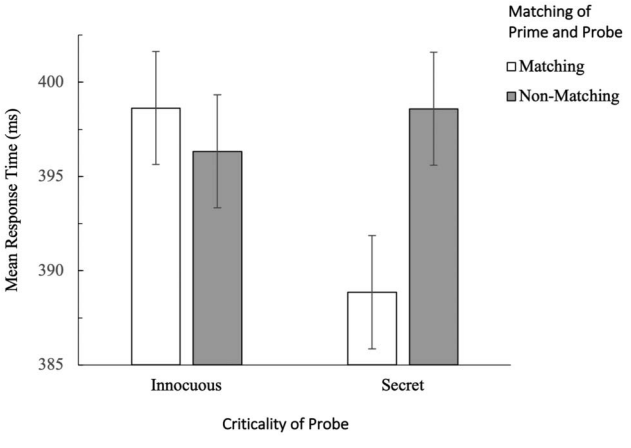

Figure 3. Mean RTs (error bars reflect SE of the mean) for executing the Go response as a function of the criticality of the probe word (secret vs. innocuous) and of the matching of prime question and probe word (matching vs. nonmatching). RT = response time.

data are available via https://osf.io/ky682/?view_only=846377030c2540cbae0c23a8c9b77e94.)

The results revealed no main effects (both Fs < 3.2) but a significant interaction of both factors, $F(1, 30) = 19.84$, $p < .001$, $\eta_p^2 = .40$. Follow-up tests showed that this effect was based on faster responses for secret probes (M = 389 ms, SD = 43) compared to innocuous probes (M = 399 ms, SD = 45) after the presentation of a prime question that matched the probe word, $t(30) = -4.03$, $p < .001$, $d_z = 0.72$. No such difference was found for nonmatching trials (M = 399 ms, SD = 49, for secret probes, and M = 396 ms, SD = 52, for innocuous probes, $|t| < 1$). Thus, in line with the hypothesis, participants were faster in executing a Go response (which was identical to the previously established suppress response) when confronted with secret information after having been primed with a question that matched the probe word.

Discussion

The aim of the present study was to test the hypothesis that critical information which was withheld in a previous communication is associated with a suppression response in memory and that this response tendency is retrieved automatically when a question is asked that directly refers to the omitted information. In line with this hypothesis, we found the predicted facilitation effect for Go responses for probe words that contained secret information (i.e., information that had to be suppressed during the first part of the experiment) after having been primed with a matching question. No such effect was found for probe words that did not match the previously presented question, leading to the conclusion that the tendency to withhold a specific piece of information is triggered in response to a certain question, and does not reflect a general tendency to suppress this specific information.

The results of the present experiment build on former studies about automatic retrieval of the knowledge of having lied (Koranyi et al., 2015; Schreckenbach et al., 2020). These results transfer the idea of an automatic memory retrieval of lies to the case of omissions, thereby extending our knowledge about implicit memory processes that relate to deception. The present findings suggest that lies of omission are stored in memory and can be automatically retrieved again, which makes them comparable to explicit lies. Importantly, and in contrast to episodes in which one has lied, omissions are coded as a tendency to suppress information. Memory for omissions apparently consists in a retrieval of the suppression of a specific type of information that had been withheld in response to a question, which is comparable to the previously mentioned binding of stop responses (Giesen & Rothermund, 2014).

However, our findings in the test phase reflect a retrieval of the knowledge that the critical information was withheld (i.e., not told) during the conversation with the partner (from part 1), and this retrieval now elicits a tendency to press the key that had been established to indicate suppression (in part 2). This indicates that knowledge of having suppressed this information before automatically elicits a general suppression tendency. No conceptual leap is taken here: The introduction of the key press as an indicator of suppression cannot change or influence the original omission episode, since the meaning of the key press was introduced only after the omission occurred.

It is important to emphasize that we established this suppression response (i.e., the key stroke) using questions that differed from the critical ones related to the secret visit in the cheating scenario. With this procedure, we ensured that only a connection between the suppression of *any* response and the corresponding key stroke could be learnt, but not the connection between the secret information (e.g., "currywurst") and the key stroke. This is an important feature of our experiment as it serves to rule out one possible alternative explanation: Participants' responses in the final Go/No-Go task cannot be attributed to the retrieval of a former episode in which the secret information was already connected to the key press. Instead, our findings reflect a retrieval of the knowledge that the critical information had been withheld (i.e., not told) during the conversation with the partner (from part 1), and this retrieval now elicits a tendency to press the key that had been established to indicate suppression (in part 2).

The retrieval effects we observed can also reflect the operation of some kind of implementation intention (e.g., Gollwitzer & Brandstätter, 1997; Gollwitzer & Schaal, 1998). Implementation intentions are self-regulatory "if-then" plans that create a strong link between a situation and an action. They have been shown to automatically direct the focus of attention toward goal-directed cues (Achtziger et al., 2012; Wieber & Sassenberg, 2006), thereby leading to a strategic automaticity during the initiation of goal-directed behavior and making the execution of the desired action effortless. The intention to withhold certain pieces of information from someone else reflects a specific type of implementation intention (e.g., "When my partner asks me whom I met last night, I will not tell her that I met my ex"). In these cases, the behavioral part of the implementation intention consists in not executing a specific behavior (cf. Chatzisarantis & Hagger, 2010; Gawrilow & Gollwitzer, 2008). If these intentions to suppress certain information in response to certain questions linger on until the end of the experiment, they can also influence performance in the final indirect test. This happens by eliciting corresponding suppression tendencies that will then be translated into corresponding

suppression behavior (i.e., pressing the space bar). To disentangle effects of implementation intentions from memory retrieval of previous acts of suppression, future studies are needed. In the critical condition of such a study, participants will be instructed to form an intention to suppress certain information (like in the present study) but then will be tested for an implicit activation of suppression tendencies without first having executed any episode of deception.

In the present study, we provided the first evidence that by omitting certain information during a conversation, an association is built between the omitting tendency and the associated question. Being confronted with that question again later leads to an automatic retrieval of this behavioral tendency to suppress information. However, there are also some limitations to the present study which need to be mentioned. First, it should be noted that the sample size in the present experiment was rather small (n = 31), which leads to a comparably low statistical power. Low power has been shown to increase the probability of committing a type II error (i.e., not rejecting the null-hypothesis although it is false), but also to lead to inflated effect sizes in the case of significant findings (Button et al., 2013; Ioannidis, 2008). Based on these considerations, the lack of power might be especially problematic when it comes to the interpretation of null effects in our design, as it was the case for nonmatching prime–probe combinations. However, because the most important finding of our study is a significant interaction effect, the small sample size does not undermine our main conclusions. Nevertheless, we cannot exclude that our effect size of $\eta_p^2 = .40$ constitutes an overestimation of the real effect. To obtain further information about the real size of the effect, future replication studies with sufficient power should be conducted.

Another potential limitation refers to the artificial setting of the present experiment. Our participants did not really cheat on their partner but only had to imagine a corresponding scenario. As a result, we do not know how well participants were able to envision this betrayal. If some participants had difficulties with this task, this, however, should rather have led to an underestimation of the effects of retrieving and re-executing omissions from memory rather than to producing these effects.

The scenario was also not representative for real-life scenarios regarding the short time span between the first omission during the simulated chat and the retrieval of this omission in the Go/No-Go task. While only few minutes passed between these tasks, longer time periods can be assumed to pass between two conversations in real life. Therefore, an important question refers to whether the effect remains stable across longer intervals between the conversation and the test phase. Another open question relates to the generalizability of the implicit memory toward other situations. According to the instance theory of automatization (Logan, 1988), and recent accounts of binding and retrieval (Frings et al., 2020), retrieval of a former bound episode should only happen when specific features of the episode are repeated. Therefore, we do not necessarily assume that the tendency to withhold secret information in future situations spreads toward situations which differ with respect to important characteristics. In the present study, it was shown that retrieval of the knowledge of having omitted a secret is bound to the critical question. However, it is possible that a similar but differently framed question requiring the same answer (or omission) does not benefit from this binding because the episode will not be retrieved and activated. Up to now, these issues remain unresolved but offer some very interesting options for follow-up investigations on this topic.

One further limitation is that the study was conducted online, which might raise questions regarding the reliability of the results. In fact, we had no control over the environment in which our participants worked on the experiment and therefore cannot rule out possible distractions induced, for example, by secondary tasks they were performing. If participants were not as focused on the task as we asked them to be, this might have had one of two possible consequences: either, the effect that we found is an underestimation of the real effect (which would have been obtained in a more controlled setting); or the effect reported here reflects a more realistic estimation than an estimation during a lab study, as it was found under conditions that are more typical for real life. Either way, the reliability and the generalizability of our findings need to be confirmed by future replications, as a single study can only be a first step into a new, promising direction.

Our findings also have interesting implications for research focusing on lie detection. Up to now, implicit paradigms of lie detection (e.g., Concealed Information Test/Guilty Knowledge Test, Lykken, 1959; autobiographical Implicit Association Test, IAT, Sartori et al., 2008) mostly rely on investigating positive signs of activation, arousal, or familiarity that are shown in response to critical information by guilty but not by innocent people (e.g., stronger physiological reactions during polygraph testing; Iacono & Ben-Shakhar, 2019). Nearly all of these measures are fairly easy to see through and are typically susceptible to faking attempts, for instance, by feigning or simulating positive responses also for noncritical information (e.g., National Research Council, 2003; Verschuere et al., 2009; for a review, see Verschuere & Meijer, 2014). Implicit measures for assessing automatic suppression tendencies have the advantage of being more indirect and thus harder to understand and control. In our view, these measures should be easy to apply, and they should be universally applicable: It is a defining feature of suspects who are guilty that they know some critical piece

of information (about the crime, about the victim, or about their own behavior) that they do not want to convey and for which they have formed an explicit suppression intention. Precisely, these suppression intentions can be identified indirectly and implicitly with the sort of paradigm we developed, which should help to distinguish guilty from innocent suspects. We expect such a measure to be mostly immune to faking due to its complexity and its indirect nature. Looking at the individual level of the measure established in the present study, one can observe that 77.4% of the participants descriptively showed the predicted pattern of results, while the remaining 22.6% showed a reverse pattern. Unfortunately, the present data are insufficient due to the lack of an innocent control group, making it impossible to conduct a more thorough discriminant analysis. However, given the chance level of 50% to guess right whether a suspect is guilty or not, it becomes obvious that the present measure is not yet sufficient for a reliable detection of guilty subjects during an interrogation. Still, we hope that with further research and a refined procedure, we will be able to reliably differentiate between guilty and innocent subjects.

In sum, our findings support the hypothesis that episodes in which a person has intentionally withheld an important piece of information in response to a question are stored in memory and are automatically retrieved by re-encountering the question again in a subsequent conversation. Retrieving this episode re-activates the tendency to suppress the same information again, which can be detected with indirect measures of suppression, as described in our study. The underlying mechanism is important for our understanding of the mechanisms of deception, and it also provides a promising approach for the implicit detection of guilty knowledge.

Electronic Supplementary Material

The electronic supplementary material is available with the online version of the article at https://doi.org/10.1027/1618-3169/a000504

ESM 1. Pictures of potential lovers, cover story, and list of questions and answers used during the decision task.

