Capturing spontaneous trait inference with the modified free association paradigm

Orghian, Diana; Smith, Anna; Garcia-Marquesa, Leonal; Heinke, Dietmar

DOI:
10.1016/j.jesp.2017.07.004

Capturing spontaneous trait inference with the modified free association paradigm

Diana Orghian
Faculdade de Psicologia, Universidade de Lisboa
Psychology Department, Harvard University

Anna Smith
Psychology Department, Birmingham University

Leonel Garcia-Marques
Faculdade de Psicologia, Universidade de Lisboa

Dietmar Heinke
Psychology Department, Birmingham University

Author Note

Correspondence concerning this article should be addressed to Diana Orghian, room 280, Department of Psychology, Harvard University, 33 Kirkland St, Cambridge, MA, 02138; E-mail: dorghian@fp.ul.pt.

The authors thank the anonymous reviewers for the insightful comments and Elizabeth Page-Gould for the suggestions provided during the editorial process.

This research was supported by Fundação para a Ciência e Tecnologia (Portuguese National Foundation for Science and Technology) under the Grant DFRH - SFRH/BD/84668/2012.

Abstract

Spontaneous Trait Inference (STI) is a psychological phenomenon that allows people to infer personality traits from others' behaviors without any intention or awareness. This phenomenon helps us organize our complex social world and ultimately it shapes our interaction with others. The goal of this manuscript is 1) to conduct a systematic analysis of the different types of measures that are currently used to detect inferences, 2) to identify their main limitations and 3) to propose a new paradigm that overcomes most of these limitations. The new paradigm is based on the modified free association task that Hourihan and MacLeod (2007) proposed as a pure conceptual implicit memory measure. This new paradigm simultaneously overcomes the contamination problem of memory measures and the dependency on data-driven processing of the activation measures. In the presented experiments we use the modified free association paradigm to detect STIs and also to compare its results to the naming task and the modified Stroop task. We show that this measure is able to reliably detect STI and discuss why it is more appropriate to study inferences. Additionally, we show that it can be used to investigate the underlying processes responsible for STI and to distinguish it from similar but different phenomena like Spontaneous Trait Transference. This manuscript focuses on STIs, but the arguments presented here apply to any research field that investigates spontaneous inference making.

Keywords: inference making, activation measures, memory measures, modified free association paradigm, conceptually-driven task.

Capturing spontaneous trait inference with the modified free association paradigm

Introduction

Spontaneous trait inference (STI) is a social inference that occurs when, for example, the man who holds the door for someone is unintentionally categorized as "friendly" or the woman who jumps the line as "rude". The inference of such traits constitutes the basis of the impressions we create about others. In turn, impressions shape the way we interact with others – perhaps avoiding the rude woman or seeking out the friendly man. Winter and Uleman (1984) provided the first empirical evidence for STI by showing that participants unintentionally inferred personality traits after reading behavioral descriptions. Since then, studies have shown that trait inference is a spontaneous process that occurs outside our awareness (Lupfer, Clark, & Hutcherson, 1990) and during the encoding of behavioral descriptions (*e.g.*, Carlston & Skowronski, 1994; Winter & Uleman, 1984; Uleman, Newman, & Moskowitz, 1996). It was shown that STI is maintained across more demanding experimental settings including short stimulus presentation timings and cognitive load (*e.g.*, Carlston & Skowronski, 1994; Todorov & Uleman, 2003) and that it lingers after explicit recall has faded (Carlston, Skowronski, & Sparks, 1995).

STI is typically investigated with the following paradigms: cued-recall (*e.g.*, Winter & Uleman, 1984; Winter, Uleman, & Cunniff, 1985; Todorov & Uleman, 2002; Van Duynslaeger, Van Overwalle, & Verstraeten, 2007), probe recognition (*e.g.*, Uleman, Hon, Roman, & Moskowitz, 1996; Newman, 1991; Ham & Vonk, 2003; Wigboldus, Sherman, Franzese, & van Knippenberg, 2004), savings in relearning (*e.g.*, Carlston & Skowronski, 1994; Carlston et al., 1995; Carlston & Skowronski, 2005; Crawford, Skowronski, Stiff, & Leonards, 2008; McCarthy & Skowronski, 2011; Schneid, Crawford, Skowronski, Irwin, & Carlston, 2015), and false recognition (*e.g.*, Todorov & Uleman, 2002, 2003, 2004; Goren & Todorov, 2009; Rim, Min, Uleman, Chartrand, & Carlston, 2013; Lee, Shimizu, & Uleman, 2015). Despite the extensive research and the valuable knowledge brought by these paradigms to the field, they present several shortcomings.

Over time, the paradigms were refined in agreement with the emerging concerns in the field and, as a consequence, some of their limitations were solved while others remain unattended. The first goal of the paper is to present the reader with a compendium of requirements that a good paradigm should satisfy in order to rigorously measure STI. The second objective is to inspect how and which requirements are fulfilled in each of the currently used paradigms. And finally, we will introduce a new paradigm, the modified free association paradigm (MFAP), as a possible solution to overcome most of the limitations discussed throughout the paper.

A paradigm, in order to rigorously detect an inference, should fulfill the following requirements:

Demand: A paradigm that measures a spontaneous process implicitly should direct participants' attention to the behavioral information without encouraging explicit trait impression goals (Uleman, 1999). To accomplish this, participants are usually instructed to familiarize themselves with the material (*e.g.*, Skowronski, Carlston, Mae, & Crawford, 1998; Crawford, Skowronski, Stiff, & Scherer, 2007; Crawford et al., 2008) or to memorize it (*e.g.*, Winter & Uleman, 1984; Newman, 1993; Todorov & Uleman, 2002; Goren & Todorov, 2009). However, even if the instruction is to memorize the material, participants might still infer traits in a strategic way in order to improve their performance in the memory task, and even more so when they are aware that the inference can actually lead to the correct response. This is especially true when all the materials used are trait-implying descriptions making it easier for the participants to guess the researcher's goal of studying personality impressions. Ideally, the trait-implying material should be inserted into a larger set of neutral sentences. In sum, this requirement concerns the type of instruction given to participants and the type of material used to disguise the scientific purpose of the experiment.

Contamination: The paradigm should discourage strategies likely to induce inference at the moment of retrieval and uniquely measure inference at encoding (Keenan, Potts, Golding, & Jennings, 1990; McKoon & Ratcliff, 1986; Wyer & Srull, 1986). In other words, the occurrence of an inference should be exclusively based on measuring unconscious recollection of earlier experience, as opposed to explicit recollection (Schacter, 1987) that might trigger inference at the moment of recollection. This is also known as the contamination problem of implicit measures by explicit recall (Jacoby, 1991). Implicit measures are said to be contaminated when the participants adopt an intentional retrieval strategy or if they are aware that the tested material is related to the studied material. To avoid contamination, researchers should refrain from using memory-based paradigms (*i.e.*, paradigms that require participants to explicitly recall past events).

Backward association: The paradigm and the material should be created in such a way that they do not allow for backward associations to take place (Wyer & Srull, 1986). Singer (1978, as cited in Keenan et al., 1990) illustrated the effect of backward association in a cued-recall experiment. In this experiment participants were first presented with sentences containing actions (*e.g.*, “stir the soup”) and later presented with two different cues to recall the sentence (*e.g.*, “spoon” or “ladle”). He showed that these cues were asymmetrically related to the action. In other words, the forward association between the action (*e.g.*, “stir the soup”) and the first cue (*e.g.*, “spoon”) is strong (leading to the inference of the instrument “spoon”) whereas the backward association is weak (meaning that the action is less likely to be generated when this cue is provided). The opposite pattern, stronger backward than forward association, was found for the second cue (*e.g.*, “ladle”). Thus, if “ladle” is used as cue, the activation that spreads from the cue (“ladle”) to the action (“stir the soup”), *i.e.*, backwards, will mimic an inference made at encoding. Note that this confounding factor is closely related with the contamination problem because the backward association is more likely to occur if the participants are required to retrieve the implying sentence. Although, backward association is not trivial to overcome, it can be addressed through the use of stimuli that have associative asymmetries, like the material used by Singer (1978), in order to test the paradigm sensitivity to backward associations.

Word-based priming: It should be guaranteed that the activation of the trait is a consequence of text-based priming and not word-based priming. In other words, the inference should result from processing the meaning of the behavior and not from processing specific words in the sentence that are individually associated with the trait (Keenan et al., 1990; Orghian, Ramos, & Garcia-Marques, n.d.). One possible way of avoiding this confound is by having control sentences that contain roughly the same words as the trait-implying ones but rearranged in such a way that they no longer imply the trait. This method equates the two types of sentences in terms of word-based priming.

Conceptual: It should be a task that is sensitive to top-down phenomena, meaning it should be able to detect concepts activated via meaning-based processing. Roediger and Blaxton (1987) named these kind of measures conceptually-driven tasks, as opposed to data-driven. The distinction between data-driven and conceptually-driven tasks concerns the distinction between tasks in which participants rely more on physical features of the stimulus, and those where the focus on the physical features is minimal and where conceptually-based and top-down processes are used (Roediger & Blaxton, 1987). This perspective assumes that any task can rely on both types of processes, meaning any task can be situated on the continuum between data-driven processes on one pole and conceptually-driven processes on the other. This view emphasizes the processes operating in each task by claiming that there are “benefits to the extent that the operations required at the test recapitulate or overlap the encoding operations during prior learning” (Roediger, Weldon, & Challis, 1989, p. 16). Applying this logic to trait inferences, it is clear that the process taking place at encoding is conceptually-driven since the meaning of a whole sentence has to be comprehended in order for the trait to be inferred. Moreover, no perceptual features can be processed because the trait is only implied and not actually presented at encoding. Thus, a conceptual task that only minimally depends of the perceptual features of the target will be more sensitive to such high-level process.

In the context of STI, we use the term data-driven, not only to refer to tasks that depend on perceptual features of the target, but also to tasks in which the performance depends on the activation of specific labels. Thus, if a researcher expects the participant to activate a certain label/trait and presents that trait for a recognition test, the detection of the inference is restricted to the activation of that specific trait. If we assume that an inference is the activation of a semantic area that can be represented by multiple traits (*e.g.*, synonyms), we should then equate the possibility of mislabeling, that is, that the trait activated via the inference made by the participant may not be the same as the trait used by the researcher at test. The label given at test has an important role because it will guide the retrieval process in a data driven or bottom-up manner. The ultimate solution for this problem is the free recall task, because no label is provided and recall happens in a top-down manner that is idiosyncratic to each participant. However, since the target word is only implied by the material given at study, there is no easy way to elicit the recall of something that was not directly presented and that the participant is not even aware of.

Delayed: It should be possible to apply the measure in delayed mode (as opposed to immediate mode), meaning the inference should be measurable even after a period of delay from the moment the sentence is read. This will allow detection of inferences that occur later on in the integration process (Till, Mross, & Kintsch, 1988; Calvo & Castillo, 1998; McKoon & Ratcliff, 1989a). This is important because we don't only want to know if the trait is activated at that specific moment in time, we need to understand if the inferred trait is integrated in the final representation of the encoded event and if it is stable over time.

In the following section we will analyze in more details the existing paradigms (see Table 1 for a summary) in the STI field. Afterwards, we will present a new paradigm and discuss its advantages compared to the paradigms already used in the field.

Paradigms and their limitations

Inferences are traditionally measured with memory or activation tasks (for a review see Keenan et al., 1990). In memory measures, participants are asked to retrieve a past event or to state whether or not a target word featured in some earlier text. Activation measures, on the contrary, do not rely on retrieval of previously presented material but on measuring the activation of the inferred concept with implicit tasks such the lexical decision task.

Memory-based measures

The cued-recall paradigm was the first measure used in STI literature and for that reason it has been one of the most recurrently used (Winter & Uleman, 1984; Winter et al., 1985; Claeys, 1990; Uleman, Moskowitz, Roman, & Rhee, 1993; Uleman, Winborne, Winter, & Shechter, 1986; Uleman & Moskowitz, 1994; Bassili & Smith, 1986). Winter and Uleman (1984) were the first to show that dispositional cues implied by (but not presented with) trait-implying sentences were more effective recall cues for those sentences than semantic associates of words presented in those sentences. For example, the recall of the sentence “The librarian carries the old woman’s groceries across the street.” was better cued by the trait word “helpful” than by the semantic cue “books” or by no cue at all. Cued-recall is a delayed measure because the measurement (the cued-recall phase) is only conducted after all the sentences are memorized, as opposed to trial by trial. The greater efficacy of the trait cue is interpreted as evidence that participants had spontaneously inferred the trait upon reading the sentence.

