The Effect of Interviewee Language on Suggestibility in a Forensic Interview Situation

Nora Helmy

Supervisor: Charlotte Alm
MASTER THESIS, 30 ECTS, 2018

STOCKHOLMS UNIVERSITET
PSYKOLOGISKA INSTITUTIONEN
THE EFFECT OF INTERVIEWEE LANGUAGE ON SUGGESTIBILITY IN A FORENSIC INTERVIEW SITUATION

Nora Helmy

Individual tendency to give in to leading and suggestive questions is of importance to eye witness accuracy and credibility. When cognitive resources are depleted, individual suggestibility has been shown to increase. The current study tested, with a sample of 51 psychology students, whether speaking a second language could lead to an increase in suggestibility due to ego-depletion. The study used a mock crime film upon which a free recall in either English or Swedish was to be delivered, followed by suggestively phrased questions. Results indicated that participants who delivered the free recall in English, compared to in Swedish, portrayed higher ego-depletion. Participants’ suggestibility was however only marginally different between the two conditions. Further results from hierarchical regression analysis showed both age and language to be significant predictors of suggestibility. It is concluded that partaking in an interrogative interview, in a non-native language, may lead to ego-depletion with subsequent suggestibility.

Eye witness testimony holds great importance in the legal system and is commonly referred to as one of the strongest but also one of the most unreliable sources of evidence in a criminal investigation (Lindholm, 2005). Errors in eyewitness testimony are prevailing and represent the leading cause of and contribution to false convictions of innocent people (e.g., Gould & Leo, 2010; Wells et al., 1998). Understanding the how’s and why’s of eyewitness psychology is of importance and concern for the administration of justice as well as for psychological research, in particular aiming to explain factors affecting witness accuracy (Lindholm, 2005). There are a variety of factors influencing eye witness accuracy and testimony, and frequently studied is peoples’ tendency to give in to leading and suggestive questions during an interview. It has specifically been shown that individuals who are highly suggestible are more likely to deliver false testimonies (Gould & Leo, 2010). Taking part in a forensic interrogative interview is for most people an unfamiliar and even an unpleasant situation, which demands an increased need of behavior regulation and control of the self (e.g., Vallacher & Wegner, 1987). Research has further shown that individuals who have limited language skills are less capable of withstanding interpersonal stress in an interrogation (Gould & Leo, 2010). No studies to date have looked at the effect of language on individual suggestibility. The present study will test whether giving testimony in a non-native language requires more self-regulatory processes potentially influencing eye witness accuracy and testimony due to subsequent suggestibility.
Factors Affecting Eyewitness Testimony

Suggestibility and its underlying mechanisms.

One vital factor affecting eyewitness testimony is post event information. Questions loaded with suggestions and misleading cues is one way post event information can influence testimony (Gudjonsson, 1984). Individuals are differently susceptible to the acceptance of leading questions and as will be made evident, there are several factors affecting this. An individual’s tendency to respond in a certain way, given a suggestive or misleading question is referred to as suggestibility. Suggestibility holds an important role in eyewitness testimonies and it is believed that a high proposition of leading questions can have distorting effects on testimony accuracy, since such questions often contain certain expectations and premises (Gudjonsson, 1984). Suggestibility effects, linked to forensic interviews, are in the literature referred to as interrogative suggestibility. In 1986 Gudjonsson and Clark defined interrogative suggestibility as: “The extent to which, within a closed social interaction, people come to accept messages communicated during formal questioning, as the result of which their subsequent behavioral response is affected” (referred in Gudjonsson, 2003, p.84). Accordingly, the definition of interrogative suggestibility involves questioning procedures typically related to past events and experiences and to recollections and states of knowledge (Gudjonsson, 2003). Today a well-used measure for interrogative suggestibility, is the Gudjonsson suggestibility scale (GSS; Gudjonsson, 1984, 1997). Besides measuring logical memory and confabulation, the scale incorporates two distinct measures of suggestibility known as shift and yield. Yield refers to a person’s tendency to give in to leading questions and shift refers to individual susceptibility to change their answers when faced with pressure in an interrogative setting (Gudjonsson, 1984, 1997, 2003). The GSS was developed as the first ever psychometric instruments to measure these two types of interrogative suggestibility, with the intention to be used to further the understanding of, and mechanisms underlying interrogative suggestibility. Today the scale is used to identify and assess individuals who are susceptible to give false accounts of events when questions (Gudjonsson, 1984, 1997).