References

Achtziger, A., Bayer, U. C., & Gollwitzer, P. M. (2012). Committing to implementation intentions: Attention and memory effects for selected situational cues. *Motivation and Emotion*, *36*(3), 287–300. https://doi.org/10.1007/s11031-011-9261-6

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(6), 365–376. https://doi.org/10.1038/nrn3475

Chatzisarantis, N. L. D., & Hagger, M. S. (2010). Effects of implementation intentions linking suppression of alcohol consumption to socializing goals on alcohol-related decisions. *Journal of Applied Social Psychology*, *40*(7), 1618–1634. https://doi.org/10.1111/j.1559-1816.2010.00632.x|

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

DeScioli, P., Christner, J., & Kurzban, R. (2011). The omission strategy. *Psychological Science*, *22*(4), 442–446. https://doi.org/10.1177/0956797611400616

Duran, N. D., Dale, R., & McNamara, D. S. (2010). The action dynamics of overcoming the truth. *Psychonomic Bulletin & Review*, *17*(4), 486–491. https://doi.org/10.3758/PBR.17.4.486

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Frings, C., Hommel, B., Koch, I., Rothermund, K., Dignath, D., Giesen, C., Kiesel, A., Kunde W., Mayr S., Moeller B., Möller M., Pfister R., & Philipp A. (2020). Binding and retrieval in action control (BRAC). *Trends in Cognitive Sciences*, *24*(5), 375–387. https://doi.org/10.1016/j.tics.2020.02.004

Gawrilow, C., & Gollwitzer, P. M. (2008). Implementation intentions facilitate response inhibition in children with ADHD. *Cognitive Therapy and Research*, *32*(2), 261–280. https://doi.org/10.1007/s10608-007-9150-1

Giesen, C., & Rothermund, K. (2014). You better stop! Binding "stop" tags to irrelevant stimulus features. *Quarterly Journal of Experimental Psychology*, *67*(4), 809–832. doi:https://doi.org/10.1080/17470218.2013.834372

Gollwitzer, P. M., & Brandstätter, V. (1997). Implementation intentions and effective goal pursuit. *Journal of Personality and Social Psychology*, *73*(1), 186–199. https://doi.org/10.1037/0022-3514/73.1.186

Gollwitzer, P. M., & Schaal, B. (1998). Metacognition in action: The importance of implementation intentions. *Personality and social psychology review: An Official Journal of the Society for Personality and Social Psychology, Inc*, *2*(2), 124–136. https://doi.org/10.1207/s15327957pspr0202_5

Hadar, A. A., Makris, S., & Yarrow, K. (2012). The truth-telling motor cortex: Response competition in M1 discloses deceptive behaviour. *Biological Psychology*, *89*(2), 495–502. https://doi.org/10.1016/j.biopsycho.2011.12.019

Iacono, W. G., & Ben-Shakhar, G. (2019). Current status of forensic lie detection with the comparison question technique: An update of the 2003 National Academy of Sciences report on polygraph testing. *Law and Human Behavior*, *43*(1), 86–98. https://doi.org/10.1037/lhb0000307

Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology*, *19*(5), 640–648. https://doi.org/10.1097/EDE.0b013e31818131e7

Koranyi, N., Schreckenbach, F., & Rothermund, K. (2015). The implicit cognition of lying: Knowledge about having lied to a question is retrieved automatically. *Social Cognition*, *33*(1), 67–84. https://doi.org/10.1521/soco.2015.33.1.67

Lane, J. D., & Wegner, D. M. (1995). The cognitive consequences of secrecy. *Journal of Personality and Social Psychology*, *69*(2), 237–253. https://doi.org/10.1037/0022-3514.69.2.237

Levine, T. R., Lapinski, M. K., Banas, J., Wong, N. C. H., Hu, A. D. S., Endo, K., Baum, K. L., & Anders, L. N. (2002). Self-construal, self

and other benefit, and the generation of deceptive messages. *Journal of Intercultural Communication Research*, *31*(1), 29–47.
Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527. https://doi.org/10.1037/0033-295X.95.4.492
Lykken, D. T. (1959). The GSR in the detection of guilt. *Journal of Applied Psychology*, *43*(6), 385–388. https://doi.org/10.1037/h0046060
McCornack, S. A., Morrison, K., Paik, J. E., Wisner, A. M., & Zhu, X. (2014). Information manipulation theory 2: A propositional theory of deceptive discourse production. *Journal of Language and Social Psychology*, *33*(4), 348–377. https://doi.org/10.1177/0261927X14534656
National Research Council. (2003). The polygraph and lie detection. In *Committee to review the scientific evidence on the polygraph*. National Academies Press.
Pittarello, A., Rubaltelli, E., & Motro, D. (2016). Legitimate lies: The relationship between omission, commission, and cheating. *European Journal of Social Psychology*, *46*(4), 481–491. https://doi.org/10.1002/ejsp.2179
Rogers, T., Zeckhauser, R., Gino, F., Norton, M. I., & Schweitzer, M. E. (2017). Artful paltering: The risks and rewards of using truthful statements to mislead others. *Journal of Personality and Social Psychology*, *112*(3), 456–473. https://doi.org/10.1037/pspi0000081
Sartori, G., Agosta, S., Zogmaister, C., Ferrara, S. D., & Castiello, U. (2008). How to accurately detect autobiographical events. *Psychological Science*, *19*(8), 772–780. https://doi.org/10.1111/j.1467-9280.2008.02156.x
Schreckenbach, F., Rothermund, K., & Koranyi, N. (2020). Quantity matters: The frequency of deception influences automatic memory retrieval effects. *Quarterly Journal of Experimental Psychology*, *73*(11), 1774. https://doi.org/10.1177/1747021820924652
Schweitzer, M. E., & Croson, R. (1999). Curtailing deception: The impact of direct questions on lies and omissions. *International Journal of Conflict Management*, *10*(3), 225–248. https://doi.org/10.1108/eb022825
Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of Experimental Social Psychology*, *27*(1), 76–105. https://doi.org/10.1016/0022-1031(91)90011-T
Tukey, J. W. (1977). *Exploratory data analysis*. Addision-Wesley.
Verschuere, B., & Meijer, E. H. (2014). What's on your mind? *European Psychologist*, *19*(3), 162–171. https://doi.org/10.1027/1016-9040/a000194
Verschuere, B., Prati, V., & Houwer, J. D. (2009). Cheating the lie detector: Faking in the autobiographical implicit association test. *Psychological Science*, *20*(4), 410–413. https://doi.org/10.1111/j.1467-9280.2009.02308.x
Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology*, *17*(7), 755–774. https://doi.org/10.1002/acp.914
Wieber, F. & Sassenberg, K. (2006). I can't take my eyes off of it – Attention attraction effects of implementation intentions. *Social Cognition*, *24*(6), 723–752. https://doi.org/10.1521/soco.2006.24.6.723
Willemsen, P. & Reuter, K. (2016). Is there really an omission effect? *Philosophical Psychology*, *29*(8), 1142–1159. https://doi.org/10.1080/09515089.2016.1225194

History
Received July 27, 2020
Revision received December 31, 2020
Accepted January 21, 2021
Published online March 4, 2021

Publication Ethics
The study was conducted in accordance with ethical standards and was approved by the Ethical Commission of the Faculty of Social and Behavioural Sciences of the University of Jena (FSV 19/44).

Open Data
The complete raw data are available at https://osf.io/ky682/?view_only=846377030c2540cbae0c23a8c9b77e94. All materials for this study are available in ESM 1.

ORCID
Franziska Schreckenbach
https://orcid.org/0000-0001-7867-533X

Franziska Schreckenbach
Institute of Psychology - Department of General Psychology II
Friedrich Schiller University Jena
Am Steiger
07743 Jena
Germany
franziska.schreckenbach@uni-jena.de

6 General Discussion

Lies can be wonderful things.
And when a lie is told artfully,
if it's done with a degree of craftsmanship,
I can't help but admire the liar.
(Patrick DeWitt)

The present thesis had the aim to broaden the understanding of the role that automatic memory processes play in deceptive and truthful communication. More specifically, it was investigated whether binding and retrieval processes occur in deceptive and truthful situations and how they are influenced by the given context. To reach this aim, a binding model of deceptive and truthful discourse was introduced, from which research questions were deduced in order to further refine the model. Specifically, it was investigated (1) how the context of a lie influences binding and retrieval, (2) which situational features can cause binding and retrieval, and (3) whether the binding of an internal control process is also possible. In the preceding chapters, these factors were investigated empirically, with their results being summarized briefly in the following section.

Chapter 2 addressed the question whether the context in which a lie occurs has an influence on the formerly observed automatic memory retrieval of the knowledge of having lied (Koranyi et al., 2015). Results provided support for the assumption that two different characteristics of deceptive situations contribute to automatic binding and retrieval processes: first, the relative frequency of lies, which in everyday life is lower than that of truths (DePaulo & Kashy, 1998), causes them to be processed in a distinct way and therefore facilitates memory storage and retrieval. Second, the specific quality of lies, which might be reflected by the cognitive resources needed to make up a convincing story as well

as to manage one's emotions during the delivery of the message (e.g., Walczyk et al., 2014), also seems to determine whether a previously made statement becomes stored in and retrieved from memory. Based on these results, it can be concluded that the context in which a lie is made has a moderating influence on binding and retrieval of the knowledge of having lied before. From the data presented in Chapter 2, one could conclude that these binding processes cannot be easily transferred to truths, as no congruency effect was observed for truth-telling, even when it was made the distinct option. However, as one could see in the subsequent chapters, there are conditions under which not deceptive but truthful statements become part of automatic memory mechanisms.

In Chapters 3 and 4, it was investigated whether other situational features besides the question that one has lied to are able to retrieve the knowledge about former deceptive or truthful statements from memory. As the knowledge about whom one deceived may be important to protect a liar from being detected, it seemed reasonable to assume that the person one has talked to can also serve as a retrieval cue. However, former research has shown that there are big differences in the prevalences of lying to familiar vs. unfamiliar persons, with the latter being deceived even more often than being told the truth (DePaulo & Kashy, 1998). In Chapter 3, this reversal of communicative behavior was shown to also alter binding and retrieval processes in such a way that having disclosed private information toward a formerly unknown person is stored in and later on retrieved from memory, while having lied to a stranger is not. To further ensure that these findings did not emerge due to the specific experimental instructions that participants received in Experiment 1, a carry-over effect of these instructions was ruled out in Experiment 2. By changing the neutral baseline, we showed that congruency effects only occurred for the actual interviewers but not for neutral persons that the same interview rules would have applied to.

In order to integrate the findings about questions and persons as retrieval cues for the knowledge of having lied or told the truth before, Chapter 4 addressed the question how a combined use of these cues works. In accordance with most findings in the field of distractor-response binding, evidence was provided for elemental binding and retrieval (e.g., Colzato et al., 2006; Giesen & Rothermund, 2014b, 2016). In this experiment, it was shown that both cues, the person one talked to and the question one had to answer in

advance, were able to retrieve the knowledge of one's former veracity independently from each other. However, in a pilot study conducted before, we observed a different pattern of results, with the question cue retrieving the knowledge of having lied before only when at the same time the person one lied to was presented. It is possible that participants in this study perceived both cues as a holistic unit, which in turn only activated the knowledge of having lied before when all features belonging to this unit were presented, but we more strongly assume this finding to be related to a lack of power in this study. In sum, our findings match former empirical studies about bindings to mainly emerge in an elemental fashion. However, given the conclusions drawn by Moeller et al. (2016) who observed a more flexible adjustment of bindings to the characteristics of a situation, additional research is surely needed to comprehensively understand the conditions that must be met for an elemental vs. configural integration of different situational cues.

Finally, Chapter 5 investigated whether bindings can also occur between the question one deceived about, its true answer, and an internal control process instead of the knowledge of having lied before. For this purpose, participants were asked to deceive by using lies of omission in a love-affair scenario, and afterwards had to perform a Go/No-Go task while being primed with questions and answers from the previous interrogation. Due to the specific structure of the experiment, faster responses in Go trials indicated an automatic retrieval of former suppression responses. As predicted, we found a facilitation effect in trials consisting of a question and its matching secret answer. We thereby further expanded the knowledge about automatic memory processes in the context of deception by providing first evidence for bindings between cues to a former deceptive situation and the suppression of a response. When comparing this study with the experiments reported in Chapter 4, it becomes obvious that we again implemented a combination of situational features (in this case, the question and the true but secret information). This time, they were bound to a suppression response of the given information, but this response was only retrieved when the secret information was preceded by the matching question. Hence, these findings could be interpreted in terms of configural binding, which can be seen as very adaptive in such a context: Thinking of the word "Berlin" as being secret with regard to where someone has been during the last weekend, a retrieval of the associated suppression response is useful in combination with the question "Where have you been last weekend?", but not in

combination with the question "What's the capitol of Germany?". Hence, this study might also be interpreted in terms of providing support for the assumption of a flexible adjustment of binding and retrieval processes to a given context.