Yet, some researchers questioned the interpretation that traits are inferred during the reading of the sentence in the cued-recall paradigm. They argue that the trait word given at test can prompt the recall of the behavior, through backward association from the trait to the behavior. Presenting the trait during the memory test can trigger the generation of typical behavioral exemplars that, in turn, can cue the recall of the correct behavior (*e.g.*, Corbett & Dosher, 1978; Srull & Wyer, 1989). This happens because the participants are explicitly asked to retrieve the sentences. Moreover, while trying to retrieve the sentences the participants can generate the inferences (or realize that the cues and the sentences are compatible) at that moment in time, a result of the contamination problem mentioned before. Another limitation of this paradigm is the lack of control for the word-based priming (Keenan et al., 1990), that is, the trait can become activated via the presence of very specific words that are part of the sentence without any consideration for the meaning of the sentence as a whole. Moreover, the sentences are usually all trait-implying, meaning the participants can easily understand the relevance of trait inference in the context of the task. Finally, some have argued that the results in this paradigm might reflect part-list cueing effects since they rely on a within-participant contrast between cued and non-cued recall (D'Agostino & Beegle, 1996). This is solved by contrasting the recall with different types of cues and ignoring the non-cue condition.

The recognition probe paradigm is another memory measure (Ham & Vonk, 2003; Newman, 1991; Uleman, Hon, et al., 1996; Van Overwalle, Drenth, & Marsman, 1999; Wigboldus, Dijksterhuis, & Van Knippenberg, 2003; Wigboldus et al., 2004) whose use was imported to social cognition (Uleman, Hon, et al., 1996) from discourse comprehension literature (*e.g.*, McKoon & Ratcliff, 1986). In this paradigm, participants are presented with trait-implying sentences under memory instructions and immediately after reading each sentence a probe word is presented. In the critical trials, the probe word is the trait implied in the sentence and the instruction is to indicate whether or not the word was part of the sentence. The logic of the task is that if the trait is inferred from the sentence it will be more activated and thus, it will be more difficult (demonstrated by longer reaction times) for the participant to state that the probe did not occur in the sentence and/or the responses will be less accurate. Contrary to cued-recall, that is a delayed measure, probe recognition is usually applied in an immediate fashion. However, it has been occasionally applied in a delayed mode (see for example Van Overwalle et al., 1999; McKoon & Ratcliff, 1986).

The recognition probe paradigm, when used with the appropriate control sentences solves the limitation of word-based priming. Uleman and colleagues (1996) used control sentences that were rearrangements of the words in the trait-implying sentences but no longer implied the traits. This ensures that the two sentences (control and trait-implying) are equal in terms of activation coming from specific words, since the words used in each are closely matched. If any difference is found between the two versions (say longer reaction times to reject the trait in the trait-implying condition), it must result from processing the meaning of the sentence, or, in other words, from inference making present in the trait-implying sentence and absent in the rearranged version. However, the use of these types of controls is not the norm but an exception.

Also, contrary to cued-recall, in this task the problem of participants inferring the trait as a strategy to improve memory is solved because inference making leads to worse, rather than better, performance. The main limitation of this paradigm is that we don't know whether the inference is drawn during encoding or during test, due to the contamination problem. The task requires the participant to check the probe against the sentence. While checking the probe against the sentence, the participant can be led to make the inference on the spot. The inference being drawn at the test moment can lead to slower responses, exactly as would be expected if the trait had been inferred at encoding. Sometimes, a deadline procedure is applied in order to deal with contamination (McKoon & Ratcliff, 1986). It is assumed that contamination is less likely to occur and the responses are more automatic if participants are compelled to answer within a brief window time.

Carlston and collaborators (Carlston & Skowronski, 1994; Carlston et al., 1995; Carlston & Skowronski, 2005) developed a different memory-based paradigm called saving in relearning. In this paradigm, participants are instructed to familiarize themselves with trait-implying paragraphs describing behaviors paired with photos representing the actor of those behaviors. After a distraction task, they have to memorize photos of people paired with personality traits. Some of the traits were implied in the paragraphs presented with those actors in the first phase, whereas others are new. The argument is as follows: if traits are inferred about the actors during the encoding of the paragraphs, then pairing the same photo with the implied trait would reinstate the initial inference which will lead to faster learning of these pairs than learning new pairs (a result called saving in relearning). In the test phase the photo is presented again, this time as a cue to recall the trait. Recall is expected to be better in the relearning trials than in the newly paired trials. This paradigm investigates, in an elegant way, not only the inference of the trait from the paragraph but also the link between the actor and the trait (for more see Carlston et al., 1995).

In saving in relearning paradigm, it is also difficult to conclude that the trait is inferred during the reading of the behavior. In the relearning phase, when the actor and the trait are presented for memorization, the participant might strategically retrieve the behavior in order to improve their memorization of the trait and thus draw the inference during this retrieval. Note however, that the authors of this paradigm argued against contamination by showing that the effects remained even when trials were dropped if the participants recalled the corresponding behaviors, as well as following a delay of up to 7 days between the study and the test (Carlston & Skowronski, 1994). This paradigm is always applied without effective control for word-based priming. Additionally, the use of paragraphs, as opposed to singles sentences, is not optimal because it was shown that pairing multiple trait-implying sentences with faces is enough to trigger explicit impression formation goals (Park, 1989), undermining in this way the spontaneity assumption.

Finally, false recognition is a memory-based paradigm developed by Todorov and Uleman (2002, 2003, 2004). In this paradigm, participants are first exposed to pairs of photos of people alongside sentences describing their behavior, under memorization instruction. After being exposed to this material, they are presented again with the same photos paired, this time, with traits. The task here is to indicate whether the trait word occurred within the sentence paired with the photo in the first phase of the study (as some of the sentences contained the actual traits within the descriptive text). Higher false recognition rate is expected if the trait was inferred from the sentence in the initial phase. This happens because, once the trait is inferred, it is part of the representation of the event and, as such, it is difficult for the participant to distinguish what was actually presented from what was inferred.

In the false recognition paradigm, the participant is also required to contrast the trait against the sentence memorized in the first phase, which means reconstructive processes can occur. When the participant tries to remember the behavior, the inference can be drawn. Moreover, no control sentences are used in this paradigm, only trait-implying sentences and trait-implying sentences that actually contain the traits, making it very prone to the “demand” limitation. The false recognition paradigm also fails to control for word-based priming because rearranged sentences are not used. Also, this paradigm has an interesting characteristic; because trait inference leads to false recognition, that is, to a poorer performance, the incentive for inference making as part of a reconstructive process is low. Both, the saving in relearning and the false recognition are delayed measures of inferences.

Activation measures

Lexical decision is an activation measure that gives the researcher access to the activated trait without explicitly asking the participant to retrieve the behavioral description. Zárate, Uleman, and Voils (2001), in order to avoid the reconstructive process criticism associated with memory measures, used lexical decision to access the inference during encoding. In this paradigm, participants are presented with strings of letters and their task is to indicate as quickly and accurately as possible whether or not the letter string is a real word. Zárate and colleagues presented two types of trials to their participants. One where the target given for lexical decision was a trait, and it was preceded by a behavioral description that implied that trait, and one where the trait was preceded by a control sentence that was unrelated to the trait. A facilitation was detected for implied traits compared to the traits unrelated to the preceding sentences.

One limitation of the lexical decision paradigm is that the type of control does not control for word-based priming. There is also some evidence that the relationship between the sentence and the trait affects the lexical decision latencies, suggesting that the two are contrasted one against each other (Balota & Chumbley, 1984; West & Stanovich, 1982; Neely, Keefe, & Ross, 1989). More precisely, these studies showed that the backward association between the target and a prime affects lexical decision latencies. And, as mentioned before, a good task should be sensitive to the prime-target associations (forward inference) but should not be sensitive to target-prime associations (backward associations). This is even more important when the measure is applied in an immediate mode (as in Zárate et al., 2001) because any relationship between the target and the prime is easier to access. However, this measure can easily be applied in a delayed mode (as done for example by Potts, Keenan, & Golding, 1988, in the text comprehension literature).

The word stem completion task is another activation measure used to study STI. Whitney and Williams - Whitney (1990) used constrained word stems to detect online (*i.e.*, drawn at encoding) inferences. Participants in this paradigm start by reading trait-implying or control paragraphs and then complete word stems (*e.g.*, F R _ _ _ _ _ _ for FRIENDLY) with the first word coming to their mind that would fit into the blanks. If the trait is activated by the trait-implying sentence then the trait will be used to complete the stem more often than it would be in the control condition (*e.g.*, Whitney & Williams-Whitney, 1990; Bassili, Smith, & MacLeod, 1989).

This word stem completion task is usually applied with effective control for word-based priming, *i.e.*, with rearranged versions of the trait-implying paragraphs (Whitney & Williams-Whitney, 1990). Moreover, this measure was shown to be unaffected by backward associations (probably because the target word is not presented in its complete format) and thus, seems to be immune to retrieval processes (for more see Whitney, Waring, & Zingmark, 1992). For these reasons this task can be considered a better measure than the lexical decision task. However, there is one important limitation we would like to point out to this paradigm that also applies to all the previous paradigms. The answer the participant is expected to provide is used as a measure of the activation of a specific trait that the sentence, previously tested, is supposed to imply. However, this is based on a questionable assumption that all the participants infer the expected pre-tested trait. For example, in Whitney and colleagues' pilot study (1992) participants were asked to generate traits for behavioral descriptions. The traits used as targets in the stem completion task were generated by the participants in the pilot study on average 54% of the time (with the least consensual being generated 35% and the most consensual 90%). This clearly shows that not all the participants infer the same traits, and this applies to all the paradigms previously discussed. However, the same target traits (or stems) are presented for all the participants at test. When a specific trait is used in the test phase, the effect will only be observed if that specific word has been lexically activated. If at encoding, rather than the pre-tested trait a synonym was activated, the expected lexical facilitation will not be observed. If this happens, it will appear that no inference occurred, when in fact the representation of the target concept was activated. In a more extreme case, the activated representation might not even overlap with the meaning of the target expected to be given in the stem completion task.

In other words, activation measures like the lexical decision and the word stem completion and memory measures like the false recognition and the probe recognition are data-driven tasks, as opposed to conceptually-driven tasks, in the sense that they are dependent on the activation of specific lexical formats. In the STI literature the paradigms are never analyzed from this perspective (data-drive versus conceptually-driven) but that is not true in implicit memory literature. For example, the word stem completion and the lexical decision tasks were characterized as being far closer to the data-driven pole of the continuum than to the conceptual pole (*e.g.*, Hamann & Squire, 1996; Roediger, Weldon, Stadler, & Riegler, 1992). Some memory measures like the free recall can be considered conceptual tests, because the participant directly accesses their representation in memory, rather than being given a trait that then drives the retrieval. However, none of the memory measures presented before are conceptual because the to-be-inferred trait is provided to the participant at some moment during the experiment.

The paradigm that seems less affected by this dependency on specific lexical formats is the saving in relearning. In this paradigm, in the test phase only the photo of the actor is presented as cue and the participant is instructed to recall the trait that was presented with that actor in the relearning phase. The effect is still very dependent of that specific trait, but the previous presentation of the trait in the relearning phase pre-activates that trait, meaning that any competition between this form and similar ones was solved before the critical measure is applied. However, as with all the memory measures presented before, saving in relearning might be contaminated by explicit retrieval. Activation measures solve, in part, the contamination problem, but they rely heavily on data-driven processes, which, as argued before, is not a good fit for a top-down mechanism like STI.

A conceptually-driven measure that is not based on memory would solve both the contamination problem and the sensitivity to conceptual processing requirement. In the following section we present a new paradigm to measure STI that solves both problems.

Modified Free Association Paradigm– MFAP

Hourihan and MacLeod's (2007) modified word association paradigm is a conceptually-driven measure of implicit memory that seems to overcome most of the limitations of the paradigms presented above. In their study, in the learning phase, participants either generated words from meaningful cues (*e.g.*, "the piece of furniture used for sitting –c?") or merely read the words (*e.g.*, "chair"). In a test phase, participants performed a word association task where they had to speak aloud the first word to come to mind upon sight of a prompt word that was either generated or read in the learning phase. Generating a word at encoding primes that word in implicit memory and this priming, or activation, extends to neighboring words in the semantic network. Thus, when the prompt is a word that was generated by the participant in the past, as opposed to a word that was just read, the generated word is faster to access and so are other words within its semantic network (Nelson & Goodmon, 2002). Translated to trait inference research, the premise is that reading a trait-implying sentence primes the implied trait word and its semantic neighbors. Thus, when the inferred trait is encountered in the free association task, this delivers faster production of an associate compared to un-primed traits.