Extensive research has tried to establish the underlying mechanism of suggestibility. Research concludes that it generally is a combination of both cognitive and socioemotional mechanisms that contribute to suggestibility (e.g., Ceci & Bruck, 1993; Gudjonsson, 2003). In addition, has Gudjonsson (2003) noted a negative correlation between memory recall and suggestibility. That is, the poorer a person’s memory the more susceptible to yield responses toward suggestions. Studies have shown that when exposed to suggestions in an interrogative setting, both adults and children frequently accept the suggestion and integrate it into their memory (e.g., Gudjonsson, 2003; Loftus & Davis, 2006; Otgaar, Candel, Merckelbach, & Wade, 2009). Moreover, age is one of the most studied individual difference factors in eyewitness suggestibility and research shows that young children are more suggestible than older children and adults, and that adults aged over 60 portray higher susceptibility levels compared to young adults (e.g. Dukala & Polczyk, 2013; Gudjonsson, 2003).
Individual compliance, has further shown to be of importance when estimating interrogative suggestibility (Gudjonsson, 2003). Compliance is an individual’s tendency to go along with propositions, instruction, or requests, for some instant instrumental gain (Gudjonsson, 2003). Individuals could for example go along with requests that they otherwise would not, with the purpose of faster release from custody, eagerness to please others or termination of a police interview (Gudjonsson, 2003). The Gudjonsson Compliance Scale (GCS; Gudjonsson, 1989) was created as a complimentary measure to the GSS. The GCS is self-report measure which requires people to rate their compliance in relation to interpersonal pressure and demand from others. Whereas suggestibility refers to overt and private acceptance of information, compliance does not necessitate a private acceptance of the request or proposition and information can thus be covertly rejected (Gudjonsson, 1989). In contrast to suggestibility, a person complying can be fully aware of that his or her responses are or can be influenced (Gudjonsson, 2003).

Self-regulation, -control and -presentation.
Within forensic assessment, the measurement and concept of interrogative suggestibility is vital and determining under which conditions suggestion are harmful to memory is of great importance in the legal field. Another area of interest has been the role of self-regulation in relation to suggestibility. Self-regulation is a broad term used to describe higher order executive function processes which are involved in attention, inhibition, working memory, decision – making, volitional control and self–control (Schmeichel, Vohs, & Baumeister, 2003). Self-control and self-regulation may be used interchangeably but if a distinction is to be made, self-control should be considered as the conscious, deliberate and effortful subset of self-regulation (Baumeister, Vohs, & Tice, 2007). Self-control can best be described as the capacity to alter one’s own responses and aligning them with the standards, ideals, morals, values and social expectations a person holds (Baumeister et al., 2007). Self-control and self-regulation hold important keys to the understanding of the function and nature of the self, and reduced self-control can be associated with weakened impulse-control and behavioral problems including alcohol and drug abuse as well as crime and violence (Baumeister et al., 2007). The experience of emotional distress can cause failure in self-control and regulation and when a large amount of self-control is required and executed, cognitive resources may become depleted (Vohs & Heatherton, 2000). The notion that self-control is a limited resource, vulnerable to deterioration over time, was first suggested by Baumeister and colleagues and is referred to as the strength model of self-regulation (Baumeister, 2002; Baumeister & Alquist, 2009; Baumeister et al., 2007; Hagger, Wood, Stiff, & Chatzisarantis, 2010; Muraven & Baumeister, 2000). Accordingly, the model argues that the engagement in self-control, draws from a limited pool of necessary resources that risk exhaustion and which, when depleted result in reduced capacity for further acts of self-regulation (Hagger et al., 2010). Self-regulation is in the model often compared to a muscle, that just as a muscle, requires energy and strength to perform and hence runs the risk of becoming tired with a reduction in further force. When engaging in self-control to the extent that it reduces our prospective abilities to engage in self-control, is in the literature referred to as ego-depletion, hence a state of temporary impairment in self-regulatory capacity (Baumeister et al., 2007; Baumeister, Bratslavsky, Muraven, Tice, & Baumeister, 1998; Baumeister & Heatherton, 1996; Muraven & Baumeister, 2000). The strength
model of self-regulation has shown to influence numerous processes relevant to an interrogative setting and hence also the accurate gathering of eyewitness testimony (e.g., Harkness et al., 2015; Sherman, Macrae, & Bodenhausen, 2000; Sherman, Groom, Ehrenberg, & Klauer, 2003). Strained cognitive resources have been shown to undermine performance on tasks that require controlled and complex cognitive processes (Schmeichel et al., 2003). When people are cognitively exhausted or even depleted, fewer processing resources are available for new information and it has been shown to especially affect the processing of emotion-laden information as well as working memory (Maranges, Schmeichel, & Baumeister, 2016). It is further believed that depleted cognitive resources influence emotion-related memory as well as peer influence (Sherman et al., 2000). Hampered effects of ego-depletion have been associated with one’s capacity for conscious, deliberate, and complex thinking, although automatic and intuitive thinking processes remain largely intact (e.g., Pocheptsova, Amir, Dhar & Baumeister, 2009). Sherman and colleagues (2003) argue that cognitive load facilitates stereotypical thinking and influences the extent to which a person is willing, as well as able to recollect information. For example, they found that ego-depletion increased a person’s reliance on their stereotypes during the recall of an event. Furthermore, Otgaar, Alberts and Cuppens (2012) found that ego-depleted participants portrayed higher levels of suggestibility, compared to non-depleted peers. Their idea was that when people have weakened executive functions, their capability to monitor the sources of their memories are equally weakened and willingness to accept suggestions increases (Otgaar et al., 2012). Their study was the first ever to show such a relation between ego-depletion and suggestibility and findings are of practical significance in interrogative settings. They conclude that interrogative interviews can appear stressful for both witnesses as well as suspects and that such interviews can evoke increased self-control of behavior, emotions and thoughts. Results are important since they imply that ego-depletion, as a result of cognitive self-control, leads to higher degrees of susceptibility to suggestions (Otgaar et al., 2012).