6.1 Theoretical Implications

The present thesis provides ample evidence for automatic memory retrieval in the context of deceptive and truthful communication. Thereby, we replicate the findings from Koranyi et al. (2015) and extend them in various directions: we found bindings for different situational cues, for response knowledge and control processes, and we showed the influence of some context factors on these bindings. On the basis of these findings, I want to refine the binding model of deceptive and truthful discourse introduced in Chapter 1 as well as discuss some implications for binding and retrieval accounts.

6.1.1 Implications for the Binding Model of Deceptive and Truthful Discourse

Having already introduced the binding model of deceptive and truthful discourse in Chapter 1, I now want to present a refined version of this model that integrates the findings of the formerly presented studies, thereby providing a more fine-grained view on episodic binding and retrieval during deception and truth-telling. In Figure 3, the formerly presented findings are summarized and adjusted to the structure of the binding model. One obvious change consists in the fact that the present version contains more components (as depicted by the several red and blue boxes), relating to different situational cues as well as response features already investigated in this thesis. This fragmentation does not only allow for a more detailed description of the components themselves, but also for the specification of bindings between single features of the model, like the depiction of elemental vs. configural binding and retrieval of situational features and the knowledge of one's former veracity or a suppression tendency, respectively. Up to now, we found an integration of different situational features with each other (connections between red boxes; see Chapter 4), with the response information (connections to blue boxes, which were included in all Chapters), and together with an internal control process (in this case, the intention to suppress a response; see Chapter 5). Furthermore, the nature of these bindings

was shown to be moderated by the specific context in which the deception or disclosure occurred, with the following conclusions being drawn:

(1) The frequency of deception influences whether binding and retrieval of the knowledge of having lied before takes place after all (Chapter 2).

(2) The familiarity with a person seems to influence whether knowledge of having lied or having told the truth becomes bound to this person (Chapters 3 and 4).

(3) Generally, different situational features are encoded in an elemental way, but there might be situations where the context moderates whether retrieval occurs already after reencountering individual features of a former situation or only in case of a complete repetition of these features (Chapters 4 and 5).

In its current state, the memory model of deception organizes the presented findings on this topic, while at the same time revealing open questions for future research. Therefore, the next section will serve to discuss the implications of the presented findings on binding and retrieval accounts and suggest future studies to further extend the acquired knowledge.

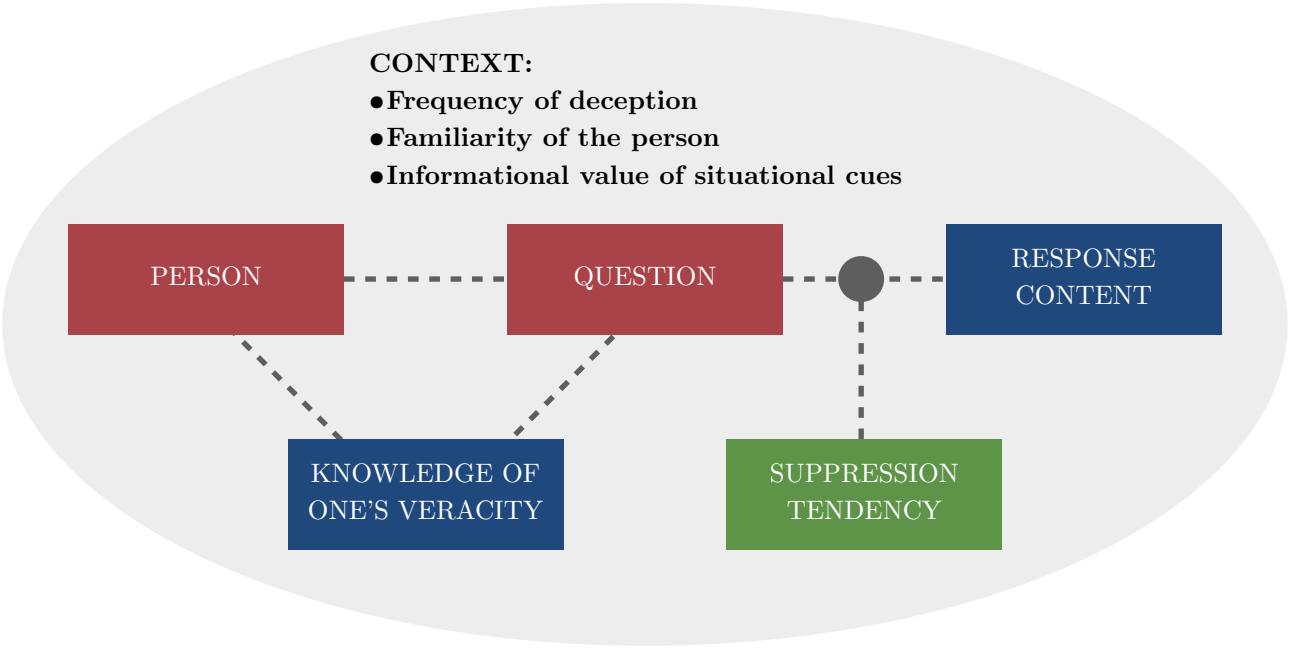

Figure 3: Integration of the findings presented in Chapters 2 to 5 into the binding model of deceptive and truthful discourse.

6.1.2 Implications for Binding and Retrieval Theories

Although the experiments presented within this thesis deviate in various aspects from typical studies reported in the literature on binding and retrieval, our findings provide further evidence to some research questions in this field. First, and maybe most obvious, we provide a new, applied perspective on binding and retrieval by investigating these processes in a quite different context than most of the studies in this field. It was already shown that stimulus-response bindings possess a great degree of complexity, with concrete object features being bound as well as abstract information, like semantic knowledge (Waszak & Hommel, 2007) or diagnostic decisions (Nett et al., 2015, 2016). However, in order to establish these bindings, most experiments (to my knowledge) only use pairings of stimuli on the PC or lists of words to learn, which depicts a strong simplification of realistic situations. Thus, although binding and retrieval theories often claim to be applicable to everyday life, the conditions where these assumptions are tested stay rather artificial. Here, we contribute substantially by showing that establishing a binding also works in more complex surroundings, e.g., in a face-to-face interview. Nevertheless, with regard to the retrieval part, we use the same artificial setting of a PC experiment as other researchers do. Still, as our participants experience at least some complexity during the encoding situation, this setting increases the ecological validity for binding and retrieval theories by going the next step toward showing that their assumptions also work in real life. Furthermore, we go beyond these findings by providing evidence for stimulus-response bindings to not only contain abstract response information that was concretely asked for during the priming situation (e.g., the decision for a concrete diagnosis, the content of a specific lie or truth), but also a semantic classification of this former response (in our case, the knowledge of having told the truth or having lied before). Interestingly, the concrete content of the semantic classification being bound varies for different situational features. While in Chapter 2, the information of having lied was bound to the question one answered before, the information of having told the truth was bound to the person one talked to in Chapter 3. These findings correspond to knowledge from pavlovian conditioning where the effectiveness of learning for conditioned stimuli depends on the unconditioned stimulus used (Garcia & Koelling, 1966). In the area of distractor-response binding, similar results have been reported, with the strength of bindings depending on

similarities between distractors and targets either on a perceptual (Laub & Frings, 2021) or on an affective level (Giesen & Rothermund, 2011). Hence, it seems that not every combination of stimulus and response can be equally easily bound together; instead, a certain fit is needed between different situational features (i.e., between a distractor and a target stimulus) or between situational features and response information (as in our studies) for a successful integration of these components into an event file.

In Chapters 4 and 5 we found first evidence for an elemental, but partially also a configural binding of different situational features, thereby leading to different retrieval mechanisms depending on the context. What we observed matches former findings about bindings being binary most of the times (Colzato et al., 2006; Giesen & Rothermund, 2014b, 2016). However, there seem to be some situations where an adaptive mechanism adjusts binding and retrieval to the demands of a given situation, thereby leading to configural bindings (Moeller et al., 2016). The findings presented in Chapter 5 can be interpreted in favor of this hypothesis by showing such a case of a configural binding. Hence, we deliver further evidence on the question how various situational features can be bound together, thereby expanding the knowledge about specific mechanisms having an impact on binding and retrieval processes. In Chapter 5, we also provide evidence for the binding and retrieval of internal control states, which matches findings from former studies on the binding of internal control states in general (e.g., Dignath et al., 2019; Egner, 2014) and, more particularly, on the binding of suppression tendencies (Giesen & Rothermund, 2014a).

It is important to notice that the presented findings provide evidence for a binding and retrieval that lasts very long in comparison to classic stimulus-response binding studies. As mentioned in the Introduction, one needs to differentiate between Logan's account of acquiring automatic behavior in a long-term way (e.g., Logan, 1988) and Hommel's account of short-term action control (Hommel, 1998, 2004). A lot of studies are concerned with the latter by studying bindings on a trial-to-trial basis, but the interest in the connections between binding and learning increased during the last years. The PEP 2.0 (as a model that combines findings from learning and binding) proposes a proportional activation of different episode nodes based on their frequency as well as their recency, thereby leading to the establishment of automatic behavior in a long-term way (Schmidt et al., 2016). The present findings support such accounts of long-term binding or a merge of

binding and (contingency) learning by showing pretty long-lasting bindings compared to classic binding and retrieval studies. While in a trial-to-trial test, bindings only need to last some seconds, our participants need several minutes to complete the interview and return to the computer before starting the priming task, and at least 10 more minutes until the last trials of the experiment. For the applied context that we are aiming at, this is not even a large enough amount of time, but we strongly exceed the time spans investigated in classic studies in the field of binding and retrieval.

When talking about differences, it is important to notice one more aspect of our experiments that deviates from classic binding and retrieval procedures. The structure of our experiments is always similar: participants first experience a specific episode where they are interrogated and have to behave either in a truthful or a deceptive way. After this episode, they perform between 150 and 250 trials in a priming task, with all pairings of primes and probes being equally probable. Thus, except for the interview episode at the beginning, no contingencies are established during the experiment. In contrast, there are two frequently used paradigms in the field of binding and retrieval: one is the trial-to-trial comparison of short-term bindings where no longer-lasting effects are investigated (see e.g., Frings et al., 2007; Giesen & Rothermund, 2011; Hommel, 1998; Moeller et al., 2016). The second paradigm consists of two different phases where the first one serves to establish bindings and the second one to test retrieval of these bindings (e.g., Horner & Henson, 2009, 2011). However, in the test phase, retrieval for each binding is demanded only once, thereby depicting a difference to our design, where bindings were tested about 20 times each. In our paradigm, we rely on the assumption that only the relevant interview situation leads to a binding, which can be retrieved but not modified during the priming part of the experiment. How does our paradigm then fit to binding and retrieval accounts at all? One possible explanation lies in the levels of processing account, published by Craik & Lockhart (1972). Within this account it is assumed that "the more deeply an item is processed, the better it will be remembered" (Baddeley, 1997, p. 47). Hence, the depth of processing might influence whether and how strong single episodes become stored in memory, with deeply encoded episodes being stored more steadily than superficially processed episodes. Transferred to our experiments, one can assume that the interview is processed more deeply than the single prime sequences, thereby leading to stronger storage and retrieval,

which cannot be simply overridden by word groupings on a screen. This deeper encoding might stem from, e.g., emotional involvement, like it is assumed in several theories on deception (e.g., Buller & Burgoon, 1996; Walczyk et al., 2014; Zuckerman et al., 1981), or from its distinctness (see Chapter 2). This idea is further supported by advancements of Craik & Lockhart's approach emphasizing that encoding many different features produces a richer and more discriminable memory trace, thereby leading to improved recognition (e.g., Cohen, 1981; Eysenck & Eysenck, 1980; Klein & Saltz, 1976). During the interview, subjects participate in a supposedly vivid situation containing numerous bits of experiences (a somehow furnished room, the face of the interviewer, the sound of her voice, the topic of the questions, ...). In contrast, the following presentation of single words on a screen does not provide such a manifold sensory experience. According to the levels of processing account and its advancements, the richer interview situation led to a more durable memory trace than the comparatively poor priming situation on the screen.