Hourihan and MacLeod presented the paradigm as a conceptual implicit memory test that provides an isolated measure of conceptual priming free from the influence of explicit retrieval. They claim that if there is nothing to retrieve from the study episode, explicit retrieval is not a problem. They also argue that if a concept was semantically processed (as we think the traits are when inferred), its semantic characteristics should be activated above a baseline, and some activation should spread to semantically associated neighbors (*e.g.*, activation of the trait "friendly" might lead to the activation of the trait "nice"). The activation of the semantic network of a trait is detected via the latencies to generate an associate. Thus, the more the semantic network of the trait was activated during the encoding of the behavior, the faster and easier it will be for the participant to generate an associate.

The word association task draws upon two mechanisms: the lexical access, which reflect access to an associate in the semantic network of the trait and the articulation of the associate. The dependent variable in this paradigm is not the response itself but the speed of response. Moreover, research has demonstrated that reaction time is a valuable measure of conceptual implicit memory (*e.g.*, Light, Prull, & Kennison, 2000). This is the first paradigm that allows us to measure the speed to access the semantic network of an inference. The idea of spreading activation through the network is, furthermore, in agreement with the halo effect usually found in the STI literature. It was shown that people do not only infer a unique trait from the behavior, but that they generalize to other close traits and use them to make judgements about the actors performing the behaviors (Carlston & Skowronski, 2005; Crawford, Skowronski, & Stiff, 2007; Crawford, Skowronski, Stiff, & Scherer, 2007; Skowronski et al., 1998).

Similar to the lexical decision task, in the test phase, the to-be-inferred trait is provided to the participant, but whereas in the lexical decision task response accuracy is key measure, here accuracy is irrelevant because there is no correct answer. The lack of a correct answer means that concerns about memorization processes, and consequently contamination by explicit retrieval, can be discounted. Even though the instruction, in the study phase, is to memorize (for more see the Procedure section in experiment 1) the material in order to avoid activating impression formation goals, this is an activation measure as opposed to memory measure because at test the participant is not required to contrast the trait with the sentence. This makes this task less prone to explicit retrieval strategies.

We cannot, nevertheless, totally exclude the possibility that backward associations from the trait to the sentence do not affect the access to the semantic network. However, as we show later, the inference is also detected in a delayed free association test (experiments 2 and 3). In a delayed test, the backward association is less likely to occur because the participant no longer has immediate access to the sentence.

Next, we describe two more paradigms: the naming task and the modified Stroop task. Even though these tasks are not traditionally used in the STI literature, we decided to approach them because they are seen by some researchers as good measures of inference (for a review see Keenan et al., 1990). We compare these two measures with the MFAP in our second experiment.

Naming

In the naming task, after reading inference or control sentences, participants are presented with words and instructed to read the words out loud as quickly as possible. The naming task can be decomposed into two processing steps, lexical access and articulation. And because these steps take place so quickly, it is difficult to imagine how the relatedness between the target and the text, *i.e.*, the backward association, can have an influence. But, for the same reason, it is difficult to imagine that it can be sensitive to higher level inferences such as text-based inferences (Norris, 1986). Thus, whereas there are some studies showing that, unlike the lexical decision task, naming is not affected by backward association (Seidenberg, Waters, Sanders, & Langer, 1984), there are authors who question its ability to detect more elaborated inferences (Keenan et al., 1990). Potts and colleagues (1988) detected inferences within the naming task. However, they admit that this measure is less sensitive than others, which is in agreement with the small effects they reported (between 11 and 17 ms facilitation). A second limitation of the naming task is the possibility that it can be completed without any access to the lexicon, via grapheme to phoneme translation rules (Keenan et al., 1990) which is clearly a data-driven mechanism. Also, note that just like in the stem completion and the lexical decision tasks, this effect depends on the activation of a very specific lexical format (the word presented at test), and so, the inference of other similar traits (*e.g.*, synonyms) might be difficult to capture with this measure.

Modified Stroop task

The modified Stroop task is an activation measure also popular in the text comprehension research. In this task participants usually read a text and then name the ink color in which the test words are written (Conrad, 1974; Dosher & Corbett, 1982;

Whitney & Kellas, 1984; Whitney, 1986). The rationale of the task is as follows: if the trait is inferred from the inferential sentence and presented for ink color naming, the participant will take longer to name the color of the inferred word than he would take to name the ink color of an unprimed word. This happens because, if the word was previously activated, it is more difficult to suppress its articulation in order to say the ink color.

The task reflects three sub-steps: lexical access, color identification, and color articulation. It is implausible that the relatedness between the target and the text is affecting color identification and articulation, meaning this measure may be a purer measure of lexical access. Keenan and colleagues argued that it might be a better measure than the naming task, because the latencies in the modified Stroop task are longer, avoiding possible floor effects (Keenan et al., 1990). These same researchers claimed that this measure can distinguish between whether people process the inference at conceptual or lexical level. If the word is primed at a conceptual level, it is more difficult to name the color ink. An example of conceptual priming is the target "doctor" preceded by a text about "hospitals", as opposed to a text about "subways". However, when the word is primed at lexical level, via explicitly presenting the target word "doctor", then facilitation is expected for ink naming (Dosher & Corbett, 1982; Whitney, 1986). The modified Stroop, like all the previous measures, is a data-driven task. In other words, the answers in this task are dependent on the activation of very specific traits.

When the word given at the test is different from the one inferred, the articulation of the word at test is not supposed to interfere with color ink naming because that specific lexical format was not activated before. In MFAP, if the trait provided at test is the one inferred by the participant, then facilitation (in terms of the latencies to generate associates) is expected at both lexical level (*e.g.*, "helpful" inferred and "help" generated as an associate) and conceptual level because the trait is part of a larger activated network ("helpful" inferred and "concerned" or "aid" generated). If the trait inferred is different from the one provided at the test, any facilitation that is eventually observed cannot be a consequence of that specific lexical format, it has to be due to conceptual priming. In other words, the facilitation happens due to the fact that the activated concept is part of a larger activated network of interconnected concepts.

If sensitive to trait inference, MFAP will be a fruitful contribution to trait inference research as it overcomes some of the limitations of the existing paradigms. First, to avoid explicit impression formation goals, the material, in this task, is presented under memorization instruction for a later unspecified test and it is embedded in a set of neutral sentences. Second, the memorization phase and the free association phase are presented as separate tasks, making is more difficult for the participant to be aware of the relation between the two. Moreover, because there is no need to explicitly retrieve past information, contamination with explicit processes is less likely to occur. Third, it is possible to control for word-based priming by including rearranged versions of the trait-implying material. Fourth, it is a conceptually-driven task, meaning the operations happening at test recapitulate, to some extent, the top-down operations taking place at encoding. And finally, it is easy to apply MFAP in delayed fashion, so that the participants read all the behavioral descriptions first and are tested in a separate phase. Furthermore, in MFAP participants remain unaware of what is being tested as the task has no explicit success criteria, and indeed the word content produced is irrelevant.

In three experiments, we test MFAP and its sensitivity to STI. In experiment 1, we predicted that the reaction time to generate an associate for a trait prompt will be faster when following trait-implying than control sentences. This is due to the fact that priming a trait word activates its semantic network and increases the accessibility of related words in implicit memory. In experiment 2, we compare the MFAP with the modified Stroop task and with the naming task. In experiment 3, we show that the MFAP can be also useful to study the relevance of the actor to the presented behavior and the link between the inferred trait and the actor. Moreover, we test the sensitivity of MFAP to differences between STI and Spontaneous Trait Transference (STT), a cognitive error that results in the transference of the inferred trait to a person who is not the actor of the behavior (Carlston & Skowronski, 2005; Crawford, Skowronski, & Stiff, 2007).

Experiment 1- MFAP

The aim of experiment 1 was to validate the use of MFAP in detecting the occurrence of STIs. Participants, in each trial, were presented with a sentence followed by a prompt word for free association. Upon sight of the prompt, they were asked to speak aloud the first word that came to their mind. In the critical trials, the sentence was either trait-implying (*e.g.*"He phoned for help while the others just screamed.") or a rearranged control (*e.g.*"He screamed for the others to help find the phone.") and the prompt was the trait word implied in the trait-implying sentence (the trait "calm"). Following Hourihan and MacLeod's (2007) rationale, we predicted that the trait-implying sentence will prime the semantic network of the trait word and lead to faster generation of an associate when compared to unprimed traits presented in the control condition.

Method

Participants. 39 undergraduate students from Birmingham University participated in this study for academic credits. 6 were males and the mean age of the sample was 19.46 (SD = 1.28) years old. In this experiment, the criteria used to decide the minimal sample size was based on the sample size reported by Hourihan and MacLeod (2007) in their first study (N = 30). The data, in this and in the following experiments, was only analyzed after the reported samples were complete.

Material. 24 pairs of sentences from Uleman, Hon, Roman and Moskowitz's (1996) studies were used in the current experiment. For each trait-implying sentence, there was a control sentence that had approximately the same words but rearranged in such a way that the trait was no longer implied.

Pre-test: To validate the material for the population in the United Kingdom, we conducted a pilot study where the material was adapted (American expressions changed to British) and given to participants for a rating task. 72 English native speakers living in the UK took part in this pilot study. Each of them was presented with 12 trait-implying sentences and 12 rearranged sentences. The two versions of the same sentence were never presented to the same participant. In each trial, the participant was provided with a trait and a sentence. The instruction was to indicate to what extent they believed the trait belonged to the person performing the action described in the sentence by using a slider (going from 0 - not at all to 100 - totally). Next, we calculated, individually for each trait, the average ratings for each trait-implying and rearranged sentences. The average difference between the ratings for the trait-implying and the rearranged sentences was 44.51, with the smallest difference being 22.46 (for the trait “ambitious”, for which the trait-implying sentence was “She held a full-time job while being a full-time student.” and the rearranged was “She couldn’t hold a full-time job while being a full-time student.”) and the largest difference being 70.67 (for the trait “sociable”, with the trait-implying sentence being “He liked parties more than films.” and the rearranged being “He liked films more than parties.”). The difference between the rearranged ($M = 37.77$, $SD = 20.35$) and the trait-implying sentences ($M = 82.29$, $SD = 6.28$) was significant, $t(23) = 10.64$, $p < .001$, and in the expected direction. All 24 pairs of sentences and the correspondent traits tested in this pilot were used in the experiment. 24 non pre-tested pairs of trait-implying and rearranged sentences were also included as filler (see the Procedure section for details). Additionally, we had 36 neutral sentences that were also used as fillers and were not meant to imply any specific traits (“She rented a summer place where she went the previous year.”). For the free association task, we had 24 critical traits (corresponding to the pairs of sentences from the pilot study; *e.g.*, “sociable”), 12 traits unrelated to any of the sentences (*e.g.*, “reliable”) and 48 non-trait words (*e.g.*, “drill”).

Procedure

Participants were tested individually in a sound proof lab on a desktop PC with 17-inch CRT monitor. Instructions and stimuli were presented in size 14 pt white font on a black screen. The software used to build the experiment was OpenSesame (Mathôt, Schreij, & Theeuwes, 2012), a multi-platform open source tool. A Sony plug-in headphones sounded the ‘bleep’ warning for the prompt in the free association task and a microphone recorded verbal reaction times digitally. The experimenter noted response content by hand. In the pilot study mentioned above, Qualtrics was used to conduct the study and Prolific platform to publicize it.

Participants were told that the experiment was an investigation about multitasking and about how people cope with the multiple requirements of the environment. More precisely, they were asked to memorize sentences while doing a concurrent task - performing associations with target words. The two tasks (sentence memorization and free association) were presented as unrelated. Participants were instructed to memorize the sentences in order to encourage attendance of the stimuli and avoid overt impression formation goals. After each sentence, a word was presented for the free association task and the instruction was to say the first word that comes to their mind upon sight of the prompt, as quickly as possible. In the free association task they were told that there was no wrong or right response. Before the study began, the experimenter helped the participants to fit the headset and microphone and instructed them to read aloud a series of words on the screen to optimize microphone recording threshold. Participants were instructed to speak clearly and to avoid clearing their throats or making any other noise that would invalidate their response. They started with four practice trials that were followed by an opportunity to ask questions.

Each trial began with a 500 ms fixation cross that was followed by the presentation of a sentence. The sentence remained on the screen for 4000 ms. Participants were instructed to read the sentence silently and memorize it. The sentence was replaced by a 500 ms blank screen, after which a bleep sound warned participants about the prompt was about to appear. Between the bleep and the prompt, a 40 ms blank screen was presented. The prompt remained on screen until a vocal response was detected. At this point, the word disappeared and a new trial began. In total there were 84 trials. The presentation order of the sentence-prompt pairs was randomized for each participant.