Pivotal for being in an interrogative setting is also the engagement in self-presentation. Self-presentation is the individual’s effort to portray a conveying and particular image of themselves to others, which in itself requires acts of self-regulation (Vohs, Baumeister, & Ciarocco, 2005). Self-presentation can be either strategic or expressive. Strategic self-presentation refers to a manipulation of the audience for ulterior means, whereas expressive motives involve claiming desired identities for the self and are consequently based on the self-presenter’s own standards and values rather than the audiences’. Expressive motives stem from when people try to prove to others that they are not what they expect them to be or to not go along with what others want them to do (Doherty & Barry, 1991; Snyder, 1974). Vohs and colleagues (2005), make the case that different types of self-presentation are differently strenuous and that a more effortful self-presentation consumes more of self-regulatory efforts. It follows that in the case of depleted self-regulatory processes, the less effective will individuals be in presenting themselves in socially desirable or normative ways, especially so if the presentation requires to overrun habitual tendencies (Vohs et al., 2005). Some self-presentations become more habitual and efficient the more they are practiced and therefore require less energy. It follows, that effortful self-presentations also deplete cognitive resources, rendering in difficulties to control and regulate one’s behavior (Vohs et al., 2005).
Research has shown that when individuals find themselves in new and unfamiliar situations they must exert increased attentional control to act accordingly and desirably (e.g., Vallacher & Wegner, 1987).

**Ethnicity and language.**

The importance of self-presentation becomes especially evident for an individual who is faced with stigmatization, skepticism and discrimination (e.g., Bye, Horverak, Sandal & Sam, 2014; Marvasti, 2006). Ethnicity is often seen as a key factor underlying discrimination within the legal system in Sweden and in other western European countries (Lindholm, 2005). In a study by Lindholm (2005), witness ethnicity (swede vs. immigrant) in relation to witness credibility and accuracy was examined. Overall results indicate that both the modality through which a testimony was presented as well the witness’ social group membership could serve as important determinants in judgments of eyewitness credibility. Results further showed that immigrants scored slightly fewer correct details on a free recall task as well were judged to be less good witnesses by participants scoring high on the Modern Racism Scale (MRS; McConahay, 1986).

Furthermore, being of a different ethnicity may imply lower mastery of a country specific language, which in turn could contribute to increased discrimination. For example, people speaking a second language due to their ethnic or social background, may be faced with increased prejudice and disbelief by people with other ethnic or social background (e.g., Collins & Clément, 2012). Several studies recognize that an individual’s social group belonging (outgroup/ingroup) is decisive of how they are judged (Elfenbein & Ambady, 2002; Lindholm, 2008; Tajfel, 1982). In legal settings, it has for example been repeatedly shown that, given the same case evidence, outgroup members are more often judged as guilty and given more sever punishments compared to ingroup members (e.g., Kemmelmeier, 2005; Sommers & Ellsworth, 2000). Further studies show that people are better at judging psychological states of confidence, of members belonging to their own national or ethnic group (ingroup) compared to people aligning with an outgroup, due to an increased understanding of their behavior (Elfenbein & Ambady, 2002; Lindholm, 2008). It may be argued that people’s different ability to judge and interpret ingroup and outgroup members can reflect witness accuracy estimations (Lindholm, 2008). That is, people should be more, or less, efficient at estimating accurate or inaccurate witness’ statements depending on group belonging, with greater favors toward one’s own ingroup (Frumkin, 2007; Lindholm, 2005; Lindholm, 2008).

Practicing a non-native language is, among much, associated with increased skepticism by one’s surroundings but also with demands on cognitive resources. Speech production and speech perception are both processes that require extensive cognitive resources (Juff & Harrington, 2011). The use and learning of a second language draws on a variety of cognitive processes, among other working memory (Juff & Harrington, 2011). Working memory is generally described as the cognitive system for temporary storing and information processing while executing higher order cognitive tasks such as learning, comprehension and reasoning (Baddeley & Logie, 1999). The role of working memory in language, both first and second, is widely established and of special importance is the phonological memory operated by the phonological loop, central in processing and temporary retention of familiar
and novel phonological information (Baddeley, Gathercole, & Papagno, 1998; Juff & Harrington, 2011). Fortkamp (2003) showed in his study that oral production of a second language (English) required both control and regulation of attention, two central cognitive processes. Fehringer and Fry (2007) found that, despite high proficiency, participants speaking a second language produced a higher overall rate of hesitation, implying that additional cognitive load is imposed when working in a second language. Moreover, Mota (2003) showed positive correlations between working memory and fluency, accuracy, and complexity. Further evidence for that operating in a second language requires more cognitive resources can be found regarding grammatical encoding. Grammatical encoding has been shown to be less automatized in a second language compared to in a person’s first language (Fehringer & Fry 2007). In a study by Lee (1986), it was shown that written recall was significantly better when performed in a person’s native language compared to in a person’s second language.