6.1.3 Implications for Cognitive Theories of Deception

When talking about cognitive theories of deception, it is important to emphasize that all implications drawn within the next section are only an approximation to the assumptions of these theories. The reason for this is simple: while, e.g., the ADCAT (Walczyk et al., 2014) investigates the real production process of lies, the studies presented here are only concerned with, at best, preparing processes for repeating a former lie. However, there are some conclusions that can be transferred to these theories and raise interesting questions for further examination.

Memory processes are at the core of the latest cognitive theories about deception, which heavily rely on working as well as long-term memory in order to explain lie construction and other parts of the deceptive and truthful communication process (McCornack et al., 2014; Walczyk et al., 2014). In the ADCAT, it is assumed that creating a lie (but also telling a truth) happens by accessing contents from long-term memory and handling them in working memory. Specifically, the authors state that in a first step, whenever a question is asked, the truth becomes activated automatically, an assumption that can be easily related to binding and retrieval accounts. However, based on the findings presented in this thesis, this hypothesis has to be questioned, at least in the case of repeated lies. In our

experiments, we did not find a special activation of former honest situations, but instead one of events that deviate from the default and therefore are perceived distinctly. Whether this event is a situation where one told the truth (see Chapter 3) or a situation where one lied (see Chapter 2, but also Koranyi et al., 2015) does not determine by itself which aspects become stored and later on activated automatically. One could argue that we only investigated abstract knowledge of having lied before, and not the activation of concrete deceptive or truthful contents. However, it seems implausible to assume that after the activation of the knowledge about a former deception the matching truth gets associated instead of the deception from the respective situation. Thus, further studies should try to resolve whether in such a situation truthful or deceptive content is activated automatically, a line of research that would help to improve both, the ADCAT as well as the binding model presented in this thesis. In order to predict which specific experience becomes activated, one could test whether the previously introduced law of recency (Giesen et al., 2020) or the unitary view on contingency learning (Schmidt et al., 2020) is able to predict which episode becomes activated after being confronted with a question demanding a serious answer. One further part of the ADCAT that our findings may cast doubt upon concerns the assumption of a decision about whether to lie or to tell the truth occurring before the construction of a lie can take place. While this might be true for first time lies, it probably does not apply to repeated lies. Based on our findings, an automatic activation of the knowledge of having lied before happens as soon as a question associated with this lie is reencountered, thereby being notably faster than the quasi-rational decision process that Walczyk et al. assume to take place before the construction of a lie. In contrast, it seems plausible to assume that with, or shortly after this knowledge the specific content of the lie gets activated too. The success of this retrieval might in turn even influence the decision whether to lie again or to tell the truth this time. It would thus be an appealing task for future research to investigate these assumptions, thereby contributing to the refinement of the ADCAT.

In contrast to the ADCAT that emphasizes the cognitive difficulty of deception, the IMT2 bases its reflections on the assumption that lies are cognitively easy. It therefore states that lies and truths are similar, as both rely on the same memory systems to arise. The presented findings about the truth being bound and retrieved under certain conditions (see

Chapter 3) can be seen as a hint toward the accuracy of this assumption (however, with our studies we only show that memory storage and retrieval processes seem to work similarly for lies and truths, while no firm conclusion about the similarity of production and execution processes can be drawn from these). Furthermore, the IMT2 suggests that lies will be only used when they are seen as the most efficient solution. To account for this claim, memory processes (e.g., the creation of a lie based on long-term memory experiences) need to happen before volitional processes (i.e., the decision to lie). It is fur sure not entirely correct to equate the memory processes of the IMT2 with the activation component in the ADCAT while equating volition of the IMT2 with decision of the ADCAT, but it is conducive to make visible how both models reverse the order of similar processes. Corresponding to my former doubts about the correctness of the assumptions in the ADCAT in this regard, I agree with the propositions of the IMT2 because they match the episodic retrieval processes demonstrated in the previous chapters. In the case of lies of omission, we even observed an automatic suppression tendency, which as a behavioral activation might strongly influence the decision of whether to repeat a lie in a subsequent situation or to act conversely by telling the truth. In the IMT2, this behavioral bias is reflected in the assumption that a lie, being delivered successfully once, will often be chosen again in subsequent situations. Based on our data, this assumption seems very reasonable, but up to now, I am not aware of any studies addressing this question. Therefore, and in order to demonstrate the relevance of our findings for ecological valid situations, it would be interesting to empirically investigate whether successful lies really are repeated with a higher probability.

6.2 Limitations and Directions for Future Research

In its current state, the binding model of deceptive and truthful discourse organizes existing findings on this topic, while at the same time revealing open questions for future research. In fact, there are some bindings that might also exist but have not yet been investigated. In Figure 4, these supposed bindings are depicted as yellow lines. As can be seen, they relate to a variety of connections between stimuli, response information, and control processes. Most importantly, there exists no study yet on the question whether the

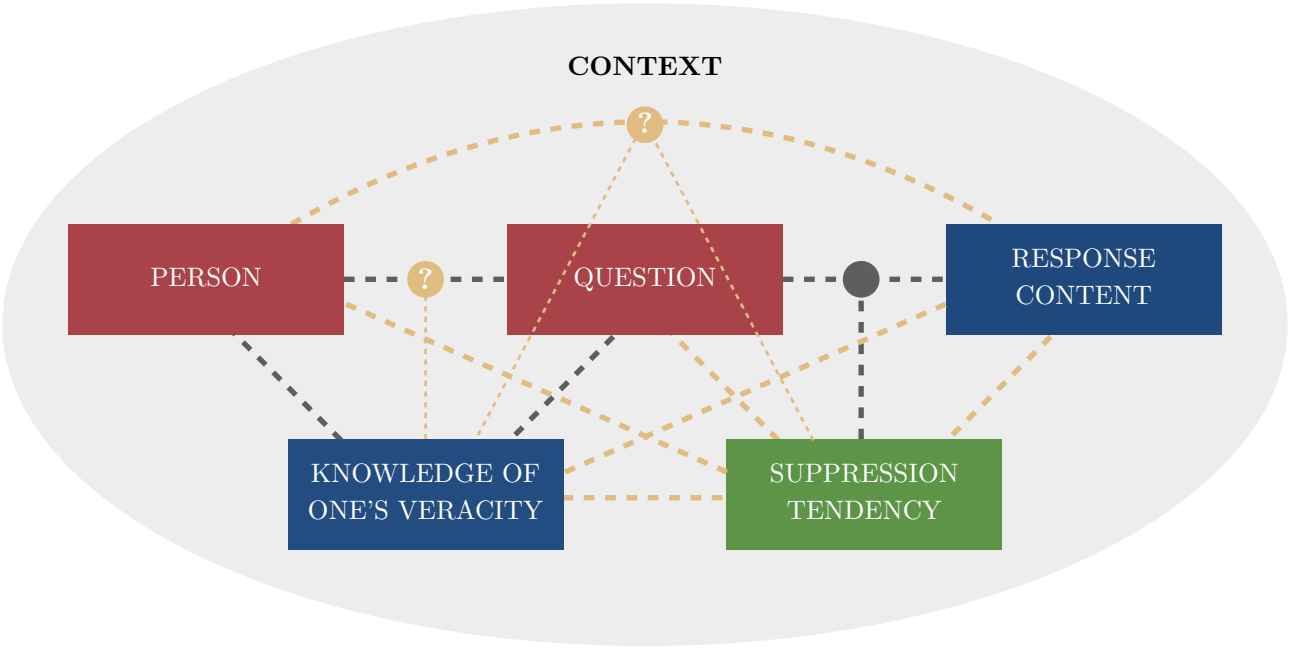

Figure 4: Illustration of open questions within the more detailed binding model of deceptive and truthful discourse, following from the findings presented in Chapters 2 to 5.

information of having lied or told the truth before and the content of the deceptive or truthful answer are bound together. However, I strongly assume this to be the case, based on the literature about binding and retrieval. In the framework of BRAC, it is stated that "stimuli, responses, and action effects are coded in a common representational format that allows us to treat features of S, R, and E interchangeably: event-file retrieval can be triggered by any type of feature" (Frings et al., 2020; p. 376). As we already found bindings between situational features and abstract response information as well as bindings between situational features and internal control processes, it seems reasonable to assume that all of these features are linked with each other, thereby enabling retrieval to also rely on the not yet tested combinations. Furthermore, as we already observed simultaneous bindings for a question, a person, and the knowledge of having lied (see Chapter 4), I assume the same for different combinations of stimuli although not having investigated them yet. Finally, the associative structure of these bindings (i.e., either elemental or configural binding) is assumed to partially depend on the context (cf. results of Chapters 4 and 5), thereby leading to the assumption that this context dependency might also apply to further combinations of stimuli and response codes not tested yet.

Besides the investigation of further bindings, there is also room for improvements with regard to the specification of the single components in the model. Therefore, the next section will serve to look at every component and consider unresolved issues. Starting with the situational features possibly being bound (red boxes in Figure 4), it was already shown that the person one talked to and the question one was asked can become stored in memory as part of the episode. Based on this, it is conceivable that more features relevant for a successful repetition of a lie can become part of this representation. However, every situation also contains irrelevant features (i.e., distractors) that might become bound and retrieved although not having any adaptive value. An example for such a distractor could be the color of the shirt the deceived person is wearing, or the taste of the tea one is drinking during the conversation. Whether distractors can also become part of an episode of deceptive communication was not tested yet, but given the similarities to binding and retrieval accounts, they can be assumed to emerge depending on the same mechanisms that were defined for distractor-response bindings elsewhere (e.g., context factors like perceptual grouping or gestalt principles, Frings & Rothermund, 2011, 2017; or affective matching, Giesen & Rothermund, 2011). Furthermore, the boundaries between relevant and irrelevant features in an interactive situation might be less clear-cut than in the artificial setting of classic computer tasks, which in turn could increase the probability of binding distractors together with relevant information. It would thus constitute an interesting line for future research to further investigate the boundary conditions for such mechanisms in the context of deception.

