Consistency and Social Identification: A Test-Retest Study of Flashbulb Memories Collected on the Day of the 2016 Brussels Bombings

In press, Memory

Aline Cordonnier [1] and Olivier Luminet [1,2]
[1] UCLouvain, Research Institute for Psychological Sciences, Belgium
[2] Fund for scientific research (FRS-FNRS)

Word count of main text only (including tables): 10328

Total word count: 12024

Abstract

Flashbulb memories are consistent and vivid representations on the way people learnt of an important, surprising and consequential event. We had two main aims for this study Our first aim was to evaluate the consistency of flashbulb memories as accurately as possible (by collecting data on the day and using an interview method at retest) and contrast these findings with other operationalisation such as vividness or confidence but also with event memory. Our second aim was to capitalise on the particular case-study that is Belgium to examine social identification. Within a few hours after the terrorist attacks in Brussels in March 2016, a small sample of students completed a questionnaire on the way they learnt of the event and their knowledge of it. Retest data was obtained fifteen months later, through an in-depth interview. Our results show a relatively high consistency over time as well as high vividness and confidence for their memory of the reception context. We also measured participants' identification at three levels: local (Brussels) – national (Belgium) – supranational (Europe). In the particular context of the Brussels bombings, social identification with Brussels and Europe correlated with measures of flashbulb memory while social identification with Belgium did not.

Keywords: flashbulb memory, consistency, social identity, vividness, immediate data collection

Consistency and Social Identification: A Test-Retest Study of Flashbulb Memories Collected on the Day of the 2016 Brussels Bombings

March 22, 2016. Time is 7.57 a.m. and Brussel's airport is already buzzing with holidaymakers, businessmen and airport workers. None of them knows that two explosions are about to tear down the departure hall. Within minutes after the attack, news and social media are reporting live on it, showing photos of debris covering the floor amidst abandoned suitcases, windows shattered under the power of the blast and shocking photos of people injured or running away. Just over an hour later, a third explosion occurs in a metro at Maelbeek station situated not too far from the European Commission headquarters of Brussels. This unexpected attack will kill 35 people and injure over 300 others. For Belgians, but also people across the world, the 22nd of March 2016 was a highly emotional day and will have long-term consequences. And for us, flashbulb memory (FBM) researchers, it was also a unique opportunity to collect data on the day of the event.

Beyond this unexpected occasion, we decided to pursue two aims in this study. Our first aim was to evaluate consistency of FBMs as accurately as possible and contrast these findings with other operationalisation of FBMs such as vividness or confidence but also with event memory (EM). Our second aim was to capitalise on the particular case-study that is Belgium to examine the role of social identification with different groups in the formation of FBMs.

Operationalisations of Flashbulb Memories and Improved Consistency

Flashbulb memories are generally defined as detailed memories of the reception context where one learns about a surprising and consequential event (Brown & Kulik, 1977). Represented as a mental photograph of the instant it happened and experienced with "almost perceptual clarity" (Brown & Kulik, 1977, p. 73), they are believed to be consistent, vivid and long-lasting. People also tend to be highly confident about the accuracy of their memory and

believe they can recall the "exact" moment they learnt about the event (Talarico & Rubin, 2007). But to this day, the literature is still in search of a consensus on how to conceptualise and operationalise FBM.

The long-lasting question on what are FBMs and whether they should be considered different from ordinary autobiographical memory is more than never on-going. One of the major difficulties resides in the way we measure them (Talarico & Rubin, 2018). In the literature, we find four principal operationalisations of FBM: specificity or number of details, consistency, vividness and confidence. Others, such as accuracy or autonoesis (the feeling of reliving the event) are more rarely examined. Yet, we do not know whether these operationalisations act as a single construct or whether they represent different dimensions. We conducted a short review of empirical studies of FBM in the last ten years[1] to examine the prevalence of these four FBM attributes. Out of the 38 articles selected, the average number of FBM attribute was just over 2 and only one third of studies measured three or more different attributes (see Table 1). Confidence (which includes measures of "belief in accuracy") was measured in 82% of studies, consistency was measured in 71%, specificity in 68% and vividness in only 57% of examined studies. This quick analysis already reveals major difference in how FBMs are assessed. More importantly, the different attributes are rarely measured together and when they are, the correlational patterns are rarely reported and seem to vary from one study to the next (for a discussion, see Dumas & Luminet, 2016). These different operationalisations of FBM hinder our capacity to properly compare studies with one another. In this research, we endeavour to be more exhaustive in our estimation of FBM by using multiple attributes and examining how they relate to one another.

[1] We used the database « PsycInfo » to run this review on November 2020. We search for all scholarly journals with the term "flashbulb" anywhere in the manuscript after January 2010 (we decided than 10 years would give us already a good overview of the recent literature on the subject). Seventy references were found and checked one by one. We removed from this list theoretical articles, literature reviews, comment pieces, errata and articles not on flashbulb memories.