Given the above, it is reasonable to conclude that people speaking a non-native language will not only exhaust cognitive resources due to the nature of language but also due to an increase in cognitive demand to come across as equally credible and accurate as their outgroup peers.

The present study.
Factors which may influence eyewitness accuracy and testimony include suggestibility, self-regulation and spoken language. In addition, researchers have reported that situational and social factors, as well as cognitive components also contribute with vital aspects of suggestibility and hence eyewitness testimony. For that reasons, the present study examined if speaking a second language (English) could affect suggestibility, due to potential ego-depletion. To experimentally test ego-depletion, a dual-task paradigm, which to my knowledge has not been used elsewhere, will be deployed in the present study. The dual-task paradigm in the present study will contain a credibility task and a language task. The study will be conducted in line with Gudjonsson’s GSS guidelines, using a mock crime film and a simulated real-life investigative interview situation. Participants will be assigned to either an experimental non-native language condition, or to a native language control condition. Language will hence serve as the main depletion manipulation. The present study will hypothesize that delivering a free recall of a crime, both credibly and in English, will increase suggestibility levels due to potential ego-depletion. It will further be hypothesized that participants in the English-speaking condition must make larger efforts to come across as credible compared to participants in the Swedish-speaking condition and will deliver shorter and less accurately detailed free recall.

Method

Participants
A sample of 57 psychology students were recruited via bulletin boards and e-mail at the Department of Psychology at Stockholm University. Six participants were excluded from the study and further data analysis as their English skills were judged to be equally good as their Swedish skills and thus, the language manipulation was
expected to have no effect. Analysis was hence based on a sample of 51 psychology students in the ages of 19 to 49 (M age = 26; SD = 6.63; 76.5% females) for whom the manipulation of the independent variable was believed to be successful. In line with the main hypothesis, participants were randomly assigned to either the Swedish-speaking condition (control) (n = 25; M age = 25; SD = 5.12; 80.0% females) or the English-speaking condition (experimental) (n = 26; M age = 26; SD = 7.88; 73.1% females). Participation in the study was part of a mandatory student course moment.

Materials

Mock crime scenario.
In order to measure suggestibility in association with an ecologically valid forensic context, participants were first presented with a 66 second short film picturing a staged crime (Lindholm, Jönsson, & Liuzza, in press). The film portrayed a robbery taking place at a bus stop in the country side of Sweden. More specifically did the film show a young man waiting for a bus when he suddenly is threatened and attacked by another young man and ultimately stabbed in the stomach, and left bleeding on the ground. Aside from the offender and the victim, there are a few people and cars passing by as well as a middle-aged woman coming to the victim’s rescue.

Free recall of the crime.
Ego-depletion was aimed to be induced by a free recall task of verbal recollection of the depicted events in the crime scenario, from a witness perspective. Participants were asked to freely recall/describe to the experimenter, as credible and as detailed, about what they as a witness had just seen. In the experimental condition (i.e. higher levels of depletion) participants were asked to deliver a free recall both credibly and in English. Participants in the control condition (lower levels of depletion) were asked to deliver a credible free recall in Swedish. A scoring template was used to score participants free recall (personal communication with supervisor Charlotte Alm and professor Torun Lindholm, 02/08/18). For correct and detailed recollection of the depicted events could a maximum score of 37 points be received and for correct and detailed description of external attributes about the offender, the victim and the aiding woman, could additional points of a maximum of 67 be received. Participants could hence receive a maximum of 104 points according to the template.

Credibility.
A self-report scale created by Lindholm (2005) was used to assess participant credibility. The original scale consists of seven items, whereof four were judged to be relevant for the present study. Specifically, participants were asked; (1) how accurately they remembered the depicted events; (2) to what extent they could convey their memory of the event during the interview; (3) how they thought that someone who would listen to the recording of their testimony would judge them to be truthful and (4) how motivated they were to recall the event. All items were to be answered on a Likert-type scale ranging from 1(not at all) to 7 (very much). Cronbach’s alpha in for the deployed credibility scale reached .70 (entire sample).
In the present study, an altered suggestibility measure, based on the already existing Gudjonsson suggestibility scale (GSS; Gudjonsson, 1984, 1997), was used to assess participants level of suggestibility. The study specific version was inspired by the Swedish translation of the GSS (Hildebrand Karlén, in press) and created by the author. The above described crime scenario and free recall phase functioned as a memory paradigm. Questions were constructed and formulated in accordance with the GSS and phrased in ways to fit the specific events depicted in the film. The questions consisted of five non-suggestive (NS) questions and 15 suggestive questions (S), of which five questions contained false alternatives. The five non-suggestive, affirmative questions (NS) were phrased in such way that there was one correct answer to be given (e.g., “Was the victim waiting at a bus stop?”) and were interspersed with the 15 suggestively phrased questions. Ten of the suggestive questions were phrased in a leading way, creating neither too strong nor too obvious expectations but including one or more prominent premises (e.g., “Did the victim’s glasses break during the robbery?”) The final five confabulated questions implied false alternatives, such as the presence of events and objects that were not depicted in the film (e.g., “Was the victim wearing gloves or mittens?”). In the present study, each of the suggestive questions that was yielded for was scored with one point, adding to a total of 15 suggestive points. The GSS has been validated extensively with adequate Cronbach’s alpha levels (e.g., Clare, Gudjonsson, Rutter, & Cross, 1994; Gignac & Powell, 2009). Cronbach’s alpha in the present study reached .75 (entire sample).