With regard to the response information, and more specifically, the knowledge of one's veracity in a former situation (see left blue box in Figure 4), we seemingly encountered some discrepancies between the conclusions drawn in Chapters 2 and 3. In Chapter 2, lies were declared to be qualitatively special, thereby containing some unique features that prevent knowledge of having told the truth from being stored and retrieved from memory even in a context where it is made the distinct option. However, in Chapter 3, we observed exactly this effect, with truths apparently being bound and retrieved together with the person that one talked to before. Moreover, we did not even put a special emphasis on truth-telling compared to lying in this study. As already discussed this discrepancy might stem from the difference in the retrieval stimuli bound to this knowledge: in one study, we

tested knowledge of having lied or told the truth toward a question, while in the other study, we tested the same knowledge toward a person. Hence, it seems like the different stimuli tend to be bound to different kinds of response information. Moreover, we also discussed that drawing the attention to different features of a situation (here, a question that is usually answered truthfully, and there, a stranger that is normally not told the bare truth), might have set the default processing mode for the whole situation. This assumption could be further tested in a study using a setup in which participants are assigned to one of two groups and prompted to lie to one of two interrogators either based on the questions asked (comparable to Koranyi et al., 2015) or based on the person asking (as it was done in Chapter 3). In the subsequent classification task, the presented primes could then also be varied orthogonally between participants by using either questions or persons. With such a manipulation, a systematic investigation of the impact of attention drawn toward a specific cue during the encoding phase as well as during retrieval would become possible. Finally, the discussion on what determines which kind of knowledge is stored and retrieved also contains the assumption that truth-knowledge in Chapter 3 resulted from the fact that participants had to disclose the truth toward an unfamiliar person. Although this explanation relies on empirical findings about the frequency of lying in general (which is comparatively low, DePaulo et al., 1996) and the frequency of lying to strangers (which is even higher than that of truth-telling, DePaulo & Kashy, 1998), it has not yet been tested for its validity. Therefore, future studies could contribute immensely to the formulation of the presented model by identifying precisely which factors determine whether the knowledge of having lied or having told the truth in a specific situation becomes stored in and retrieved from memory.

Looking at the concrete response content (the right blue box in Figure 4), we provided first evidence for bindings to occur between a question, its true but secret answer, and a control process to suppress this response (see Chapter 5). It is important to notice that although the truthful answer was entailed in the experiment described in Chapter 5, it was not tested for an automatic activation itself. Instead, we presented it as a potential retrieval cue for the suppression response. With this manipulation, we demonstrated that the secret information was also part of the deceptive episode stored in memory, but neither do we know whether it would also retrieve automatic activation (although BRAC suggests so),

nor is it clear whether the storage and retrieval of the truth is specific for omissions or whether it applies to fabrications too. Indeed, a situation where one decides to tell a lie differs in important aspects from an omission: in addition to the truth, an alternative story exists now, which is actually told in the situation and therefore could be well assumed to be stored as part of the deceptive episode. For the further development of this model, it would be crucial to investigate whether a formerly created fabrication becomes part of this episode, as its content is what ensures the successful repetition of a given lie at all. Assuming that this is the case, the question remains what happens to the truthful but secret information. Given that the preoccupation model of secrecy suggests that omissions are a special case increasing the accessibility of the truth (Lane & Wegner, 1995), one might assume that it is not stored as port of this episode as soon as a lie is formulated. However, relying on the claim of the ADCAT about the truth to become activated always as soon as one encounters a critical question (Walczyk et al., 2014), it could be also assumed that it becomes part of every deceptive episode as well. Obviously, there are a lot of unresolved issues with regard to the response component in the current version of the model, which should be subject to future research.

Internal control processes (the green box in Figure 4), which were shown to be bound in Chapter 5, can be compared to situational characteristics, and thereby possibly consist of several features too. We have not yet investigated these thoroughly, but based on the literature suggesting different internal control states, like task-sets, attention states or conflict detection processes (Egner, 2014), there is a whole range of experimental manipulations conceivable.

Let us now turn to the last part of the model, which is the context having an impact on binding and retrieval processes. It is probably the most difficult component because it can come in so many different shapes that it is almost impossible to investigate them comprehensively. This concern is reflected by the empirical findings presented in this thesis, which strongly suggest an influence of various external factors on the processing of deceptive and truthful communication but are not well suited to define the exact conditions of these factors and their respective effects on binding and retrieval (with an exception being the influence of frequency described in Chapter 2). In Chapter 3 it was observed that for the picture of a person being used as a prime, a congruency effect

emerges especially in the case of having told the truth toward this person. As discussed above, we assume this to be caused by the fact that in the experiments reported, interviewers were always unfamiliar to participants. However, to gain direct proof for this assumption, an empirical follow-up study would be important. Besides, one can observe an impact of context on the associative structure of bindings between different features when comparing Chapters 4 and 5. While in Chapter 4, congruency effects were observed based on individual retrieval cues, in Chapter 5, a configural use of retrieval cues is suggested as only the combination of prime question and secret answer led to a retrieval of the formerly established suppression response. For future research, these findings raise the question what exactly determines whether a specific retrieval cue is bound at all and which associative structure these bindings develop. One likely candidate might be the concept of intentional weighting, which comprises the idea that "codes of features on a dimension that is (assumed to be) relevant for the presently relevant task will have a stronger impact on representing an event then codes of features related to currently irrelevant dimensions" (Hommel, 2019, p. 2141; see also Memelink & Hommel, 2013). Likewise, it was shown that instructing participants with a specific task set helps them to ignore irrelevant information in a given situation (Dreisbach & Haider, 2008). Applied to the context of deception, and more specifically, to the experiments presented in this thesis, intentional weighting may have determined which combinations of features became encoded at all during the interview. In the instructions, we told participants which situational features were going to be important in the following interview, thereby providing them with information possibly leading to an intentional weighting of relevant features. As a consequence, they may have focused on these dimensions while ignoring other features that would have been included otherwise. Therefore, it is possible that in Chapter 3 no information about questions was bound at all because it was not necessary, while in Chapters 2, 4, and 5, this was a critical feature to focus on in order to successfully complete the interview, thereby leading to an integration in episodic memory structures. However, to gain more knowledge about these assumptions, further studies on the topic are needed.

Finally, one last limitation regarding the overall conceptualization of the binding model of deceptive and truthful discourse has to be mentioned. Up to this point, it was taken for granted that episodic binding and retrieval was investigated in all studies presented in this

thesis, and accordingly, all results were combined in a model of binding and retrieval. However, this claim is debatable at least in one case. In the discussion of Chapter 5, it was already mentioned that the observed effects might alternatively occur due to formed implementation intentions. Implementation intentions are action plans that are often construed in the form of if-then connections (e.g., *if* my partner asks me about what I did during the weekend, *then* I won't mention my visit in Berlin). By automatically directing the focus of attention towards goal-directed cues (Achtziger et al., 2012; Wieber & Sassenberg, 2006), implementation intentions lead to a strategic automaticity during the initiation of goal-directed behavior and enable an effortless execution of the desired action. In Chapter 5, this might have been the case because participants in the first part of the experiment received the instruction to suppress certain information, and later formed an association between a key press and the act of suppressing information. Hence, in the final Go/No-Go task they may have acted due to the intention built during the instruction to suppress secret information whenever it was prompted. However, while this might apply to the findings in Chapter 5, we explicitly ruled out this kind of alternative explanation for the binding between a person and the knowledge of one's truthfulness in Chapter 3. Furthermore, in Koranyi et al. (2015) the same test was conducted for the binding between a question and the knowledge of having lied before, showing that no overlap from the instructions could have accounted for the congruency effects in the priming task. Therefore, it is plausible to assume that in Chapters 2 and 4, episodic retrieval played a dominant rule too, although the direct test of this hypothesis was not repeated. However, the experimental paradigm in Chapter 5 differed from those used before, which is why we cannot conclude safely that episodic retrieval was observed here too. But, as the conceptual structure of the study was similar to earlier ones, I still assume that episodic retrieval at least contributes to the observed congruency effect, even if it was not the only factor. To resolve this issue, further studies are needed that transfer the control conditions used in Chapter 3 as well as in Koranyi et al's study to the new paradigm.

6.3 Conclusion

In the present thesis, a new binding model of deceptive and truthful discourse was introduced, proposing episodic binding and retrieval to contribute to an automatic regulation of deceptive and truthful conversations. In general, situational features were assumed to be bound with abstract response information, i.e., the knowledge of having lied or told the truth in a former situation, as well as with specific responses or internal control states. In the empirical chapters, some of the assumed bindings were tested and the impact of contextual factors was investigated. Specifically, it was shown that the frequency of deception influences automatic memory retrieval of the knowledge of having lied before in a way that it disappears when lying becomes the default mode in a given situation (Chapter 2). Furthermore, it was demonstrated that not only questions but also persons can become bound to the truth status of a former statement (Chapters 3 and 4) and that both of these situational features can be bound and retrieved simultaneously in an elemental way (Chapter 4). Finally, we found first hints for binding and retrieval to also take place in the case of omissions, and especially for the behavioral tendency to suppress a response (Chapter 5). The presented findings enabled an improvement of the binding model of deceptive and truthful discourse by refining its single components. Furthermore, it was possible to identify open questions requiring further investigation in the future based on this model. Besides, the increase of knowledge due to this model contributes to the advancement of binding and retrieval accounts in an applied area, thereby demonstrating the relevance of these theories for everyday behavior.

References

Achtziger, A., Bayer, U. C., & Gollwitzer, P. M. (2012). Committing to implementation intentions: Attention and memory effects for selected situational cues. *Motivation and Emotion, 36*(3), 287–300. https://doi.org/10.1007/s11031-011-9261-6

Baddely, A. (1997). *Human Memory: Theory and Practice, Revised Edition.* Exeter, UK: Psychology Press.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*, 203–242. https://doi.org/10.1111/j.1468-2885.1996.tb00127.x

Cohen, R. L. (1981). On the generality of some memory laws. *Scandinavian Journal of Psychology, 22*, 267–282. https://doi.org/10.1111/j.1467-9450.1981.tb00402.x

Colzato, L. S., Raffone, A. & Hommel, B. (2006). What do we learn from binding features? Evidence for multilevel feature integration. *Journal of Experimental Psychology: Human Perception and Performance, 32*(3), 705 – 716. https://doi.org/10.1037/0096-1523.32.3.705

Craik, F. I. M. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684. https://doi.org/10.1016/S0022-5371(72)80001-X

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science*, *23*(4), 517–542. https://doi.org/10.1016/S0364-0213(99)00014-2

Denkinger, B., & Koutstaal, W. (2009). Perceive-decide-act, perceive-decide-act: How abstract is repetition-related decision learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 742-756. https://doi.org/10.1037/a0015263

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Harris, C. (2003). Cues to deception. *Psychological Bulletin, 129*, 74–112. https://doi.org/10.1037/0033-2909.129.1.74

DePaulo, B. M. & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of Personality and Social Psychology, 74*(1), 63-79. https://doi.org/10.1037//0022-3514.74.1.63

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. Journal of Personality and Social Psychology, 70, 979–995. https://doi.org/ 10.1037/0022-3514.70.5.979

DeScioli, P., Christner, J., & Kurzban, R. (2011). The Omission Strategy. *Psychological Science, 22*(4), 442–446. https://doi.org/10.1177/0956797611400616

Dignath, D., Johannsen, L., Hommel, B., & Kiesel, A. (2019). Reconciling cognitive-control and episodic-retrieval accounts of sequential conflict modulation: Binding of control-states into event-files. *Journal of Experimental Psychology: Human Perception & Performance, 45*(9), 1265-1270. https://doi.org/10.1037/xhp0000673.