Because we did not want the two versions (trait-implying and the rearranged) corresponding to the same trait to be presented to the same participant we created two versions of the experiment. In version A, 12 traits were preceded by their trait-implying sentences and the other 12 by the rearranged/control sentences. In version B, the 12 traits that in version A were preceded by trait-implying sentences were preceded here by controls and the 12 that were preceded by controls in version A, in version B were preceded by trait-implying sentences. The assignment of the trait to one version or another was done randomly and participants were randomly assigned to one of the two versions.

Thus, in each version of the experiment there were 12 *critical trait-implying sentences* followed by the correspondent *implied traits* for free association, 12 *critical rearranged sentences* followed by the correspondent *traits* (implied in the trait-implying versions), 12 *non-critical trait-implying sentences* followed by *non-trait words*, 12 *non-critical rearranged sentences* followed by *non-trait words*. These last 24 non-critical trials were fillers and were presented in order to avoid a situation where the only times that trait words were presented for free association they were following trait-implying sentences. This would have made the study interest in trait inference easier for participants to detect. Moreover, in 24 trials *neutral sentences* were followed by *non-trait words*. Finally, in the last 12 fillers *neutral sentences* were followed by *traits*, so that not all the neutral sentences were followed by non-traits. All these fillers have the purpose of hiding the main relationship being investigated - the relation between the trait-implying sentences and the correspondent implied traits. This relationship was present in only 12 of the 84 trials.

During the free association task, the experimenter coded as invalid those trials that were affected by technical issues or those where participants' "lip popping" noises were recorded in error as their vocal response. After all trials were completed, to keep with the cover story regarding the memory test for the sentences, participants were given a memory test consisting of 20 pairs of similarly worded sentences. Their task was to identify which of the two versions they had seen earlier. The whole experiment lasted approximately 20 minutes, after which the participants were thanked and debriefed.

Results and Discussion

Next, we analyzed the reaction times (RTs) in the trials where the target words were traits. This included the trials where the prompts were preceded by trait-implying sentences, by rearranged control sentences and by neutral controls sentences. Trials where the prompts were non-trait words were not included in these analyses first, because we don't have any hypotheses for them and second, because any differences compared to the trait trials would be difficult to interpret since the two types of prompts are too different from each other. Trials where recording problems (*e.g.*, those flagged by the experimenter as affected by an undersensitive or oversensitive microphone) were detected were excluded, accounting for 5.4%.

Next, we applied an outlier elimination method based on median absolute deviation with a cutoff of 2 (MAD; Leys, Ley, Klein, Bernard, & Licata, 2013; Rousseeuw & Croux, 1993). As opposed to the standard deviation from mean criterion that is highly affected by the presence of outliers (those same outliers that one wants to eliminate), this method is less affected by outliers because it is based on the deviation from the median instead of the mean. By using this method, we eliminated 16.70% of the responses.

Due to the amount of missing data, a Little's MCAR test (Little, 1988) was conducted in order to verify if the missing values are missing completely at random. The Little's MCAR test obtained for this experiment data resulted in a $Chi-square = 644.41$, $df = 672$, $p = .772$, which means that no identifiable pattern exists to the missing data.

A mixed effects ANOVA was conducted. The type of sentence that preceded the trait (trait-implying, control-rearranged, control-neutral sentences) was the within-participant independent variable, the version of the material was the between-participant independent variable and the RT in the free association task was the dependent variable. All the p-values reported are two-tailed, except when replication results are reported (experiment 2).

We found a significant effect of type of sentence, $F(2, 74) = 5.50$, $p = .006$, $\eta_p^2 = .13$, no significant version effect, $F(2, 74) = .40$, $p = .529$, $\eta_p^2 = .01$, and no significant interaction between version and type of sentence, $F(2, 74) = .91$, $p = .407$, $\eta_p^2 = .02$. This means the version of the material did not lead to different patterns of results. As shown in figure 1, the reaction time to generate a word was significantly shorter when the preceding sentence implied the trait ($M = 1628.79$, $SD = 226.45$) than when it was a rearranged version ($M = 1715.29$, $SD = 277.28$), $t(38) = 2.27$,

$p = .029$, or when it was a neutral sentence ($M = 1745.17$, $SD = 279.42$), $t(38) = 3.10$, $p = .004$. The neutral condition did not vary from the rearranged one, $t(38) = .88$, $p = .385$. From this result, we can conclude that reading trait-implying sentences, as opposed to rearranged or neutral, leads to the activation of the traits and their semantic networks, which facilitates the generation of associates in the free association task.

Thus, as predicted, this measure seems to be sensitive to trait inference since a significant difference in the latencies to generate an associate between the trait-implying sentences and the two types of control sentences was found.

The data of the experiments reported in the paper can be downloaded here: osf.io/c5ju9.

Experiment 2

In the second experiment, we compared the efficacy of MFAP with other two activation measures used to detect inferences: the naming task and the modified Stroop task.

In experiment 1, the free association task was applied immediately after reading the sentence, but testing inference in delayed mode can give us access to inferences that need more time to develop. Some researchers think that stabilized inferences that become part of the representation of the event occur only during a later stage of information processing (Kintsch, 1988). Thus, in the current experiment we tested MFAP in a delayed fashion. In this experiment we also added a priming manipulation where the trait itself was sometimes presented as a subliminal prime before being presented as a target. The goal was to prime perceptual features of the word and thus facilitate the access to it. This is relevant, especially if the inference is only partially encoded as argued by minimalist hypothesis theorists (McKoon & Ratcliff, 1986, 1989b, 1990). In the minimalist hypothesis, it is claimed that an inferences that is not necessary for text coherence is only minimally or partially encoded. Thus, by providing a subliminal prime, the inference can be instantiated more firmly. This means stronger facilitation is expected with the prime than without it in both MFAP and naming task. In the modified Stroop task, the prime is expected to lead to opposite results. Since the actual presentation of the target activates the lexical format of the inference and as discussed in the introduction section, it should lead to facilitation instead of interference in ink color naming.

Method

Participants. 77 students from the University of Lisbon took part in this experiment in exchange for a 10 euro voucher. 27 were males and the sample's average age was 22.67 (SD = 3.41) years old. In this experiment, we used a larger sample size than in the previous experiment because naming is described in the literature as a less sensitive task and thus, a larger sample might be necessary to detect the effect.

Material. This experiment was conducted with participants that spoke Portuguese, so the material (sentences and traits) was pre-tested for this population. The critical material consisted of 36 pairs of sentences and the corresponding traits. 12 of these pairs were selected from Pre-test 1 and the other 24 from the Pre-test 2 (Orghian, Ramos, Reis, & Garcia-Marques, in press). The purpose of both pre-tests was to select pairs of sentences in which the rearranged sentence sounded as natural as the trait-implying sentence, but implied the trait less than the trait-implying one (or ideally did not imply the trait at all). Because this objective was not accomplished in the first pre-test, the second one was conducted in order to create new sentences.

Pre-test 1: in this pre-test 48 pairs of rearranged and trait-implying sentences were used. Each *trait-implying sentence* was presented to 110 participants and each *rearranged sentence* was presented to other 41 participants. The participants were instructed to read the sentences and indicate the three personality traits that best described the person performing the behavior in the sentence. Next, the traits generated in the first position were analyzed. After detecting the most frequently generated trait for each trait-implying sentence, we calculated how frequently the same trait was generated for the rearranged version. The traits that were most frequently generated in the trait-implying sentences (M = .56, SD = .15) were significantly less generated in the rearranged versions (M = .16, SD = .19), $t(47) = 13.74$, $p < .01$. From the 48 pairs, 12 pairs of sentences were selected to be further used in the experiment. The criteria for this selected was a subjective evaluation of the framing of the rearranged sentences, performed by the first author. In order to compare the generation of the critical trait in the selected and in the non-selected sentences, a mixed effects ANOVA was conducted, where the within-item factor was the sentence (trait-implying or rearranged) and the between-item factor was being selected or not for the experiment. As expected, and in agreement with the t-test presented above, a main effect of sentences was found, $F(1, 46) = 127.90$, $p < .001$, $\eta_p^2 = .74$. However, no significant interaction between the sentences and the selection for the experiment was found, $F(1, 46) = 2.20$, $p = .15$, $\eta_p^2 = .05$.

Pre-test 2: the second pilot was conducted with the purpose of obtaining trait-implying sentences illustrative of a large set of personality traits (223). 293 participants took part in this pre-test. Each participant was presented with 15 personality traits randomly chosen from this set and their task was to generate a typical behavior for each trait. Participants were instructed to think about people they knew and concrete behaviors. They were also told to avoid using adjectives, and to be as specific as possible in their behavioral descriptions. Two independent judges each received half of the presented traits and the correspondent behavioral descriptions generated by the participants in the pre-test. After eliminating answers that were not behavioral descriptions and redundancies, the judges selected 2 to 3 behavioral descriptions that better illustrated each trait. Traits with similar behavioral descriptions, usually synonym traits, were grouped under the same trait label. This resulted in 154 pairs of trait-behavioral descriptions. Two new judges received these pairs and were instructed to select the best behavioral description, that is, the sentence that most strongly implied the trait. Moreover, the same two judges created rearranged versions for each of the 154 trait-implying sentences. An effort was made to use as many words from the trait-implying sentences in their rearranged versions as possible. The newly created rearranged sentences were then presented to four new judges. The first two judges (group A) were asked to indicate the first word that came to their mind when reading the sentences and the other two (group B) were instructed to evaluate to what extent the sentence was related to the trait using a 9-point scale ranging from not related (1) to very related (9). If at least one of the judges from group A generated the target trait for the control sentence, that pair of sentences (rearranged/implying) was excluded from the final set. Moreover, if both judges from group B rated the relation between the rearranged sentence and the trait with a value higher than 5, that pair of sentences was also excluded from the set. This filtering left us with 124 pairs of trait implying/rearranged sentences, from which 24 were used in the current experiment. The pairs of sentences in which the rearranged version sounded more natural to the authors were selected. In order to compare the selected and the non-selected sentences, the average of the ratings of the two judges in group B was calculated for each rearranged sentence. The ratings for the selected sentences ($M = 2.94$, $SD = 1.25$) were not significantly different from the non-selected ones ($M = 2.80$, $SD = 1.37$), $t(122) = .465$, $p = .463$. Note that for both selected and non-selected sets the average ratings are low because they correspond to the rearranged sentences, which are not expected to be related with the critical traits.

Aside from the 36 critical pairs of trait-implying and rearranged sentences, we also used 36 neutral sentences. Some of the neutral sentences were not even behavioral descriptions (*e.g.*, "In the winter the days are shorter"). In addition to the 36 traits that corresponded to the critical sentences, we also utilized a set of 36 non-trait words.

Procedure

Participants were told they were going to take part in four unrelated studies. All participants started with a sentence-memory task that corresponded to the learning phase during which the behavioral information was presented. The following three tasks that corresponded to the three test phases, were counterbalanced in terms of the order they were performed in. The three test phases were the naming task (reading words), the modified free association task (saying the first word that came to mind upon the sight of provided words) and the modified Stroop task (saying the color ink in which words were written). The cover story for the sentence learning phase was that we wanted to investigate *long-term memory* of complex material and that is why they were participating in 3 other studies before being tested for that material. Note that contrary to experiment 1, in experiment 2 the measurements are delayed, meaning the measurements do not happen immediately after reading each sentence, but during a later phase after memorizing all the sentences. The participants were also randomly assigned to the prime or no-prime condition. The only difference between the two is that in the prime condition, the target words presented for naming, free association or modified Stroop tasks were also presented as subliminal primes prior to the presentation of the sentence in the learning phase. This was designed to strengthen the inferential process in the critical trials.

The sentence learning task started with six practice trials that were followed by 72 experimental sentences. From the full set of 36 sentences, 18 critical trait-implying sentences (and their correspondent traits) were randomly selected for each participant.

The remaining 18 critical pairs and the corresponding traits were used in the rearranged condition. This prevented the same participant from seeing the two versions (the trait-implying *and* the rearranged sentences) of the same trait. Thus, in this phase, each participant memorized 18 trait-implying sentences, 18 rearranged sentences, and 36 neutral sentences that did not imply traits.