Compliance.
The study deployed a Swedish version of the GCS (Gudjonsson, 1989; Hildebrand Karlén, in press). The GCS is a self-measurement of compliance and consists of 20 statements, which are to be answered either as True or False (e.g., “I tend to go along with what people tell me even when I know they are wrong”). Scores range from zero to 20 where higher scores indicate higher tendency to comply. Obtained Cronbach’s alpha in the present study was .81 (entire sample). Compliance has been shown to correlate significantly with suggestibility and the GCS was thus included as a measure to exclude that suggestibility levels were driven by compliance and not by the language depletion manipulation (Gudjonsson, 1989).

Manipulation check of ego-depletion.
Four self-report questions, created by Englert, Persaud, Oudejans and Bertrams (2015) and translated into Swedish by the author, were used to assess degree of self-rated ego-depletion (e.g., “How effortful did you find the task?”). Answers were to be given on a Likert scale ranging from 1 (not at all) to 7 (very much). Preliminary analysis reached a Cronbach’s alpha of .34. Further analysis indicated that an exclusion of the item “How much did you have to adjust your presentation to come across as credible” would generate in an increased Cronbach’s alpha of .54. An inspection of the bivariate intercorrelations revealed further low non-significant correlations with the other three items. It was therefore decided to excluded the item from an ego-depletion index and analysis. Three items were used hence as manipulation check in analysis (Cronbach’s alpha = .54)
English skills, sensation, and demographic information.
Two items were used to establish English skills (e.g., “How comfortable are you with English?”). Two further items were used to measure how nervous participants felt during the free recall and how they experienced being questioned about the event. Demographic information (age and sex) was as well collected. All items were to be answered on a Likert-type scale ranging from 1(not at all) to 7 (very much).

Procedure
The study was conducted in accordance with the Ethical law of research concerning human beings (2003:460) and written informed consent was obtained from each participant before the study. The study was presented as a police investigative study where witness memory and credibility was to be tested. Participants were tested individually, by the same test leader, and all tests were administered in accordance with GSS guidelines (Gudjonsson, 1984, 1997; Hildebrand Karlén, in press). As suggested by the manual, participants were first asked to freely talk about their memory in order to create a good working alliance between the test leader and participant. The study was thereafter administered as follows. Participants were instructed to watch the 66 second short film on the computer of the test leader. They were told to watch attentively and that they afterwards were to describe as detailed and as much of the film as they could remember. No further information was given. After having watched the film, participants were informed that they had just witnessed a serious crime and were asked to deliver the free recall (Swedish/English) of the depicted events in a credible way from their perspective as a witness. Participants were informed that the suggested time for the free recall was six minutes and that their statement would be recorded for further analysis. Participants were thereafter instructed to answer the 20 questions (whereof 15 suggestive), posed orally by the test leader, as correctly as possible. Once suggestibility had been administered the credibility questionnaire, dispersed with the manipulation check and demographic variables was distributed. Conclusively, the GCS was administered and participants were briefed. The test leader accompanied the participants throughout the entire procedure and answered questions if necessary.

Results
A series of independent sample t-tests were conducted, using Bonferroni adjusted alpha levels of .01 ($\alpha^* = \alpha/k$), to compare suggestibility, ego-depletion, credibility and free recall for participants in the English-speaking condition and participants in the Swedish-speaking condition. Main results suggest significantly higher ego-depletion levels in the English-speaking condition compared to the Swedish-speaking condition, with a medium to large effect size (Cohen’s $d = .66$). No significant differences were found between the groups regarding suggestibility and credibility in the present sample. Further results are presented in table 1.
Table 1. Independent sample t-test (two-tailed) on outcome measures between the English-speaking and Swedish-speaking group.