Dreisbach, G. & Haider, H. (2008). That's what task sets are for: Shielding against irrelevant information. *Psychological Research, 72*, 355-361. https://doi.org/10.1007/s00426-007-0131-5

Egner, T. (2014). Creatures of habit (and control): A multi-level learning perspective on the modulation of congruency effects. *Frontiers in Psychology, 5,* Article 1247. https://doi.org/10.3389/fpsyg.2014.01247

Eysenck, M. W. & Eysenck, M. C. (1980). Effects of processing depth, distinctiveness, and word frequency on retention. *British Journal of Psychology, 71*, 263-274. https://doi.org/10.1111/j.2044-8295.1980.tb01743.x

Frings, C., Hommel, B., Koch, I., Rothermund, K., Dignath, D., Giesen, C. G. ..., Philipp, A. (2020). Binding and retrieval in action control (BRAC). *Trends in Cognitive Sciences, 24*(5). 375-387. https://doi.org/10.1016/j.tics.2020.02.004

Frings, C. & Rothermund, K. (2011). To be or not to be ... included in an event file: Integration and retrieval of distractors in stimulus-response episodes is influenced by perceptual grouping. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(5), 1209-1227. https://doi.org/10.1037/a0023915

Frings, C. & Rothermund, K. (2017). How perception guides action: Figure-ground segmentation modulates integration of context features into s-r episodes. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, 43(11), 1720-1729. https://doi.org/10.1037/xlm0000403

Frings, C., Rothermund, K., & Wentura, D. (2007). Distractor repetitions retrieve previous responses to targets. *The Quarterly Journal of Experimental Psychology, 60*(10), 1367-1377. https://doi.org/10.1080/17470210600955645

Garcia, J. & Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science, 4*, 123–124. https://doi.org/10.3758/BF03342209

Gibbons, H. & Stahl, J. (2008). Early activity in the lateralized readiness potential suggests prime-response retrieval as a source of negative priming. *Experimental Psychology, 55*(3), 164-172. https://doi.org/10.1027/1618-3169.55.3.164

Giesen, C., Frings, C., & Rothermund, K. (2012). Differences in the strength of distractor inhibition do not affect distractor-response bindings. *Memory & Cognition, 40*, 373-387. https://doi.org/10.3758/s13421-011-0157-1

Giesen, C., & Rothermund, K. (2011). Affective matching moderates S-R binding. *Cognition and Emotion, 25*(2), 342–350. https://doi.org/10.1080/02699931.2010.482765

Giesen, C., & Rothermund, K. (2014a). You better stop! Binding "stop" tags to irrelevant stimulus features. *The Quarterly Journal of Experimental Psychology, 67*, 809-832. https://doi.org/10.1080/17470218.2013.834372

Giesen, C. & Rothermund, K. (2014b). Distractor repetitions retrieve previous responses and previous targets. Experimental dissociations of distractor-response and distractor-target bindings. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(3), 645-659. https://doi.org/10.1037/a0035278

Giesen, C. & Rothermund, K. (2016). Multi-level response coding in stimulus-response bindings: Irrelevant distractors retrieve both semantic and motor response codes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(10), 1643-1656. https://doi.org/10.1037/xlm0000264

Giesen, C, Scherdin, K., & Rothermund, K. (2017). Flexible goal imitation: Vicarious feedback influences stimulus-response binding by observation. *Learning & Behavior, 45*, 147-156. https://doi.org/10.3758/s13420-016-0250-1

Giesen, C. G., Schmidt, J. R., & Rothermund, K. (2020). The law of recency: An episodic stimulus-response retrieval account of habit acquisition. *Frontiers in Psychology, 10*, 2927. https://doi.org/10.3389/fpsyg.2019.02927

Goel, V., & Pirolli, P. (1992). The structure of design problem spaces. *Cognitive Science*, *16*(3), 395–429. https://doi.org/10.1207/s15516709cog1603_3

Greene, J. O., O'Hair, H. D., Cody, M. J., & Yen, C. (1985). Planning and control of behavior during deception. *Human Communication Research*, *11*(3), 335–364. https://doi.org/10.1111/j.1468-2958.1985.tb00051.x

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Hartwig, M. & Bond, C. F., Jr. (2014). Lie detection from multiple cues: A meta-analysis. *Applied Cognitive Psychology*, *28*(5), 661–676. https://doi.org/10.1002/acp.3052

Henson, R. N., Eckstein, D., Waszak, F., Frings, C., & Horner, A. J. (2014). Stimulus-response bindings in priming. *Trends in Cognitive Sciences, 18*(7), 376-384. https://doi.org/10.1016/j.tics.2014.03.004

Herwig, A. & Waszak, F. (2012). Action-effect bindings and ideomotor learning in intention- and stimulus-based actions. *Frontiers in Psychology, 3,* 444. https://doi.org/10.3389/fpsyg.2012.00444

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition, 5*, 183-216. https://doi.org/10.1080/713756773

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8*, 494-500. https://doi.org/10.1016/j.tics.2004.08.007

Hommel, B. (2019). Theory of Event Coding (TEC) V2.0: Representing and controlling perception and action. *Attention, Perception, & Psychophysics, 81,* 2139–2154. https://doi.org/10.3758/s13414-019-01779-4

Hommel, B. & Colzato, L. S. (2009). When an object is more than a binding of its features: Evidence for two mechanisms of visual feature integration. *Visual Cognition, 17*(1/2), 120-140. https://doi.org/10.1080/13506280802349787

Horner, A. J. & Henson, R. N. (2009). Bindings between stimuli and multiple response codes dominate long-lag repetition priming in speeded classification tasks. *Journal*

of Experimental Psychology: Learning, Memory and Cognition, 35, 757-779. https://doi.org/10.1037/a0015262

Horner, A. J. & Henson, R. N. (2011). Stimulus-response bindings code both abstract and specific representations of stimuli: Evidence from a classification priming design that reverses multiple levels of response representation. *Memory & Cognition, 39,* 1457-1471. https://doi.org/10.3758/s13421-011-0118-8

Hunt, R. R. (2006). The concept of distinctiveness in memory research. In R. R. Hunt & J. B. Worthen (Eds.), *Distinctiveness and Memory* (pp. 3-25). Oxford University Press.

Klein, K. & Saltz, E. (1976). Specifying the mechanisms in a levels-of-processing approach to memory. *Journal of Experimental Psychology: Human Learning and Memory, 2*(6), 671-679. https://doi.org/10.1006/ccog.1996.0009

Koranyi, N., Schreckenbach, F., & Rothermund, K. (2015). The implicit cognition of lying: Knowledge about having lied to a question is retrieved automatically. *Social Cognition, 33*, 67-84. https://doi.org/10.1521/soco.2015.33.1.67

Lane, J. D., & Wegner, D. M. (1995). The cognitive consequences of secrecy. *Journal of Personality and Social Psychology*, *69*(2), 237–253. https://doi.org/10.1037/0022-3514.69.2.237

Laub, R. & Frings, C. (2021). Brightness versus darkness: The influence of stimulus intensity on the distractor-response binding effect. *Acta Psychologica, 212.* https://doi.org/10.1016/j.actpsy.2020.103224

Levine, T. R., Lapinski, M. K., Banas, J., Wong, N. C. H., Hu, A. D. S., Endo, K., Baum, K. L., & Anders, L. N. (2002). Self-Construal, Self and Other Benefit, and the Generation of Deceptive Messages. *Journal of Intercultural Communication Research, 31*(1), 29– 47.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492-527. https://doi.org/10.1037/0033-295X.95.4.492

Logan, G. D. (1992). Attention and preattention in theories of automaticity. *American Journal of Psychology, 105*(2), 317-339. https://doi.org/10.2307/1423031

Logie, R. H., Brockmole, J. R., & Vandenbroucke, A. R. E. (2009). Bound feature combinations in visual short-term memory are fragile but influence long-term

learning. *Visual Cognition, 17*(1/2), 160-179. https://doi.org/10.1080/13506280802228411

Mayr, S. & Buchner, A. (2006). Evidence for episodic retrieval of inadequate prime responses in auditory negative priming. *Journal of Experimental Psychology: Human Perception and Performance, 32*(4), 932-943. https://doi.org/10.1037/0096-1523.32.4.932

McCornack, S., Morrison, K., Paik, J. E., Wisner, A. M., & Zhu, X. (2014). Information manipulation theory 2: A propositional theory of deceptive discourse production. *Journal of Language and Social Psychology, 33*, 348–377. https://doi.org/10.1177/0261927X14534656

Memelink, J. & Hommel, B. (2012). Intentional weighting: A basic principle in cognitive control. *Psychological Research, 77*, 249-259. https://doi.org/10.1007/s00426-012-0435-y

Moeller, B. & Frings, C. (2014). Long-term response-stimulus associations can influence distractor-response bindings. *Advances in Cognitive Psychology, 10*(2), 68-80. https://doi.org/10.5709/acp-0158-1

Moeller, B. & Frings, C. (2017). Overlearned responses hinder s-r binding. *Journal of Experimental Psychology: Human Perception and Performance, 43*(1), 1-5. https://doi.org/10.1037/xhp0000341

Moeller, B., Frings, C., & Pfister, R. (2016). The structure of distractor-response bindings: Conditions for configural and elemental integration. *Journal of Experimental Psychology: Human Perception and Performance, 42*(4), 474-479. https://doi.org/10.1037/xhp0000158

Nett, N., Bröder, A., & Frings, C. (2015). When irrelevance matters: Stimulus-response binding in decision making under uncertainty. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(6), 1831-1848. https://doi.org/10.1037/xlm0000109

Nett, N., Bröder, A., & Frings, C. (2016). Distractor-based stimulus-response bindings retrieve decisions independent of motor programs. *Acta Psychologica, 171*, 57-64. https://doi.org/10.1016/j.actpsy.2016.09.006

O'Hair, H. D., Cody, M. J., & McLaughlin, M. L. (1981). Prepared lies, spontaneous lies, Machiavellianism, and nonverbal communication. *Human Communication Research, 7*(4), 325–339. https://doi.org/10.1111/j.1468-2958.1981.tb00579.x

Pittarello, A., Rubaltelli, E., & Motro, D. (2016). Legitimate lies: The relationship between omission, commission, and cheating. *European Journal of Social Psychology, 46*(4), 481–491. https://doi.org/10.1002/ejsp.2179

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*(3), 219–235. https://doi.org/10.1037/0096-3445.118.3.219

Rogers, T., Zeckhauser, R., Gino, F., Norton, M. I., & Schweitzer, M. E. (2017). Artful paltering: The risks and rewards of using truthful statements to mislead others. *Journal of Personality and Social Psychology, 112*(3), 456–473. https://doi.org/10.1037/pspi0000081

Rothermund, K., Wentura, D., & De Houwer, J. (2005). Retrieval of incidental stimulus-response associations as a source of negative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(3), 482-495. https://doi.org/10.1037/0278-7393.31.3.482 2005, Vol. 31, No. 3, 482–495

Schmidt, J. R. (2013). The Parallel Episodic Processing (PEP) model: Dissociating contingency and conflict adaptation in the item-specific proportion congruent paradigm. *Acta Psychologica, 142*, 119-126. https://doi.org/10.1016/j.actpsy.2012.11.004

Schmidt, J. R., De Houwer, J., & Rothermund, K. (2016). The Parallel Episodic Processing (PEP) model 2.0: A single computational model of stimulus-response binding, contingency learning, power curves, and mixing costs. *Cognitive Psychology, 91*, 82-108. https://doi.org/10.1016/j.cogpsych.2016.10.004

Schmidt, J. R., Giesen, C. G., & Rothermund, K. (2020). Contingency learning as binding? Testing an exemplar view of the colour-word contingency learning effect. *Quarterly Journal of Experimental Psychology, 73*(5), 739-761. https://doi.org/10.1177/1747021820906397

Schweitzer, M. E., & Croson, R. (1999). Curtailing deception: The impact of direct questions on lies and omissions. *International Journal of Conflict Management, 10*(3), 225-248. https://doi.org/10.1108/eb022825

van Bockstaele, B., Verschuere, B., Moens, T., Suchotzki, K., Debey, E., & Spruyt, A. (2012). Learning to lie: effects of practice on the cognitive cost of lying. *Frontiers in Psychology, 3,* 526. https://doi.org/10.3389/fpsyg.2012.00526

Verbruggen, F. & Logan, G. D. (2008). Long-term aftereffects of response inhibition: Memory retrieval, task goals, and cognitive control. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1229-1235. https://doi.org/10.1037/0096-1523.34.5.1229

Verschuere, B., Spruyt, A., Meijer, E. H., & Otgaar, H. (2011). The ease of lying. *Consciousness and Cognition: An International Journal*, *20*(3), 908–911. https://doi.org/10.1016/j.concog.2010.10.023

von Restorff, H. (1933). Über die Wirkung von Bereichsbildungen im Spurenfeld [The effects of field formation in the trace field]. *Psychologische Forschung, 18*, 299-342.