In the no-prime condition, each trial started with 800 ms fixation dot, followed by the presentation of the sentence for 3500 ms and ended with a 500 ms blank screen. In the prime condition, before the presentation of the sentence, and after the fixation dot, a 150 ms mask preceded the prime which was presented for 16 ms. The prime was followed by another mask that was presented for another 16 ms. Next the sentence was presented for 3500 ms, followed by the 500 ms blank screen. 37 participants took part in the prime condition and the remaining 40 took part in the non-prime condition. After the sentence learning task, the naming, the free association and the Stroop tasks followed. The procedures in the three tasks were identical in the prime and in the no-prime condition (this manipulation takes place in the learning phase). All the three tasks started with six practice trials that were also used to adjust the microphone to the participant's voice. In the naming task, participants were told that the goal of the task was to study reading skills. Participants were instructed to read as fast as possible the words presented on the screen. Before the word appeared, a fixation dot was presented for 800 ms and a 40 ms blank screen followed. The word was presented next and stayed on the screen until the microphone detected a response. The experimenter took notes about the accuracy of the responses. After the response was detected, and before the following trial, a 500 ms blank screen was displayed. This procedure was identical in the free association and the modified Stroop task except for the instructions and the fact that the words in the Stroop task were written in blue, yellow, green or red ink. In the association task, the participants were told that the goal of the task was to investigate the speed with which people generate new words. And thus, when the words appeared on the screen they were instructed to say as fast as possible the first word that came to their mind after reading the prompt. They were also told that there was no right or wrong response and to say any word that came to their mind regardless of how ridiculous it may sound. In the modified Stroop task, the instruction was to say as fast and accurate as possible the color ink in which the word was written. And the cover story for the Stroop was that we were interested in investigating color naming abilities. All the participants did the three tasks, and from the 18 traits that were implied in the 18 trait-implying sentences memorized in the first part of the study, six were presented in the naming task, six in the free association task and six in the modified Stroop task. A further six traits that corresponded to the rearranged sentences learned in the first phase, were presented in each of the tasks. In each task, besides the 12 traits (six in the rearranged and six in the trait-implying condition), there were 24 non-trait words unrelated to any of the sentences presented before. After participants finished the three tasks, a memory test regarding the sentences was performed, just to remain consistent to the initial cover story. This test consisted of 16 pairs of similarly worded sentences and the task was to identify which of the two versions was seen earlier. The presentation order of the trials in each of the five phases (the learning phase, the three test phases and the memory test for the sentences) was randomized for each participant individually.

Results and Discussion

After eliminating trials affected by recording problems and incorrect responses (5.1%) from the critical trials (trait prompts), the same outlier removal method used in experiment 1 was applied (9.6% of responses were eliminated based on this criteria).

A mixed effects ANOVA was conducted, with the presence of the prime (present or absent) being the between-participant independent variable, the task (free association, Stroop or naming) being the first within-participant independent variable, the type of sentence (trait-implying or rearranged) being the second within-participant independent variable and the RT being the dependent variable.

Besides a main effect of task, $F(2, 144) = 860.60$, $p < .001$, $\eta_p^2 = .92$, and a main effect of type of sentence, $F(2, 144) = 10.09$, $p = .002$, $\eta_p^2 = .12$, an interaction between the task and the type of sentence was found, $F(2, 144) = 11.50$, $p = .001$, $\eta_p^2 = .14$. As shown in figure 2 the participants performed differently in the three tasks. Moreover, there was no significant prime effect, $F(2,144) = 2.56$, $p = .11$, $\eta_p^2 = .03$. The interaction between the task and the prime, $F(2,144) = 1.59$, $p = .207$, $\eta_p^2 = .02$, and the interaction between the type of sentence and the prime, $F(2,144) = 1.93$, $p = .170$, $\eta_p^2 = .03$, did not reach significance. A three-way interaction was also not detected, $F(2,144) = .44$, $p = .648$, $\eta_p^2 = .01$.

To better understand the results underlying the interaction between the task and the type of sentence, pairwise comparisons were performed. They revealed a strong inference in the free association task, with larger RTs in the rearranged condition ($M = 1887.76$, $SD = 451.17$) than in the trait-implying condition ($M = 1721.54$, $SD = 377.41$), $F(2,144) = 11.28$, $p = .001$, $\eta_p^2 = .14$. No such difference was found in the modified Stroop task, $F(2,144) = .08$, $p = .772$, $\eta_p^2 = .001$ (trait-implying: $M = 671.07$, $SD = 71.18$; rearranged: $M = 669.24$, $SD = 72.93$), or in the naming task, $F(2,144) = .44$, $p = .509$, $\eta_p^2 = .01$ (trait-implying: $M = 581.51$, $SD = 61.79$; rearranged: $M = 578.52$, $SD = 59.54$). The result obtained in the free association task replicated the results from our first experiment.

Even though the three-way interaction was not significant, we conducted separate pairwise comparisons for the prime and no-prime conditions. In the free association task, an inference was observed, $F(2,144) = 8.90$, $p = .004$, $\eta_p^2 = .11$ (trait-implying: $M = 1641.14$, $SD = 359.24$; rearranged: $M = 1852.66$, $SD = 490.19$) when the prime was present and when prime was absent, $F(2,144) = 3.07$, $p = .042$ (one-tailed), $\eta_p^2 = .04$ (trait-implying: $M = 1801.94$, $SD = 393.57$; rearranged: $M = 1922.86$, $SD = 410.52$). In the Stroop task, no inference was found in the prime condition, $F(2,144) = 1.84$, $p = .179$, $\eta_p^2 = .03$ (trait-implying: $M = 646.60$, $SD = 58.62$; rearranged: $M = 658.84$, $SD = 60.68$) and a marginal interference was detected in the condition without prime, $F(2,144) = 3.27$, $p = .075$, $\eta_p^2 = .04$ (trait-implying: $M = 695.54$, $SD = 81.24$; rearranged: $M = 679.64$, $SD = 82.83$). In the naming task, no inference was detected in the prime condition, $F(2,144) = .1.59$, $p = .212$, $\eta_p^2 = .02$ (trait-implying: $M = 575.30$, $SD = 47.21$; rearranged: $M = 583.42$, $SD = 55.79$), but an opposite of an inference effect was detected in the condition without the prime, $F(2, 144) = 5.05$, $p = .028$, $\eta_p^2 = .07$ (trait-implying: $M = 587.73$, $SD = 72.91$; rearranged: $M = 573.61$, $SD = 62.84$). We don't have an explanation for this inverted pattern of results.

The results of experiment 2 suggest that the naming task is the least sensitive task to STI, while Stroop task and, in particular, the new paradigm we are introducing, MFAP, performed in the expected way, that is, served to distinguish the trials where inference is expected to occur from those where it is not.

Experiment 3

In experiment 3, we explored how the information inferred from the sentence is integrated with other relevant material. Our hypothesis was that the information is differently integrated depending on the relevance of the actor to the behavior.

This final study also allows us to test the paradigm's sensitivity to spontaneous trait transference (STT), the tendency to transfer the inferred trait to a person who only describes the behavior of a third party or who is simultaneously presented with the behavioral information (*e.g.*, Carlston et al., 1995; Skowronski et al., 1998). In fact, Brown and Bassili (2002) published an article whose title contained the expression "superstitious banana", where they showed that trait transference can even be extended to inanimate objects. Goren and Todorov (2009) showed that STT occurs even when participants are told that the behavioral description was randomly paired with the person. Moreover, evidence suggests that trait information is differently processed in STI and STT. Differences include a stronger effect of STI than STT (*e.g.*, Brown & Bassili, 2002; Skowronski et al., 1998; Goren & Todorov, 2009), the reduction or elimination of STT effects when the actor is presented concurrently (Crawford, Skowronski, & Stiff, 2007; Goren & Todorov, 2009; Todorov & Uleman, 2004) and a generalization of trait impressions in the form of halo effect in STI but not STT (Crawford, Skowronski, & Stiff, 2007; Carlston & Skowronski, 2005; Skowronski et al., 1998). These differences have led to an ongoing debate about the nature of the cognitive processes that underlie them, with some proposing a two-process approach (an attribution-based (or inferential) process in STI and an associative process in STT; *e.g.*, Carlston & Skowronski, 2005; Crawford, Skowronski, & Stiff, 2007) and others, a contrasting single-process model based on different strengths of association between the trait and the person (Brown & Bassili, 2002; Bassili & Smith, 1986; Orghian, Garcia-Marques, Uleman, & Heinke, 2015). A paradigm that is sensitive to both STI and STT can make a useful contribution to this debate. By using MFAP, we predict that there will be a stronger STI effect than STT (Brown & Bassili, 2002; Goren & Todorov, 2009; Skowronski et al., 1998) because a larger activation of the semantic network of the inferred trait is expected in STI and, in particular, a larger integration of these activations within the representation of the actor. Consequently, this will lead to faster reaction times to deliver an associate word when the prompt is the trait that was inferred about that actor.

Method

Participants. 51 first-year undergraduate students from Birmingham University, UK, participated for academic credits. 2 were males and the average age of the sample was 18.86 (SD = .89) years old. As opposed to experiment 1, where we were looking for a main effect, in this experiment we were looking for an interaction (as explained below) and for this reason we slightly increased the sample size.

Material. The same 24 trait-implying sentences tested in the pilot study from experiment 1 are used in this experiment. Along with the 24 critical sentences, 48 neutral sentences were given for memorization in the learning phase. For the free association task, the 24 critical traits were used, 24 non-trait words that were associated with the sentence (*e.g.*, for the sentence “She ordered a bottle of wine during dinner.” the prompt word was “restaurant”) and 24 non-trait words that were part of 24 neutral sentences (*e.g.*, the sentence “Because she was robbed, she had no money to pay the bills.” contains the prompt word “robbed”). As explained later, the main purpose of these fillers is to explore the sensitivity of the paradigm to different levels of activation. A word that is part of the sentence is expected to be more activated than a word that is related with the sentence. Moreover, we want to compare the trait inference effect with the activation of other non-trait words inferred or associated with the sentence. In addition to the words and sentences, 72 black and white photos of people with neutral expressions were used (Lundqvist, Flykt, & Öhman, 1998).

Procedure

Participants were told that the goal of the study was to investigate how people memorize different types of information and how they multitask. The study consisted of two phases, a learning phase and a free association test phase. In the learning phase participants were instructed to memorize photo-sentence pairs of material. In this phase there were two conditions: a STI (relevant) condition where the person presented in the photo was said to be the actor of the action described in the sentence and a STT (irrelevant) condition where the person in the photo was said to be unrelated to the behavioral sentence. STT occurs if the trait inferred from the sentence is transferred to the irrelevant person in the photo. Henceforth, this is referred to as the relevance manipulation. As in Goren and Todorov's (2009) studies, in the STT condition the participants were told that the person and the sentence were randomly paired instead of saying that the person in the photo is communicating the behavior of a third party, as it is commonly done in the literature (*e.g.*, Skowronski et al., 1998; Crawford, Skowronski, Stiff, & Scherer, 2007; Crawford et al., 2008). We used the random pairing method because otherwise the participant might make assumptions about the relationship between the communicator and the person being described and consequently infer that the communicator has similar (or dissimilar) traits to those he is describing (*e.g.*, the communicator chose to talk about behaviors he agrees with (or disagrees), or because they have person relevance of the communicator, *etc.*). Photo relevance was made explicit in the learning phase by presenting the photos in one of two colored borders. Instructions explained that when the sentence described the actions of the person in the photo, the photo would be presented in a blue border and when the text was randomly paired with the photo, it would be shown in a red border (see figure 3). The participants were explicitly instructed to memorize all the information regardless of the type of photograph. In the four practice trials, reminder text regarding the meaning of the borders' colors was shown underneath each photo. In the STT condition, the gender of the person in the photo was opposite to the gender of the pronoun used in the sentence, to emphasize that there was no relation between the person and behavior. The learning phase consisted of 72 trials (36 STI and 36 STT). Each trial started with a 500 ms fixation dot that was followed by the simultaneous presentation of the sentence and the photo for 6000 ms. The photo was presented in the center of the screen and the sentence was presented underneath the photo. After the sentence and the photo disappeared, a new trial began. A "spot the differences" distractor task lasting three minutes followed the learning phase. Next, participants read the instructions for the test phase and had an opportunity to ask questions before and after four practice trials. In the test phase we reminded them that we were interested in investigating their multitasking abilities and thus, they had to perform two different tasks simultaneously in each trial. 1) They were presented with the same photos again and were instructed to use this second chance to better learn the photos and 2) a word would be presented and they had to say the first word coming to their mind. The free association trials began with a 500 ms fixation cross followed by a bleep warning. Next, one of the photos presented earlier, this time without an identifying border, appeared in the center of the screen for 1000 ms. The prompt word then appeared directly underneath the photo and remained until a vocal response was detected. Once a response was recorded a blank screen was shown for 500 ms, followed by the final screen which stated 'press C to continue'.