<table>
<thead>
<tr>
<th></th>
<th>English (n = 26)</th>
<th>Swedish (n = 25)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Suggestibility¹</td>
<td>4.38</td>
<td>2.79</td>
<td>2.68</td>
<td>2.35</td>
</tr>
<tr>
<td>Ego-depletion²</td>
<td>12.81</td>
<td>2.56</td>
<td>10.60</td>
<td>2.74</td>
</tr>
<tr>
<td>Credibility³</td>
<td>16.42</td>
<td>3.30</td>
<td>18.28</td>
<td>2.70</td>
</tr>
<tr>
<td>Time (min)⁴</td>
<td>2.92</td>
<td>1.15</td>
<td>2.86</td>
<td>1.23</td>
</tr>
<tr>
<td>Free recall⁴</td>
<td>26.58</td>
<td>8.28</td>
<td>29.12</td>
<td>9.29</td>
</tr>
</tbody>
</table>

p < .01  ° df = 49

Note. English coded as 0. 1) Suggestibility scores as measured by GSS, ranged from 0 to 10 (achievable maximum of 15). 2) Ego-depletion levels as measured by index ranged from 5 to 17 (achievable maximum of 21). 3) Credibility levels as measured by index ranged from 9 to 25 (achievable maximum of 28). 4) Total time for the free recall ranged from 1 to 7.20 minutes with correct remembered free recall scores ranging from 15 to 56 (achievable maximum of 104).

Hierarchical multiple regression analysis was performed to examine if language (English/Swedish) and credibility could predict suggestibility levels after controlling for compliance and age, both of which have shown to influence suggestibility levels. Preliminary analysis’ were performed to ensure no violation of the assumption of normality, linearity, multicollinearity or homoscedasticity. Age and compliance were entered in model 1 and showed statistical significance F (2, 48) = 5.21; p = .009 and explained 18% of variance in suggestibility, with age making a significant unique contribution to the model. After entry of language and credibility in model 2, total variance explained by the model as a whole was 30%, F (4, 46) = 4.91; p = .002 with an additional explained variance of 12% ΔF (2, 46) = 3.96; p = .026 by language and credibility together. To examine if free recall (time and score) could further serve as potential suggestibility predictors, they were entered in model 3. Results showed an additional simultaneously explained variance of 3% ΔF (2, 44) = .927; p = .403 by time of free recall and correct scored details in the free recall. In the final adjusted model, age and language were statistically significant, with age recording a higher Beta value β = -.53, p = .002 than language β = -.35, p = .012. Compliance fell short of statistical significance in all models, indicating that compliance in the present sample did not predict suggestibility. Results are presented in table 2.

To test whether differences in suggestibility levels, as a result of spoken language, was mediated by ego-depletion, regression analysis were conducted as recommended by Baron and Kenny (1986). Results of bivariate regression analysis showed that although language was associated with suggestibility levels, β = -.319, p < .05 and ego-depletion levels β = -.391, p < .05, suggestibility levels were not significantly associated with ego-depletion levels, β = .02, p = .89.
Table 2. Summary of hierarchical Regression analysis, with standardized beta values, for variables predicting suggestibility (n =51).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th></th>
<th>Model 2</th>
<th></th>
<th>Model3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SE B</td>
<td>β</td>
<td>SE B</td>
<td>β</td>
<td>SE B</td>
<td>β</td>
</tr>
<tr>
<td>Compliance</td>
<td>.09</td>
<td>.09</td>
<td>.09</td>
<td>-.006</td>
<td>.10</td>
<td>-.05</td>
</tr>
<tr>
<td>Age</td>
<td>.06</td>
<td>-.37*</td>
<td>.06</td>
<td>-.44*</td>
<td>.06</td>
<td>-.53*</td>
</tr>
<tr>
<td>Language</td>
<td>.71</td>
<td>-.35*</td>
<td>.72</td>
<td>-.35*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Credibility</td>
<td>.12</td>
<td>-.03</td>
<td>.13</td>
<td>-.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td></td>
<td></td>
<td>.44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free recall</td>
<td></td>
<td></td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>.18*</td>
<td></td>
<td>.30*</td>
<td></td>
<td>.33*</td>
<td></td>
</tr>
<tr>
<td>$R^2$change</td>
<td>N/A*</td>
<td></td>
<td>.12*</td>
<td></td>
<td>.03</td>
<td></td>
</tr>
</tbody>
</table>

Note. Language was coded as 0 for English and 1 for Swedish. *non-applicable

Discussion

The present study examined the effect of interviewee language on suggestibility, with the main hypothesis being that; participants who delivered the free recall in English would portray higher suggestibility levels due to ego-depletion. Positive results suggested that participants in the English-speaking condition were significantly more ego-depleted than participants in the Swedish-speaking condition. However, could merely marginally supportive results be found in difference of suggestibility between the English-speaking participants and the Swedish-speaking participants, after Bonferroni correction was made. Further results from regression analysis, indicated that ego-depletion had no mediating effect on the relationships between language and suggestibility, although language was found to be associated with both suggestibility and ego-depletion, in the present sample. It was further hypothesized that delivering the free recall in English would imply less accurate and detailed descriptions of the depicted crime as well result in shorter time spent on the free recall. Following this, participants’ self-rated credibility would therefore also be affected by interviewee language. Regarding free recall, no differences between the groups were obtained, whether in correct amount scored details nor in time spent on the free recall. Results further portray no differences in self-rated credibility and credibility was further not shown to be a significant predictor of suggestibility.
As this study has argued, speaking a second language is effortful and could lead to ego-depletion and therefore also influence suggestibility. Results show that language was associated with suggestibility and contribute thus with new evidence of that speaking a second language is effortful and may influence suggestibility. The fact that participants in the English-speaking condition were shown to be more ego-depleted is an important finding albeit not portraying higher suggestibility in this sample. Results do not align with previous findings by Otgaar and colleagues (2012), who in their study found associations between ego-depletion and suggestibility. It should though be stressed that results do not align after Bonferroni corrections and that language indeed was associated with suggestibility. Results are important since they point to the effects speaking a non-native language can have on both suggestibility and ego-depletion and the potential rendering consequences thereof.