Walczyk, J. J., Harris, L. L., Duck, T. K., & Mulay, D. (2014). A social-cognitive framework for understanding serious lies: Activation-decision-construction-action theory. *New Ideas in Psychology*, *34*, 22–36. https://doi.org/10.1016/j.newideapsych.2014.03.001

Waszak, F. & Hommel, B. (2007). The costs and benefits of cross-task priming. *Memory & Cognition, 35*(5), 1175-1186. https://doi.org/10.3758/bf03193487

Wieber, F. & Sassenberg, K. (2006). I can't take my eyes off of it – Attention attraction effects of implementation intentions. *Social Cognition, 24*(6), 723–752. https://doi.org/10.1521/soco. 2006.24.6.723

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. *Advances in experimental social psychology, 14*, 1–59. https://doi.org/10.1016/S0065-2601(08)60369-X

Appendix A: Supplemental Material for Chapter 2

This Appendix presents the supplemental online material of the following article:

Schreckenbach, F., Rothermund, K., & Koranyi, N. (2020). Quantity matters: The frequency of deception influences automatic memory retrieval effects. *Quarterly Journal of Experimental Psychology, 73*(11), 1774-1783. https://doi.org/10.1177/1747021820924652

This material was subject to peer-review and is thus presented in the form that has been accepted for publication.

Appendix

English Version of the Questions Used in the Experiment

The questions of one topic were always asked in the depicted order. The assignment of topic to frequency condition was counterbalanced, the order of topics during the interview was randomized.

Topic 1: University Life

1. Why did you choose your subject of study?
2. What profession do you want to pursue after you have finished your studies?

Topic 2: Leisure Time

1. How often do you meet friends?
2. What did you do last Sunday?

Topic 3: Money

1. How do you finance your studies?
2. How do you like to spend your money?

Topic 4: Family

1. How long have you been living with your parents?
2. How many siblings do you have?

Filler Trials in the RT task that were not asked during the Interview

1. When do you usually get up in the morning?
2. How old are you?
3. What is your favorite dish?
4. Where do you usually go for lunch?
5. What is your shoe size?
6. When was the last time you were at the hairdresser?

Mean Response Time (RT) as a Function of Prime Question (Truth vs. Lie), Probe Word (Honest vs. Dishonest), and Context (Rare Lie vs. Rare Truth)

	Prime Question			
	Truth		Lie	
	Probe Word			
Context	Honest	Dishonest	Honest	Dishonest
Rare Lie	649 (98)	647 (113)	643 (78)	626 (80)
Rare Truth	656 (154)	639 (137)	668 (165)	663 (165)

Note. Standard deviations appear in parentheses.

Appendix B: Supplemental Material for Chapter 3

This Appendix presents the supplemental online material of the following article:

Schreckenbach, F., Rothermund, K., & Koranyi, N. (2019). And remember the truth that once was spoken: Knowledge of having disclosed private information to a stranger is retrieved automatically. *Experimental Psychology, 66*(1), 12-22. https://doi.org/10.1027/1618-3169/a000427

This material was subject to peer-review and is thus presented in the form that has been accepted for publication.

Supplementary Material

English Version of the Questions Used in the Experiments

The questions of one topic were always asked in the depicted order. The assignment of topic to truth condition as well as the order of topics during the interview was counterbalanced in both studies.

Topic 1: University

1. How important are good grades to you?
2. Did you ever fail an examination?
3. Why did you choose your subject of study?
4. What profession do you want to pursue after you have finished your studies?

Topic 2: Family

1. What is your mother's first name?
2. How long have you been living with your parents?
3. How was your children's room furnished?
4. Do you want to have children someday?

English Version of the Questions that were Used to Check Participants' Memory of Instructions and their prior Knowledge about their Interviewers

Memory of Instructions:

- Do you remember whom you were supposed to tell the truth? Please type your answer into the textbox below.

Knowledge about Interviewers:

1. Did you know any of the persons that appeared in the Experiment in advance?

If participants agree to the prior question, a picture of each person is depicted together with the question:

2. Do you know this person? Press J for yes and F for no.

If participants again agree, a second screen appears with a picture of the person and the following request:

3. Please type the name of this person into the textbox and also where you know this person from.

Appendix C: Supplemental Material for Chapter 4

This Appendix presents the supplemental online material of the following manuscript:

Schreckenbach, F. & Rothermund, K. (2021). *Feature specific binding of the knowledge of having lied before: Evidence for elemental integration.* MS submitted for publication.

This material is subject to peer-review and is thus presented in the form that has been submitted for publication.

Supplementary Material – Pilot study

In a first attempt to test whether bindings in the context of lying are configural or elemental, we used an experimental setup where participants met only one interrogator and had to lie about one topic while telling the truth about a second topic. During the subsequent classification task, participants were presented with pictures of either their actual interrogator or an unknown person. On top of this picture, a question from the interview appeared, thereby leading to four different combinations of primes: (a) the interrogator was presented with a question that participants had told the truth to during the interview, (b) the interrogator was presented with a question that participants had lied to, (c) an unknown person was presented with a truthfully answered question, or (d) an unknown person was presented with a deceptively answered question.

Method

Participants and design. The study was part of the same ethic approval as the main experiment (FSV 19/18). Based on the same power analysis (effect size $\eta_p^2 = .13$, $\alpha = .05$, $1 - \beta = .8$), we recruited 69 students from a German University. The data of five subjects could not be meaningfully analyzed due to excessive error rates (> 30 %) in the priming task (average 3.2 % errors, without excluded participants). The final sample then consisted of 64 subjects (39 female) with an average age of $M = 23.4$ years ($SD = 3.61$). All participants gave informed consent via keypress at the beginning of the experiment. For their participation they received 4 € and a chocolate bar. We used a 2 x 2 x 2 factorial design with the within-subjects factors prime picture (facial photo of the person one has met during the interview vs. an unfamiliar person), prime question (question that one has told the truth vs. lied to during the interview) and probe word (honest vs. dishonest).

Procedure and materials. The procedure was the same as in the main experiment with the following exceptions: The material for the interview comprised 12 questions, four of which

related to one of two different topics (family, university, and vacation). Counterbalanced across participants, questions from two of the topics were selected for the interview and participants were instructed to reply untruthfully to all questions concerning one specified topic and truthfully to all questions from the second topic. Participants again received instructions on the screen, but this time only one female interrogator was presented on the left side of the screen, while on the right side the interviewer's name, the topics that participants would be asked about, and the instruction when to tell the truth and when to lie was presented (e.g., "This is Julia. Julia is going to ask you some questions about the topics 'family' and 'university'. Please tell the truth to all questions about family, but lie to all questions about university.") The order of instructions as well as the assignment of topic or interrogator to response instruction was counterbalanced across participants, as was the order of topics during the subsequent interviews. After the instructions, participants were guided to a separate room where the interview took place.

The procedure of the priming task was identical to that of the main experiment. The picture of the interrogator as well as the picture of an unknown person served as familiar vs. unfamiliar prime stimuli during the priming task. The priming task comprised 160 experimental trials in an individually randomized order. The prime pictures were presented 80 times each, 10 times preceding each of the 8 questions from the interview. Every combination of primes was half of the times followed by the degraded probe word *ehrlich* and half of the times by the degraded probe word *gelogen*. The memory task had to be performed in 16 of the experimental trials and 30 filler trials requiring a genuine true/false decision were randomly intermixed among the other trials.

Results

Data treatment. All response latencies that were more than one and a half interquartile ranges above the third quartile of an individual's reaction time distribution were categorized as outliers (Tukey, 1977) and discarded (5.6% of all responses). All response latencies below the threshold of 250 ms were discarded (0.1%), as well as erroneous responses (3.2% of all responses). For each participant, we calculated the difference in response times for *honest* and *dishonest* probe words as an indicator of retrieval effects ($RT_{honest} - RT_{dishonest}$; positive values reflect faster identification of the probe word dishonest, thus indicating retrieval of the knowledge of having lied), separately for each combination of prime picture and prime question in the factorial design (see Figure 1 for the pattern of means).

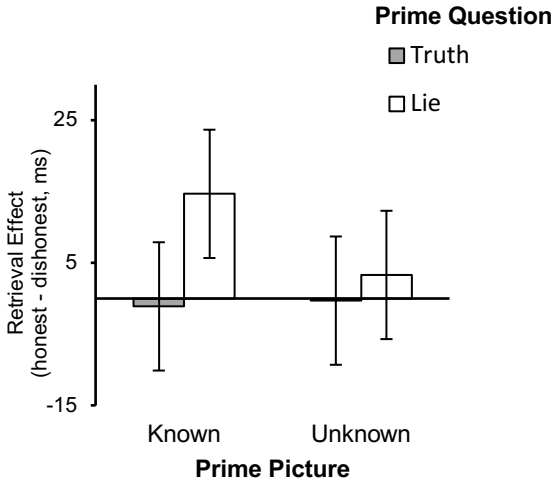

Figure 1. *Mean retrieval effects (positive values reflect faster identification of the probe word dishonest, thus indicating retrieval of the knowledge of having lied) as a function of prime picture (known vs. unknown) and prime question (truth vs. lie. Error bars represent 95 % CIs calculated for RM interaction effects as suggested in Jarmasz & Hollands (2009).*

Analyses of variance. To test our assumptions, average retrieval effects were submitted to a 2 (prime picture: known vs. unknown) x 2 (prime question: truth vs. lie) ANOVA with repeated measures on all factors. A significant interaction of prime picture and prime question emerged, F (1,63) = 4.51, p = .038, η_p^2 = .07 (90% CI [.01, .18]), thereby lending support to the assumption of an integrated use of cues.

To further disentangle this interaction, we ran separate analyses for known vs. unknown prime pictures. For known pictures the predicted main effect reached significance, F (1,63) = 2.83, p = .049, η_p^2 = .04 (90% CI [.00, .15]) (one-tailed) [1], with the descriptive pattern pointing in the predicted direction (see Figure 1). Correspondingly, follow-up tests showed a significant retrieval effect indicating faster responding to the probe word *dishonest* for prime questions to which one had lied ($M_{retrieval}$ = 15 ms), t(63) = 2.28 p = .026, d = .28 (95% CI [.03, .53]), whereas no difference in responding to honest and dishonest probes was found after prime questions to which one had told the truth ($M_{retrieval}$ = -1 ms), t < 1. For unknown pictures no significant results emerged (all F < 1). This pattern of results lends first support to the assumption of configural binding and retrieval processes under certain conditions.

Bayesian statistics. The Bayes factors and the posterior probabilities (computed assuming equal prior probabilities with the same software as used in the main experiment) are displayed in Table 1. Constraints for hypotheses 1 (corresponding to an elemental binding account assuming the use of questions as retrieval cue regardless of the presented person) and 2 (corresponding to configural binding, i.e., storage and retrieval of both, persons and questions as retrieval cues) were defined in a way that they display retrieval effects for both prime pictures for hypothesis 1 and retrieval effects only for known pictures for hypothesis 2. As can be seen,

[1] Methodologically, this F-test is equivalent to a t-test that tests the difference between congruent and incongruent combinations of prime questions and probe words against zero. Thus, given our specific predictions, a one-tailed test is recommended in order to increase the power of the test (Maxwell & Delaney, 1990, p.144).

H_2 is supported more than H_1. A Bayes Factor of 37.095 is usually seen as very strong evidence in favor of H_1. The Bayesian error probability associated with preferring H_2 equals .098.