There were different types of critical trials in the test phase. Half the photos were photos from the STT condition in the learning phase and the other half from the STI condition. In the critical trials, the prompt was the trait implied in the trait-implying sentence presented during the learning phase. In the filler trials, the word wasn't a trait, rather it was a word that was related to a neutral sentence or a word that was part of a neutral sentence presented in the learning phase. In this phase, there were two types of pairings, the match and the mismatch/control trials. In match trials, the photo was presented with the trait (word) implied (associated with or presented) in the sentence presented with that photo in the learning phase. In the mismatch trials, the photo was presented with a trait (word) that was inferred (associated with or presented) in a sentence presented in the learning phase with a different photo. Thus, each type of target word could be in one of the four relevance-pairing conditions; relevant/match, relevant/mismatch; irrelevant/match, irrelevant mismatch. Again, this manipulation is based on Goren and Todorov's (2009) and Todorov and Uleman's (2002) work. An example of the irrelevant condition with the two types of pairing is shown in figure 3. The difference between the match and the mismatch condition informs us about the integration of the trait into the representation of the actor or into the representation of the irrelevant person. In the case of the irrelevant person, the integration is expected to be much less. If the trait is inferred from the sentence and bound to the person, then, in the test phase, when the trait is presented with that person, a facilitation is expected. In other words, the photo will re-instantiate the inferred trait and its network, making it easier to generate an associate when the trait is presented for the free association task. For each participant, there were 24 critical trials, 12 STI (from which 6 were match and 6 mismatch) and 12 STT (6 match and 6 mismatch) trials. The assignment of the sentences to the photos and the presentation order were randomized individually for each participant. Note that in this case, we don't have rearranged sentences, the control is the mismatch condition. Not using rearranged sentences might seem a problem, but the fact that we already used these sentences in experiment 1 and demonstrated they were able to detect inference means we can be confident the material captures inference beyond the word-based priming. We have also shown, in previous research, that not using rearranged controls is not a serious problem when the measure is delayed (Orghian et al., n.d.) because the activation which arises from specific words in the sentence decays with time, while the inference itself does not suffer such decay.

Results and Discussion

Spoiled trials (recording problems flagged by the experiment) were eliminated, making a total of 9.12% of the whole set of responses. Next, the same MAD method was applied, eliminating 14.38% of the remaining responses.

In this experiment, once more, the amount of missing values is larger than 20%. A Little's MCAR test was conducted in order to verify if the missing values are missing completely at random. The Little's MCAR test obtained for this experiment data resulted in a $Chi - square = 1427.45$, $df = 1490$, $p = .875$, which tells us that no identifiable pattern exists to the missing data.

A repeated measure ANOVA was conducted. All the trials were included in the analysis: - the critical trials where the prompt words are traits implied in the correspondent sentences, - the trials where the prompts are non-trait words that are part of the sentences and - the trials where the prompts are non-trait words that are only associated with the correspondent sentences. The first factor is the type of trial (critical, non-trait present, or non-trait associated), the second is the relevance of the photo (actor face – STI condition or irrelevant face – STT condition) and the third factor is the pairing between the face and the trait (match or mismatch). As presented in figure 4, a significant three-way interaction was found between the type of trial, the relevance and the pairing, $F(2, 100) = 3.49$, $p = .034$, $\eta_p^2 = .07$.

Pairwise comparisons were conducted for each type of trial. For the critical trials, a larger facilitation was detected for the actor in the match condition ($M = 1556.49$, $SD = 278.74$) than in the mismatch condition ($M = 1672.70$, $SD = 379.57$), $F(2, 100) = 4.69$, $p = .035$, $\eta_p^2 = .09$. The opposite pattern was detected in the irrelevant condition, with a larger facilitation in the mismatching pairs ($M = 1549.88$, $SD = 282.93$) than in the matching pairs ($M = 1643.98$, $SD = 333.48$), $F(2, 100) = 3.21$, $p = .079$, $\eta_p^2 = .06$.

A similar result was obtained in the trials where the prompt word is an associate of the sentence. When the photo presented at test was the actor, there was a significant difference between the match ($M = 1393.33$, $SD = 192.50$) than the mismatch condition ($M = 1482.00$, $SD = 228.02$), $F(2, 100) = 7.84$, $p = .007$, $\eta_p^2 = .14$. The opposite pattern was detected in the irrelevant condition, with a larger facilitation in the mismatch ($M = 1344.92$, $SD = 206.73$) than in the match condition ($M = 1422.79$, $SD = 241.99$), $F(2, 100) = 3.77$, $p = .058$, $\eta_p^2 = .07$.

Finally, for the presented targets, the difference between match ($M = 1278.91$, $SD = 190.74$) and mismatch condition ($M = 1296.43$, $SD = 173.66$) when the person is the actor was not significant, $F(2, 100) = .42$, $p = .521$, $\eta_p^2 = .008$, and the same was true for the irrelevant person (match condition: $M = 1323.42$, $SD = 207.42$ and mismatch condition: $M = 1330.12$, $SD = 208.34$), $F(2, 100) = .028$, $p = .868$, $\eta_p^2 = .001$.

The main difference between STI and STT conditions is in the amount of integration of the trait information into the representation of the person. The trait becomes strongly associated with the representation of the actor, but it does not with the representation of the irrelevant person. More specifically, in this experiment we show that the MFAP can also be used to study the link between the trait and the actor. The paradigm led to distinguishable effects in STI and STT conditions. This dissociation is originated by an unexpected pattern of results found in the STT condition. Instead of a smaller effect we find an opposite effect, that is, there is more facilitation for the mismatch condition than for the match condition. This opposite pattern might be interpreted as a reflection of an inhibitory mechanism triggered by the presence of an irrelevant person. When the person in the photo is not the actor, the participant might inhibit the trait inferred and its linkage to the person. Thus, in the test phase, when the photo is presented, the trait inferred from the sentence presented with that irrelevant person might be suppressed and so its semantic network. When the trait does not much the photo, a process similar to the release from inhibition (MacLeod, 1989; Geiselman & Bagheri, 1985) might be taking place. The same dissociation was found for the trials where the targets were non-trait words associated with neutral sentences. However, the fact that this same unexpected effect is not found for the words that were actually part of the sentences might be suggesting that what is being inhibited is not the sentence *per se*. We cannot say that the photo of the irrelevant person is being inhabited either, because in that case the inhibition would be detected in all the trials, including the fillers where the prompts were part of the neutral sentences. So, a possible but speculative explanation is that the *link* between the photo and the inferred concept is being inhibited.

In sum, experiment 3 shows that MFAP can we used to investigate other facets of STI and, in particular, that the paradigm is sensitive to other relevant information about the actor that is provided to the participant in the context of the task.

General Discussion

Our experiments were motivated by a desire to overcome some of the main limitations of the traditional paradigms used to study STI. An important characteristic of STI paradigms is that all of them need to use cover stories in order to make the participant process the trait-implying behavioral descriptions without triggering explicit goals of impression formation. One popular way of doing this is by instructing the participants to memorize the material for a later unspecified memory test. Later, during a memory test, the participants are required to contrast a target (the trait) against the memorized material (the behavioral sentence). For such a contrast to be possible, the participant has to retrieve (or attempt to) the memorized material. However, retrieving the sentence at the test might trigger trait inference at that moment in time. If this happens, it means the inference is not a result of processing the sentence and spontaneously inferring the trait from it, but a result of retrieval processes happening later on. This problem is called contamination with explicit retrieval. Paradigms based on procedures that require retrieval of memorized material are called memory-based measures. Cued-recall (Winter & Uleman, 1984), probe recognition (Uleman, Hon, et al., 1996), savings in relearning (Carlston & Skowronski, 1994) and false recognition (Goren & Todorov, 2009) are memory-based measures that can be easily contaminated by explicit retrieval processes. A way to overcome contamination is by accessing the inference without requiring the participants to explicitly retrieve the learned information (Keenan et al., 1990). Activation measures are usually presented as the solution for the contamination problem, but these measures have their own drawbacks. Interestingly, activation measures are uncommon in STI research and we suspect that it is because they might be less sensitive to top-down phenomena like STIs. In other words, activation measures rely on superficial processing of the material, and thus might be less sensitive to STI. Moreover, some activation measures like the naming task and the lexical decision task rely on very fast responses and, for that reason, floor effects might be more frequent than in memory measures (Potts et al., 1988).

In the present paper, we introduced a new paradigm, the Modified Free Association Paradigm (MFAP), a conceptually-driven measure. The measure follows Hourihan and Macleod's (2007) rationale in which it is claimed that if a word is conceptually processed, the spread of activation will extend to other words within the semantic network. This priming of the semantic network will lead to faster reaction times when the participant has to generate an associate for the inferred trait. This task seems capable of detecting activations present in the semantic area that sustains the inference and not just the inferred concept *per se*.

In experiment 1, after reading trait-implying and rearranged sentences, participants were presented with targets and instructed to say the first word that came to their mind. The RTs to generate a word upon the sight of a target trait were shorter when the target followed a sentence that implied the trait than when the target followed a sentence that did not imply that trait. We interpreted this difference in RTs as being due to the fact that the trait is spontaneously inferred during the reading of the implying sentences and, as a consequence, the semantic network of that trait becomes activated, facilitating the generation of associates in the free association task.

In experiment 2, we used a delayed version of the MFAP together with a naming task and a modified Stroop task. We did not find STIs with the naming task, which might be due to its lower sensitivity to conceptual activation or due to a floor effect (Norris, 1986; Keenan et al., 1990). Potts and colleagues (1988), for example, also suggested in their work that naming might not be sensitive to inference, which is consistent with the very short RTs they reported with this task (423 ms). We did, however, find some evidence that making an inference interfered with performance in the modified Stroop task. And, replicating the result in experiment 1, a facilitation to generate associates for the traits implied in previously seen sentences was detected with the MFAP. Crucially, this experiment suggests that the MFAP allows for trait inference detection even when applied in a delayed fashion.

The aim of experiment 3 was to validate the paradigm's sensitivity to the relevance of the actor, and thus to Spontaneous Trait Transference (STT). STT is a neighboring effect of STI that consists of the association of the inferred trait to the wrong person. As predicted, and in agreement with the strong STI effect usually found in the literature, a facilitation effect was found in the trials where the person presented together with the behavioral description was the actor of the actions described in the sentence. An opposite effect was found for trials where an irrelevant person was presented together with the behavioral information. Contrary to the usual difference found in the magnitudes of the two effects (*i.e.*, STI larger than STT, Brown & Bassili, 2002; Goren & Todorov, 2009; Skowronski et al., 1998), we found a dissociation between the two phenomena that suggests the existence of two processes, one responsible for STI and one for STT. This finding is in agreement with past evidence showing differences between STI and STT across a range of situations (Crawford, Skowronski, & Stiff, 2007; Goren & Todorov, 2009; Todorov & Uleman, 2004). However, instead of assuming that a complex process such attributional thinking is taking place in STI (Skowronski et al., 1998), it is also possible that the process in STI is purely associative and something else is happening in STT, like inhibition, an attempt (conscious or not) to suppress the erroneous association between unrelated information. This interesting speculation that seems to emerge from our data needs further testing.

Even though this paradigm apparently leaves us with more questions than answers in what concerns the processes underlying STI and STT, we believe that it can bring us a step closer to understanding the mechanisms and the differences between these two phenomena. Moreover, past research shows that the introduction of a new paradigm can uncover new findings and new perspectives of the investigated phenomena. The creation of savings in relearning paradigm for instance, that was intended to provide a measure of the actor-trait link, inadvertently led to the identification of the spontaneous trait transference phenomenon (Carlston et al., 1995, experiment 4).

In summary, the experiments presented in this paper suggest that the modified free association paradigm may be a viable tool to use in future trait inference research.

The task presents characteristics that makes it appealing when compared to the more traditional paradigms. To start with, participants process the material under memory instruction without encouraging explicit trait inference, a crucial condition to grasp a spontaneous phenomenon. The task is an activation measure, as opposed to a memory measure. This means it is implausible that the inference results from contrasting the trait against the memorized sentence during test, since the participant is not asked to retrieve the sentence or to consider any past events while performing the free association task. This measure, when used with rearranged controls, can ensure that "real" inferences, and not just activation coming from word-based priming, are being detected. Filler trials are also used making the goal of the task less obvious to the participant. An important advantage of this task is that it allows us to measure the access to the semantic network of a trait and not only the activation of the trait itself. This is of particular relevance in the context of the minimalist hypothesis (*e.g.*, McKoon & Ratcliff, 1986) that claims that more elaborate inferences are only partially activated. Thus, if an inference is not totally computed, being able to access related neighbors that support that inference might be informative in regard to the activation state of that concept. Moreover, contrary to the lexical decision and naming tasks, in which the responses are given in a very short time window (between 600 and 700 ms in the lexical decision and between 500 and 600 ms in the naming task), and are thus less sensitive to top-down thinking, MFAP takes longer (because generating words is more difficult, it take usually more than 1500 ms) leaving more room for top-down influences to be detected. Moreover, MFAP is a conceptually-driven task that fits better with a conceptual mechanism taking place in STI.