As to why suggestibility was only marginally significant in the present study, it is of essence to consider the scale used to measure suggestibility. It should be noted that the scale was created by the author and had not been used elsewhere. The GSS is created to measure individual suggestibility, no matter of spoken language. It has been argued that portraying oneself in a credible way would affect suggestibility due to an increased employment of self-regulatory resources, indifferent of spoken language. The study aimed to induce ego-depletion in both conditions where the English-speaking participants would portray higher ego-depletion levels than their Swedish counterparts. Mean suggestibility scores for the English-speaking group were well in line with mean values for yield 1 in the general population (Gudjonsson, 2003; Hildebran Karlén, in press). However, did suggestibility scores in the Swedish speaking group not reach as high, which could imply that the used suggestibility measure was not successful enough. Although the scale presented acceptable Cronbach’s alpha level, it is arguable that the scale itself did not manage to encapsulate participant’s suggestibility. It should further be noted that the deployed suggestibility measure was not as extensive as the original version of the GSS. To assess total suggestibility, most studies measure several aspects of suggestibility including yield and shift. The original version of the GSS deploys a memory paradigm of a narrative upon which a free recall is requested, similar to the deployed crime scenario in the present study. After having delivered the free recall, participants are asked to do so again after a retention interval of 50 minutes. Participants are thereafter asked 20 questions intended to tap into the yield aspect of suggestibility. After having answered the 20 questions, participants receive negative feedback and are told that they have made a few errors and are asked to answer the 20 questions again. This step aims at the shift aspect of suggestibility. In the present study, only the yield aspect of suggestibility was measured. Although research is divided on whether shift is a reliable measure of suggestibility, the present study should be viewed in light of the sole measure of yield regarding suggestibility. Due to practical limitations as well as ethical considerations, negative feedback was for example not practiced in the study and therefore excluding the shift aspect. Bearing the above in mind, it would be premature to draw conclusions on whether there were any differences or not between the two groups, regarding suggestibility in a general more common sense.
Compliance measures in the sample were well in line with Gudjonsson’s (1989) reported mean value for the general population and did report even higher alpha levels than those most frequently obtained (e.g., Drake & Egan, 2017). Following this, it can be concluded that the deployed Swedish version of the GCS in the study was successful in measuring compliance. Given that compliance did not explain any unique variance in suggestibility it can further be assumed that compliance did not influence suggestibility levels in the current sample.

Although participants delivering the free recall in English reported a significantly higher degree of ego-depletion compared to their Swedish counterpart, it was not possible to establish whether suggestibility levels were due to the increased demand of self-regulatory processes and consequently ego-depletion. To establish mediation the following conditions must be met: The independent variable (language) must affect the mediator (i.e. ego-depletion), the independent variable (language) must be shown to affect the dependent variable (suggestibility) and last must the mediator (ego-depletion) be shown to affect the dependent variable (suggestibility) (Baron & Kenny, 1986). Due to that conditions for mediation were not met, ego-depletion could not be established as a mediator. As to why ego-depletion was not found to mediate language and suggestibility, there are potential reasons worth considering. First, the scale used to measure ego-depletion obtained only acceptable Cronbach’s alpha and is somewhat unconventional compared to more frequently used ego-depletion measures. Such include solvable anagrams, math or mental arithmetic, handgrip and modified Stroop (Hagger et al., 2012). Besides not being deployed in this study, due to practical reasons, the more frequently used measures are debated. Carter, Kofler, Forster and McCullough (2015), suggest in their meta-analysis of the depletion effect, that ego-depletion effects are dependent on the measurement used to operationalize it and stress the importance of using standardized tests to enable predictions of ego-depletion. The ego-depletion manipulation check used in this study is only one way to measure ego-depletion and it is possible that a different manipulation check could have generated other results regarding ego-depletion levels. Given that there were still obtained significant ego-depletion differences, between the two conditions, it is arguable that the deployed manipulation check was successful in capturing self-reported estimations of perceived difficulty, exhaustion and strain related to the performed depletion task. Further reason as to why mediation could not be established, could be that portraying oneself in a credible way and speaking a second language did not sufficiently engage self-regulatory resources to result in ego-depletion subsequently affecting suggestibility in this context. Given that the study was conducted in a university setting were English is a universal language, it can be argued that requesting participants to speak English was not as effortful as it could have been for individuals in a different setting with different backgrounds.