Table 2. *Bayes factors and posterior probabilities for Hypotheses 1 and 2, each in comparison to H_a which assumes unconstrained means in all conditions.*

Hypothesis	BF.a	posterior probability
H_1	3.462	.073
H_2	37.047	.903
H_a		.024

Appendix D: Supplemental Material for Chapter 5

This Appendix presents the supplemental online material of the following article:

Schreckenbach, F., Sprengholz, P., Rothermund, K., & Koranyi, N. (2020). How to remember something you didn't say: Lies of omission can be stored and retrieved from memory. *Experimental Psychology, 67*(6), 364-372. https://doi.org/10.1027/1618-3169/a000504

This material was subject to peer-review and is thus presented in the form that has been accepted for publication.

https://doi.org/10.1027/1618-3169/a000504

Supplemental Material

Pictures of potential lovers:

Heterosexual female and homosexual male participants had to choose their lover from three male faces:

Homosexual female and heterosexual male participants had to choose their lover from three female faces:

https://doi.org/10.1027/1618-3169/a000504

Cover Story including instructions about what information had to be remembered

The vignette depended on participants' input. The following example text (translated from German) was shown if a female heterosexual subject had a partner called Tom, a friend named Anna and a lover named Paul:

> Please try to imagine the following Scenario. Read carefully.
>
> Since days you were looking forward to this weekend. Your partner Tom had brought you to the train station wishing you a lot of fun. He only knew that you wanted to visit Anna, an old friend in Hamburg. But that was only half the truth.
>
> When the train arrived in Hamburg, Anna picked you up. You had some fish and later rented a boat for sightseeing. At night you visited the Reeperbahn.

> The next day Anna had to work but that did not bother you. You wanted to travel to Berlin and see Paul again. You two met some weeks ago and the initial flirt had turned into a hot affair.
>
> You took the bus to Berlin and when you arrived there, Paul was happy to see you. You rented a Trabi, visited the Bundestag and ate a Currywurst together.

https://doi.org/10.1027/1618-3169/a000504

Unfortunately you had to return home after some hours. You did not like to say goodbye, but the last train was already waiting for you.

Probably Tom will chat with you in a moment. Of course, he will ask about your weekend and how the trip was. You should not lie to him, but you better omit information about Berlin.

Consequently you should suppress that you were driving a Trabi, ate Currywurst and visited the Bundestag.

You better tell him about Hamburg, that you travelled by Boat, ate Fish and visited the Reeperbahn.

Memorize which information you can share with Tom (marked green) and which must be omitted (marked red).

https://doi.org/10.1027/1618-3169/a000504

List of Questions and Answers Used During the Decision Task

Question	**Answer 1** Utterance expected	**Answer 2** Suppression expected
Do you think I'm attractive?	Yes	No
Do you love me?	Yes	No
Would you rather be with someone else?	No	Yes
Do you cheat on me?	No	Yes
Do you think others are more attractive than I am?	No	Yes
Whom did you visit this weekend?	Friend's name	Lover's name
Where were you this weekend?	Hamburg	Berlin

Publikationen

Die mit einem Sternchen (*) versehenen Arbeiten sind Teil dieser Dissertation

Submitted

*Schreckenbach, F. & Rothermund, K. *Feature-specific binding of the knowledge of having lied before. Evidence for elemental integration.* Manuscript submitted for publication.

Sprengholz, P., Schreckenbach, F., Giesen, C., & Rothermund, K. *How to (not) get away with murder. Improving the reaction time based concealed information test by considering automatic retrieval of guilty knowledge.* Manuscript submitted for publication.

2020

*Schreckenbach, F., Sprengholz, P., Rothermund, K., & Koranyi, N. (2020). How to remember something you didn't say. Lies of omission can be stored and retrieved from memory. *Experimental Psychology, 67*(6), 364-372. https://doi.org/10.1027/1618-3169/a000504

*Schreckenbach, F., Rothermund, K., & Koranyi, N. (2019). Quantity matters. The frequency of deception influences automatic memory retrieval effects. *Quarterly Journal of Experimental Psychology, 73*(11), 1774-1783. https://doi.org/10.1177/1747021820924652

2019

*Schreckenbach, F., Rothermund, K., & Koranyi, N. (2019). And remember the truth that once was spoken: Knowledge of having disclosed private information to a stranger is retrieved automatically. *Experimental Psychology, 66*(1), 12-22. https://doi.org/10.1027/1618-3169/a000427

2015

Koranyi, N., Schreckenbach, F., & Rothermund, K. (2015). The implicit cognition of lying: Knowledge about having lied to a question is retrieved automatically. *Social Cognition, 33*, 67-84. https://doi.org/10.1521/soco.2015.33.1.67

Invited Talks

Schreckenbach, F. (2018, January). *Lügner brauchen ein gutes Gedächtnis. Untersuchung automatischer Abrufprozesse von gelogenen und ehrlichen Aussagen.* Vortrag im Rahmen des Forschungskolloquiums der Abteilung für Sozial- und Rechtspsychologie der Johannes Gutenberg-Universität Mainz, Deutschland.

Conference Contributions

Schreckenbach, F., Koranyi, N., & Rothermund, K. (2018, September). If you always tell the truth you don't have to remember anything. New insights into the implicit memory of lying [Poster presentation]. 51st Kongress der Deutschen Gesellschaft für Psychologie (DGPs), Frankfurt a.M., Germany.

Schreckenbach, F., Koranyi, N., & Rothermund, K. (2018, July). Lügengeschichten. Untersuchung automatischer Abrufprozesse von gelogenen und ehrlichen Aussagen [Conference session]. 7th Doktoranden-Workshop Allgemeine Psychologie (A-Dok), Mainz, Germany.

Schreckenbach, F., Koranyi, N., & Rothermund, K. (2018, March). If you always tell the truth you don't have to remember anything. New insights into the implicit memory of lying [Poster presentation]. 60th Conference of Experimental Psychologists (Tagung experimentell arbeitender Psychologen, TeaP), Marburg, Germany.

Schreckenbach, F., Koranyi, N., & Rothermund, K. (2016, March). Telling ordinary lies. The role of distinctiveness in implicit retrieval of lies. [Conference session]. 58th Conference of Experimental Psychologists (Tagung experimentell arbeitender Psychologen, TeaP), Heidelberg, Germany.

Schreckenbach, F., Koranyi, N., & Rothermund, K. (2015, March). The implicit cognition of lying [Conference session]. 57th Conference of Experimental Psychologists (Tagung experimentell arbeitender Psychologen, TeaP), Hildesheim, Germany.¶

Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass mir die geltende Promotionsordnung der Fakultät für Sozial- und Verhaltenswissenschaften der Friedrich-Schiller-Universität Jena bekannt ist.

Ferner erkläre ich, dass ich die vorliegende Arbeit selbst und ohne unzulässige Hilfe Dritter angefertigt, keine Textabschnitte eines Dritten oder eigener Prüfungsarbeiten ohne Kennzeichnung übernommen und alle von mir benutzten Hilfsmittel, persönlichen Mitteilungen und Quellen in meiner Arbeit angegeben habe. Insbesondere habe ich für diese Arbeit nicht die Hilfe eines Promotionsberaters in Anspruch genommen. Dritte haben weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die studentischen Hilfskräfte der Abteilung für Allgemeine Psychologie II des Instituts für Psychologie der Universität Jena haben mich bei der Erhebung der Daten unterstützt. Die Manuskripte habe ich jeweils gemeinsam mit Prof. Dr. Klaus Rothermund, in drei Fällen mit Dr. Nicolas Koranyi sowie in einem Fall mit Philipp Sprengholz erstellt (vgl. Abschnitt *Angaben zum Eigenanteil*).

Die vorliegende Dissertation wurde noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht. Ferner habe ich weder die gleiche noch eine in wesentlichen Teilen ähnliche oder eine andere Abhandlung bei einer anderen Hochschule oder Fakultät als Dissertation eingereicht.

Ich versichere, dass ich nach bestem Wissen die reine Wahrheit sage und nichts verschwiegen habe.

Ort, Datum	Franziska Schreckenbach

Bemerkung zu Kapitel 3

Die in Kapitel 3 vorgestellte Studie macht eine gesonderte Bemerkung notwendig. Die Fragestellungen der vorliegenden Dissertation bauen inhaltlich auf meiner Masterarbeit auf. Im Rahmen dieser Masterarbeit habe ich bereits das Experiment durchgeführt, das in Kapitel 3 als Experiment 1 beschrieben wird. Als Doktorandin habe ich dann die für die Publikation maßgeblichen inhaltlichen Ideen entwickelt, eine leicht veränderte Art der Auswertung vorgenommen sowie Experiment 2 durchgeführt, ausgewertet und in das Manuskript aufgenommen. Aufgrund der inhaltlichen Kohärenz zu den anderen Kapiteln und aufgrund der Tatsache, dass dieses Kapitel einen wichtigen inhaltlichen Punkt bedient, habe ich mich für eine Aufnahme des Artikels in diese Arbeit entschieden.

Laut der Hinweise zur kumulativen Dissertation der FSU gilt: „Publikationen, die vorrangig Ergebnisse aus der Abschlussarbeit eines der Promotion vorausgegangenen Studiums des Doktoranden/der Doktorandin darstellen, können nicht Bestandteil einer kumulativen Dissertation sein.“ Das in Kapitel 3 vorgestellte Experiment 1 wurde bereits im Rahmen einer früheren Qualifikationsarbeit konzipiert; auch die Daten wurden in diesem Rahmen erhoben. Zusätzlich wurden jedoch weitere Daten erhoben sowie ein veränderter inhaltlicher Schwerpunkt bei der Erstellung der Publikation gesetzt. Sollten die Gutachter trotzdem Bedenken haben, sollte dieses Kapitel aus der Bewertung ausgeschlossen werden. Auch ohne Kapitel 3 sind die Anforderungen an eine kumulative Dissertation erfüllt (nicht weniger als drei Artikel, die Mehrzahl davon publiziert).

Angaben zum Eigenanteil

Die vorliegende Dissertation wurde im Rahmen einer publikationsbasierten Promotion erstellt. Teil dieser Arbeit sind vier Artikel, die im Rahmen meiner Tätigkeit als wissenschaftliche Mitarbeiterin am Lehrstuhl für Allgemeine Psychologie II bei Prof. Dr. Klaus Rothermund entstanden sind, und deren Erstautorin ich bin. Drei Artikel, die Teil dieser Arbeit sind, wurden in Zeitschriften mit Peer-Review-Verfahren publiziert. Der vierte Artikel wurde in einer solchen Zeitschrift zur Publikation eingereicht.

Alle vier Artikel wurden in Zusammenarbeit mit Prof. Dr. Klaus Rothermund erstellt, drei davon außerdem in Zusammenarbeit mit Dr. Nicolas Koranyi; bei einem Manuskript wirkte Philipp Sprengholz als Ko-Autor mit. Bei allen vier empirischen Kapiteln hatte ich mehrheitlich Anteil daran, die Hypothesen und Experimentaldesigns zu entwickeln. Die Datenerhebung der Studien in den Kapiteln 2 bis 4 habe ich vollständig selbständig organisiert; unterstützt wurde ich bei der Erhebung selbst von den wissenschaftlichen Hilfskräften des Lehrstuhls für Allgemeine Psychologie II. Die Daten des Experiments in Kapitel 5 wurden von Philipp Sprengholz erhoben. Alle benötigten Datenanalysen wurden von mir ausgeführt. Als Erstautorin der empirischen Artikel habe ich jeweils die erste Version aller Manuskripte (und ggf. Revisionen) verfasst. Diese wurden anschließend von meinem Koautor Prof. Dr. Klaus Rothermund, in drei Fällen von Dr. Nicolas Koranyi, sowie in einem Fall von Philipp Sprengholz überarbeitet. Alle Artikel dieser Dissertation spiegeln die geteilten Überzeugungen aller Autoren wider.

_____ _____

Ort, Datum Franziska Schreckenbach