Although our findings suggest that the free association task is sensitive to trait inference, its validity requires further testing. In particular, the next important step is to investigate how MFAP is affected by backward associations. The backward priming is an interesting aspect of trait inference because it relates to a theoretical dichotomy between inductive (from behavior to trait) and deductive (from trait to behavior) inferences. Maass, Colombo, Colombo and Sherman (2001) introduced the term induction-deduction asymmetry (IDA) after showing that people draw more behavior-to-trait inferences than trait-to-behavior inferences. Note that backward association is exactly the generation of a behavior from a given trait; it is a deductive inference. The rationale behind this result is the fact that a single trait may find expression in many different behaviors whereas a behavior is usually diagnostic on one or few traits. In other words, a trait is linked to many behaviors while a behavior is linked to fewer traits. And assuming that activation is inversely related to the number of existing relations/links, the activation of a concept/node results in a “fan effect”, that is, each link receives only part of the entire activation. The fan effect is suggested to be larger for the deductive inference because the trait is linked to many behaviors and thus, the connection of the trait to each of those behaviors is weaker. This finding suggests that what we have been calling backward association might be weaker than the forward association in the context of trait inference. However, it is still necessary to verify if results similar to the ones obtained in this paper can be explained purely based on backward associations.

Moreover, in the third experiment, one can argue that during the test phase, when the photo is presented for the second time, it is conceivable to think that the participant might try to recall the behavior as a strategy to improve their memory of the photo. This recollection of the behavior can trigger the inference at test. A subliminal priming of the photo could solve this problem by decreasing the probability of explicit sentence recall, a modification that should be considered in future research. And finally, even if fillers are used and the relation between the memorization phase and the free association phase is not obvious, we don’t know for sure that the participants are not aware of it. Thus, in order to verify if the participants are aware of the purpose of the experiment, a questionnaire or a structured interview should be applied after the test phase.

To sum up, our final remark is that there is no such thing as a perfect paradigm, and because of that, researchers, when conducting their research, should 1) consider the different paradigms they have available, 2) know their limitations and advantages, and 3) combine more than one paradigm to answer the same research question in order to gain confidence over the outcome.

References

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? the role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 340–357.

Bassili, J. N., & Smith, M. C. (1986). On the spontaneity of trait attribution: Converging evidence for the role of cognitive strategy. *Journal of Personality and Social Psychology*, *50*(2), 239–245.

Bassili, J. N., Smith, M. C., & MacLeod, C. M. (1989). Auditory and visual word-stem completion: Separating data-driven and conceptually driven processes. *The Quarterly Journal of Experimental Psychology*, *41*(3), 439–453.

Brown, R. D., & Bassili, J. N. (2002). Spontaneous trait associations and the case of the superstitious banana. *Journal of Experimental Social Psychology*, *38*(1), 87–92.

Calvo, M. G., & Castillo, M. D. (1998). Predictive inferences take time to develop. *Psychological Research*, *61*(4), 249–260.

Carlston, D. E., & Skowronski, J. J. (1994). Savings in the relearning of trait information as evidence for spontaneous inference generation. *Journal of Personality and Social Psychology*, *66*(5), 840–856.

Carlston, D. E., & Skowronski, J. J. (2005). Linking versus thinking: evidence for the different associative and attributional bases of spontaneous trait transference and spontaneous trait inference. *Journal of Personality and Social Psychology*, *89*(6), 884–898.

Carlston, D. E., Skowronski, J. J., & Sparks, C. (1995). Savings in relearning: Ii. on the formation of behavior-based trait associations and inferences. *Journal of Personality and Social Psychology*, *69*(3), 420–436.

Claeys, W. (1990). On the spontaneity of behaviour categorization and its implications for personality measurement. *European Journal of Personality*, *4*(3), 173–186.

Conrad, C. (1974). Context effects in sentence comprehension: A study of the subjective lexicon. *Memory and Cognition*, *2*(1), 130–138.

Corbett, A. T., & Dosher, B. A. (1978). Instrument inferences in sentence encoding.

Journal of Verbal Learning and Verbal Behavior, *17*(4), 479–491.

Crawford, M. T., Skowronski, J. J., & Stiff, C. (2007). Limiting the spread of spontaneous trait transference. *Journal of Experimental Social Psychology*, *43*(3), 466–472.

Crawford, M. T., Skowronski, J. J., Stiff, C., & Leonards, U. (2008). Seeing, but not thinking: Limiting the spread of spontaneous trait transference ii. *Journal of Experimental Social Psychology*, *44*(3), 840–847.

Crawford, M. T., Skowronski, J. J., Stiff, C., & Scherer, C. R. (2007). Interfering with inferential, but not associative, processes underlying spontaneous trait inference. *Personality and Social Psychology Bulletin*, *33*(5), 677–690.

D'Agostino, P. R., & Beegle, W. (1996). A reevaluation of the evidence for spontaneous trait inferences. *Journal of Experimental Social Psychology*, *32*(2), 153–164.

Dosher, B. A., & Corbett, A. T. (1982). Instrument inferences and verb schemata. *Memory and Cognition*, *10*(6), 531–539.

Geiselman, R. E., & Bagheri, B. (1985). Repetition effects in directed forgetting: Evidence for retrieval inhibition. *Memory and Cognition*, *13*(1), 57–62.

Goren, A., & Todorov, A. (2009). Two faces are better than one: Eliminating false trait associations with faces. *Social Cognition*, *27*(2), 222–248.

Ham, J., & Vonk, R. (2003). Smart and easy: Co-occurring activation of spontaneous trait inferences and spontaneous situational inferences. *Journal of Experimental Social Psychology*, *39*(5), 434–447.

Hamann, S. B., & Squire, L. R. (1996). Level-of-processing effects in word-completion priming: A neuropsychological study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(4), 933–947.

Hourihan, K. L., & MacLeod, C. M. (2007). Capturing conceptual implicit memory: The time it takes to produce an association. *Memory and Cognition*, *35*(6), 1187–1196.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513–541.

Keenan, J. M., Potts, G. R., Golding, J. M., & Jennings, T. M. (1990). Which elaborative inferences are drawn during reading? a question of methodologies. In D. A. Balotta, G. B. F. d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (p. 377-402). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review*, *95*(2), 163–182.

Lee, H., Shimizu, Y., & Uleman, J. S. (2015). Cultural differences in the automaticity of elemental impression formation. *Social Cognition*, *33*(1), 1.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766.

Light, L. L., Prull, M. W., & Kennison, R. F. (2000). Divided attention, aging, and priming in exemplar generation and category verification. *Memory and Cognition*, *28*(5), 856–872.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The karolinska directed emotional faces-kdef. cd-rom from department of clinical neuroscience, psychology section, karolinska institutet, stockholm, sweden* (Tech. Rep.). ISBN 91-630-7164-9.

Lupfer, M. B., Clark, L. F., & Hutcherson, H. W. (1990). Impact of context on spontaneous trait and situational attributions. *Journal of Personality and Social Psychology*, *58*(2), 1239–1249.

Maass, A., Colombo, A., Colombo, A., & Sherman, S. J. (2001). Inferring traits from behaviors versus behaviors from traits: The induction–deduction asymmetry. *Journal of Personality and Social Psychology*, *81*(3), 391–404.

MacLeod, C. M. (1989). Directed forgetting affects both direct and indirect tests of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(1), 13–21.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). Opensesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324.

McCarthy, R. J., & Skowronski, J. J. (2011). What will phil do next?: Spontaneously inferred traits influence predictions of behavior. *Journal of Experimental Social Psychology*, *47*(2), 321–332.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(1), 82–91.

McKoon, G., & Ratcliff, R. (1989a). Inferences about contextually defined categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1134–1146.

McKoon, G., & Ratcliff, R. (1989b). Semantic associations and elaborative inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(2), 326–338.

McKoon, G., & Ratcliff, R. (1990). Dimensions of inference. *Psychology of Learning and Motivation*, *25*, 313–328.

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1003–1019.

Nelson, D. I., & Goodmon, L. B. (2002). Experiencing a word can prime its accessibility and its associative connections to related words. *Memory and Cognition*, *30*(3), 380–398.

Newman, L. S. (1991). Why are traits inferred spontaneously? a developmental approach. *Social cognition*, *9*(3), 221–253.

Newman, L. S. (1993). How individualists interpret behavior: Idiocentrism and spontaneous trait inference. *Social Cognition*, *11*(2), 243–269.

Norris, D. (1986). Word recognition: Context effects without priming. *Cognition*, *22*(2), 93–136.

Orghian, D., Garcia-Marques, L., Uleman, J. S., & Heinke, D. (2015). A connectionist model of spontaneous trait inference and spontaneous trait transference: Do they have the same underlying processes? *Social Cognition*, *33*(1), 20-66.

Orghian, D., Ramos, T., & Garcia-Marques, L. (n.d.). Activation is not always inference: word-based priming and spontaneous trait inferences. *Manuscript submitted for publication.*

Orghian, D., Ramos, T., Reis, J., & Garcia-Marques, L. (in press). Acknowledging the role of word-based activation in spontaneous trait inferences. *Análise Psicológica*.

Park, B. (1989). Trait attributes as on-line organizers in person impressions. In J. N. Bassili (Ed.), *On-line cognition in person perception* (pp. 39–59). Hillsdale, NJ: Erlbaum.

Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language*, *27*(4), 399–415.

Rim, S., Min, K. E., Uleman, J. S., Chartrand, T. L., & Carlston, D. E. (2013). Seeing others through rose-colored glasses: An affiliation goal and positivity bias in implicit trait impressions. *Journal of Experimental Social Psychology*, *49*(6), 1204–1209.

Roediger, H. L., & Blaxton, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word-fragment completion. *Memory and Cognition*, *15*(5), 379–388.

Roediger, H. L., Weldon, M. S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger & F. I. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of endel tulving* (pp. 3–41). Hillsdale, NJ: Erlbaum.

Roediger, H. L., Weldon, M. S., Stadler, M. L., & Riegler, G. L. (1992). Direct comparison of two implicit memory tests: Word fragment and word stem completion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(6), 1251–1269.

Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, *88*(424), 1273–1283.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *13*, 501–518.

Schneid, E. D., Crawford, M. T., Skowronski, J. J., Irwin, L. M., & Carlston, D. E. (2015). Thinking about other people. *Social Psychology*.

Seidenberg, M. S., Waters, G. S., Sanders, M., & Langer, P. (1984). Pre-and postlexical loci of contextual effects on word recognition. *Memory and Cognition*, *12*(4), 315–328.

Skowronski, J. J., Carlston, D. E., Mae, L., & Crawford, M. T. (1998). Spontaneous trait transference: Communicators take on the qualities they describe in others. *Journal of Personality and Social Psychology*, *74*(4), 837–848.

Srull, T. K., & Wyer, R. S. (1989). Person memory and judgment. *Psychological Review*, *96*(1), 58–83.

Till, R. E., Mross, E. F., & Kintsch, W. (1988). Time course of priming for associate and inference words in a discourse context. *Memory and Cognition*, *16*(4), 283–298.

Todorov, A., & Uleman, J. S. (2002). Spontaneous trait inferences are bound to actors' faces: evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*, *83*(5), 1051–1065.

Todorov, A., & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology*, *39*(6), 549–562.

Todorov, A., & Uleman, J. S. (2004). The person reference process in spontaneous trait inferences. *Journal of Personality and Social Psychology*, *87*(4), 482–493.

Uleman, J. S. (1999). Spontaneous versus intentional inferences in impression formation. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 141–160). New York, NY: Guilford.

Uleman, J. S., Hon, A., Roman, R. J., & Moskowitz, G. B. (1996). On-line evidence for spontaneous trait inferences at encoding. *Personality and Social Psychology Bulletin*, *22*(4), 377–394.

Uleman, J. S., & Moskowitz, G. B. (1994). Unintended effects of goals on unintended inferences. *Journal of Personality and Social Psychology*, *66*(3), 490–501.

Uleman, J. S., Moskowitz, G. B., Roman, R. J., & Rhee, E. (1993). Tacit, manifest, and intentional reference: How spontaneous trait inferences refer to persons. *Social Cognition*, *11*(3), 321–351.