Time spent on the free recall task as well as the correct recollected details of the depicted events in the film did not differ between the groups, contrasting the results in Lindholm’s (2005) study. In general, time spent on the free recall was low for both groups compared to the suggested six minutes. Correct scored details were also substantially lower than possible to achieve by the scoring template. As to why, it can be speculated that independent of language, it was not evident to participants in which extent they were to recall the staged robbery. That is, participants did not
focus on as many details as required by the scoring template to reach high scores as well as an average time of six minutes, even though they were instructed to report as many details possible. It can be assumed that in a more real-life simulated interrogative setting the extent to which details are requested is more obvious. Further it should be noted that participation in the study was a required course moment for participants, leading to the assumption of that enthusiasm and seriousness not were at their potential peaks. Following this could potentially be linked to how participants rated their credibility.

In the present sample result suggest that there was not any difference in how participants rated their credibility. An inspection of the mean values in credibility indicate that neither group rated themselves as particularly credible. These results could be because of the setting the study was conducted in. Although the study aimed to resemble an interrogative setting as much as possible, fact remain that the study was conducted in a student environment between two students. Regarding credibility, it can also be assumed that a more real-life setting could yield higher credibility ratings amongst participants. Credibility was further not shown to be a potential predictor of suggestibility in the current sample. Results should further be interpreted considering the scale used to measure credibility, with a merely acceptable obtained Cronbach’s alpha. Studies have shown that when asked to portray oneself in a specific way it can influence susceptibility of suggestibility (e.g., Vohs et al., 2005). Deploying a stronger and more systematic measure of credibility, with more items, could potentially generate greater differences between the groups as well serve as a better predictor in suggestibility levels.

As previously mentioned, age differences within suggestibility levels are one of the most studied individual differences in suggestibility and studies show that younger children and older adults are more susceptible to suggestibility (e.g. Dukala & Polczyk, 2013). In the present sample of psychology students with a mean age of 26 years, age constituted the largest predictor in suggestibility levels. Findings show that younger age per se may be a predictor of suggestibility and that even in this rather age homogenous sample, younger participants were significantly more susceptible to suggestibility. Language was however still shown to significantly contribute to and predict suggestibility levels, indicating that although age differences matter, language does to. Results evoke questions as to what potentially happens in terms of development in the early 20s that could affect suggestibility and should encourage further research on age differences in more narrow age ranges.

Although the current findings are promising and inspiring some limitations remain. Overall limitations include the small sample size and future studies should be conducted on larger sample sizes to increase internal and external validity of the results. However, when looking at the strength of the obtained effect size, in differences within ego-depletion levels, results reveal that even with the current sample interviewee language was associated with ego-depletion. Another limitation is that only the yield aspect of suggestibility was measured. As standard measures of suggestibility generally are more extensive, future studies should replicate the present study measuring all aspects of suggestibility in an interrogative setting. Furthermore, memory capacity was not tested for in the present study, although it
has been shown to play a role in suggestibility (Gudjonsson, 2003). It is arguable that such a measure could have contributed with useful information in terms of a potential predictor or confounder of suggestibility. Hagger and colleagues (2010) mention skill, fatigue, motivation, self-efficacy and affect as alternative explanations to observed self-regulatory failures measured in ego-depletion experiments. Alternatives which are believed to compete more or less with ego-depletion and it can only be speculated in whether they could serve as alternative explanations in the present study. To further generalizability of the results of that speaking a second language may increase suggestibility, the above mentioned alternative explanations should be taken into consideration and tested for in future studies. Future studies should try to replicate results with a different sample, including different ethnical backgrounds, in a more real-life forensic setting. Nevertheless, does the study contribute with findings highlighting potential predictors of suggestibility to be regarded in future more extensive research projects.

The present study used adjusted Bonferroni alpha levels to decrease the risk of type I errors and multiple comparison problems. Bonferroni corrections are frequently argued to be conservative as well can they increase the risk of type II errors. Results from the obtained study should be interpreted with this in mind. Using Bonferroni correction resulted in marginally significant suggestibility differences between the English-speaking participants and the Swedish-speaking participants and conclusions thereof should hence be made with caution.

Conclusion
As made evident, it is of essence to consider under which conditions the present study was conducted in. One strength of the study was the used filmed with the staged robbery. The film very much resembled a real-life situation facilitating the participant to engage in their role as a witness, rendering in an increase of ecological validity. It should be kept in mind that the environment in which the study was conducted probably is experienced as more safe and familiar than that of an interrogative forensic interview. The fact that results point to a difference in ego-depletion levels, depending on interviewee language, in this context, adds to the strength of the study as well to the insight of the delicate nature of factors influencing ego-depletion and not only in an obviously stressful and taxing situation. Interrogative suggestibility is a complex area which, as portrayed, can be influenced in many ways. The effect of speaking a second language on ego-depletion and suggestibility has not previously been studied and results contribute with valuable information to consider and take into account when conducting an interrogative interview. The study contributes to the evidence that engaging in self-regulatory processes makes people more prone to ego-depletion and thus also potentially suggestibility. Conclusively does the study show that ego-depletion levels increase when an interview is conducted in a second language compared to in one’s native language. Not being able to take part in an interrogative interview in one’s native language may potentially, and in worst case scenario, lead to less accurate and potentially false testimonies.
References