Uleman, J. S., Newman, L. S., & Moskowitz, G. B. (1996). People as flexible interpreters: Evidence and issues from spontaneous trait inference. *Advances in Experimental Social Psychology*, *28*, 211–279.

Uleman, J. S., Winborne, W. C., Winter, L., & Shechter, D. (1986). Personality differences in spontaneous personality inferences at encoding. *Journal of Personality and Social Psychology*, *51*(2), 396–403.

Van Duynslaeger, M., Van Overwalle, F., & Verstraeten, E. (2007). Electrophysiological time course and brain areas of spontaneous and intentional trait inferences. *Social Cognitive and Affective Neuroscience*, *2*(3), 174–188.

Van Overwalle, F., Drenth, T., & Marsman, G. (1999). Spontaneous trait inferences: Are they linked to the actor or to the action? *Personality and Social Psychology Bulletin*, *25*(4), 450–462.

West, R. F., & Stanovich, K. E. (1982). Source of inhibition in experiments on the effect of sentence context on word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*(5), 385–399.

Whitney, P. (1986). Processing category terms in context: Instantiations as inferences. *Memory and Cognition*, *14*(1), 39–48.

Whitney, P., & Kellas, G. (1984). Processing category terms in context: instantiation and the structure of semantic categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 95–103.

Whitney, P., Waring, D. A., & Zingmark, B. (1992). Task effects on the spontaneous activation of trait concepts. *Social Cognition*, *10*(4), 377–396.

Whitney, P., & Williams-Whitney, D. (1990). Toward a contextualist view of elaborative inferences. *Psychology of Learning and Motivation*, *25*, 279–293.

Wigboldus, D. H., 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*(3), 470–484.

Wigboldus, D. H., Sherman, J. W., Franzese, H. L., & van Knippenberg, A. (2004). Capacity and comprehension: Spontaneous stereotyping under cognitive load. *Social Cognition*, *22*(3), 292–309.

Winter, L., & Uleman, J. S. (1984). When are social judgments made? evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology*, *47*(2), 237–252.

Winter, L., Uleman, J. S., & Cunniff, C. (1985). How automatic are social judgments? *Jouranl of Personality and Social Psychology*, *49*(4), 904-917.

Wyer, R. S., & Srull, T. K. (1986). Human cognition in its social context. *Psychological Review*, *93*(3), 322–359.

Zárate, M. A., Uleman, J. S., & Voils, C. I. (2001). Effects of culture and processing goals on the activation and binding of trait concepts. *Social Cognition*, *19*(3), 295-323.

Table 1

Paradigms and their limitations.

Paradigm	Demand	Contamination	Backward association	Word-priming	Conceptual	Delayed
Cued-recall	NO	NO	NO	NO	YES	YES
Probe recognition	YES	NO	NO	YES	NO	YES
Saving relearning	YES	NO	NO	NO	YES	YES
False recognition	NO	NO	NO	NO	NO	YES
Lexical decision	YES	YES	NO	NO	NO	YES
Stem completion	YES	YES	NO	YES	NO	YES
Naming*	YES	YES	YES	YES	NO	YES
Stroop*	YES	YES	NO	YES	NO	YES
MFAP*	YES	YES	NO	YES	YES	YES

Note. * new paradigms in the context of trait inference research; "NO" - limitation not solved, "YES" - limitation solved.

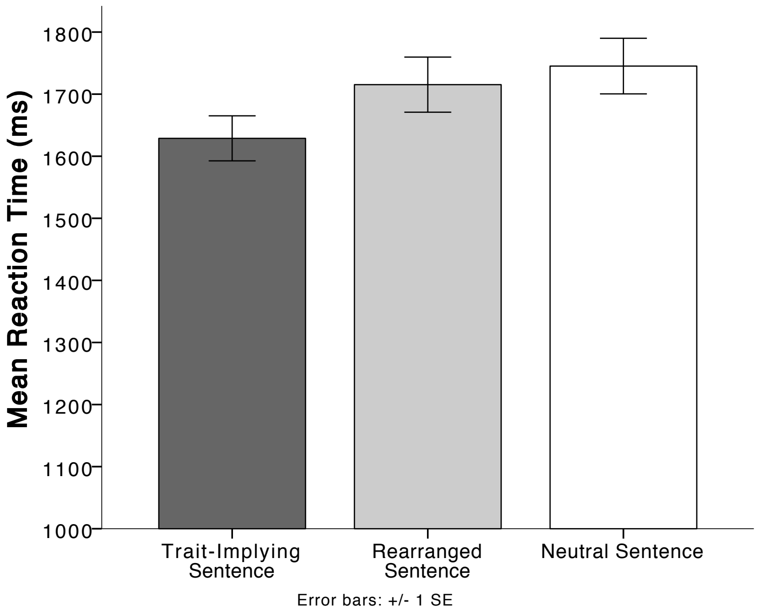

Figure 1. Mean reaction time in function of type of sentence in Experiment 1.

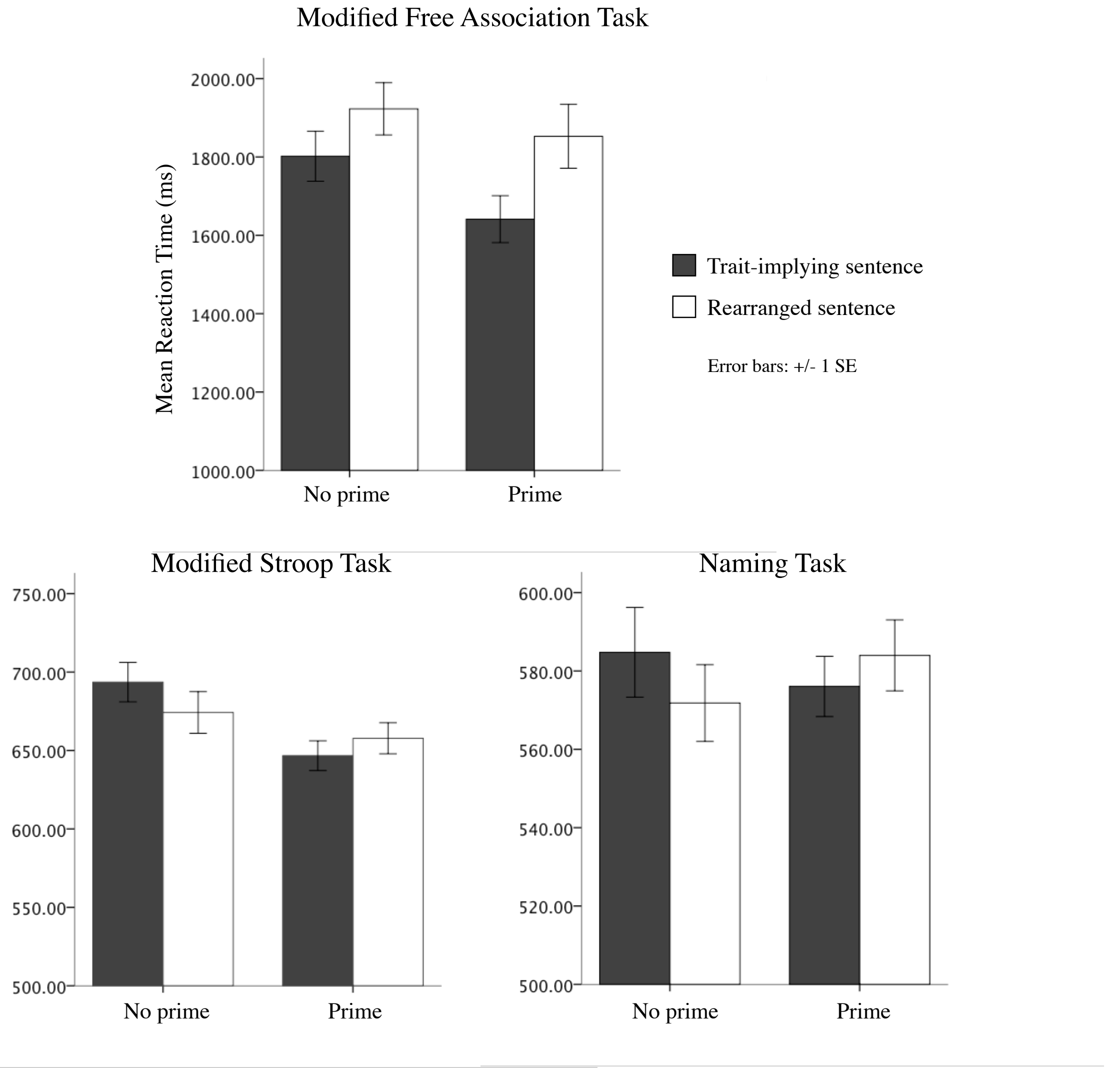

Figure 2. Mean reaction time in function of task, prime condition and type of sentence in Experiment 2.

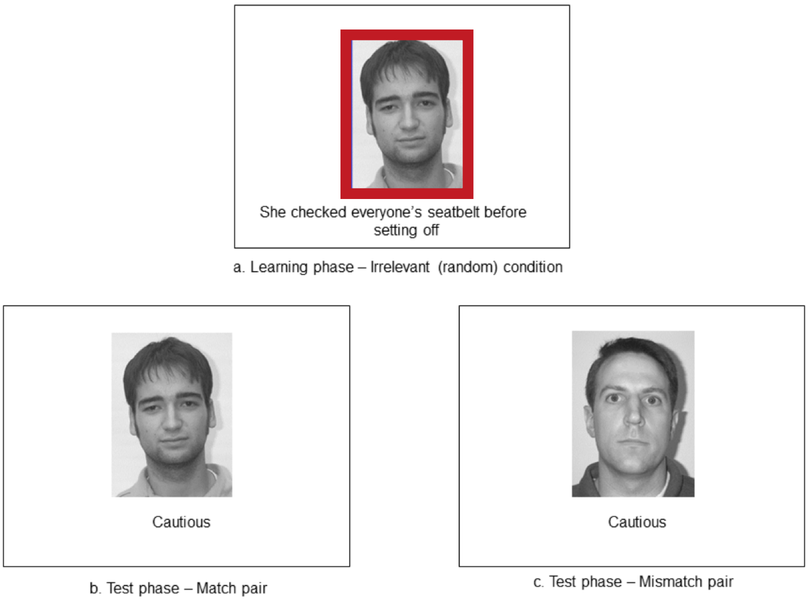

Figure 3. Example screen views of one learning trial and two free association trials showing different pairing conditions. Image a. depicts an irrelevant learning trial (identified by border color and pronoun gender); images b. and c. illustrate a match or mismatch/control pairing, respectively.

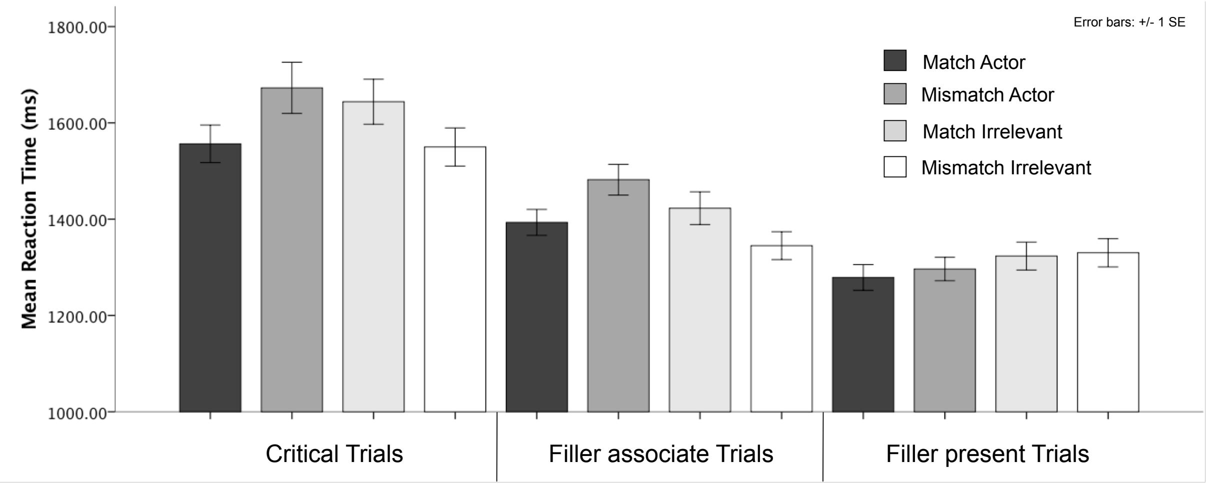

Figure 4. Mean reaction time in function of type of trial, relevance and pairing in Experiment 3.