One important operationalisation of FBM that has received very little scrutiny is that of accuracy. As discussed above, people tend to be highly confident in the accuracy of their memory, however, from a research standpoint, it is almost impossible to evaluate whether the details recalled are actually accurate. In some rare cases, researchers have been able to test the validity of some claims with archival data. For example, Berntsen and Thomsen (2005) compared participants' recall of the weather when Denmark was invaded and when it was liberated during the Second World War with official data. Whereas two thirds of veterans could correctly recall the weather, some memories were negatively biased when discussing the invasion and positively biased when discussing the liberation. Similarly, even though George W. Bush indicated watching on TV the first plane hit the World Trade Center on September 11, 2001, Greenberg (2004) highlights the fact that his memory cannot be accurate as, at that time, such footage did not exist.

Unfortunately, this method only works for a small range of specific details or for event memory details, but not for the many idiosyncratic details we generally find within descriptions of the reception context. To evaluate the accuracy of these elements over time, researchers have instead relied on measures of consistency across time. Yet, to be able to use consistency as a proxy for accuracy, researchers would need to obtain reliable baseline details on the exact circumstances in which participants learnt the news. But by definition, events likely to trigger the formation of FBMs are highly surprising – and thus, highly unexpected. By the time researchers have reached participants, a certain number of hours or days have passed. Autobiographical memories always rely on a reconstructive process and, as such, with the passing of time – especially when the memory is likely to be told and retold – the details of the event are modified, elements are added while other are forgotten. These "natural variations" can even be found for personally significant events, such as one's first date or 18th birthday party (Temler et al., 2020).

In this study, we attempted to improve our measure of consistency by implementing two methodological changes: we collected baseline data for our test-retest design *on the day of the event* and we used an *interview style* to collect data at retest. We explain our rationale in the following sections.

Collecting Baseline Data on the Day

Researchers have always endeavoured to collect baseline data as soon as possible after the event happened when using a test-retest design. Hence, most data have been collected at best within the first two or three days. However, memories tend to consolidate overnight (Schonauer et al., 2015; Stickgold & Walker, 2013). Diekelmann and Born (2010) define consolidation as "the process that transforms new and initially labile memories encoded in the wake state into more stable representations that become integrated into the network of pre-existing long-term memories" (p. 114). Seminal studies have reported that emotional memories are particularly influenced by sleep (e.g., Hu et al., 2006; Lipinska et al., 2019; Payne et al., 2008). This consolidation process is important when examining the accuracy and the consistency of FBMs.

With regards to the accuracy of autobiographical memories, we know that a selection is done during sleep. Our mind selects what to recall, what to forget and reorganises our memories to fit pre-existing schemas (Landmann et al., 2014; Stickgold & Walker, 2013). In the case of emotional memories, background details are particularly susceptible to forgetting during the consolidation process (Payne et al., 2008). It is for this reason, amongst others, that eyewitnesses should be interviewed on the day of the event to get the most accurate testimony (Tuckey & Brewer, 2003). Therefore, to examine whether FBM are accurate over time, researchers should strive to collect baseline data on the day of the event, before the consolidation phase has occurred.

With regards to measures of consistency, as memory becomes more crystallised after a night of sleep, we would expect higher consistency of FBMs when the first data collection is not on the day of the event. In accordance with this hypothesis, Winningham et al. (2000) showed that longer delays between the event and data collection increased the consistency of FBMs. To the best of our knowledge, their study on O.J. Simpson's trial verdict is the only one who collected data *on the day of the event*. However, the studied event was not "unexpected" (the result of the verdict might have been surprising, but people were expected to hear the verdict on that day) and its consequentiality for most people might have been limited. We thus regard our study as the first one to assess FBM on the day the event happened for an event that is scoring high on the important parameters of FBM: importance, unexpectedness and emotionality.

Collecting Retest Data in an Interview

Consistency is measured by comparing participants' answers to the canonical features (e.g., with whom, where, when, doing what) at two different points in time (with the first time point hopefully collected closely after the event's occurrence). The difficulty of the task resides in deciding when an answer is consistent or not, especially when few details have been shared. In the vast majority of studies, these answers are collected through the use of questionnaires. Yet, open-ended questions in questionnaires can lead to a significant amount of unclear or missing data (Reja et al., 2003). Answers also tend to be short and unprecise, making the estimation of consistency even more arduous. To the best of our knowledge, no study has actually used in-person interviews to collect retest data. Interviews provide the opportunity for researchers to ask follow-up questions or request clarifications. Thus, in order to get a more detailed and reliable evaluation of consistency across time we used in our study a detailed interview approach.

Contrasting FBM with Event Memory (EM)

Recent research has started to take an interest in the accuracy, and more rarely the consistency, of EMs – that is memories of the facts about the studied event. Some authors considered that FBMs and EMs belong to different memory system, the former relying more on episodic memory and the later relying more on semantic memory (Bohannon & Symons, 1992; Curci et al., 2001; Finkenauer et al., 1998; Pezdek, 2003). FBMs are, after all, autobiographical memories; the episode of receiving the news is directly experienced and encoded in memory.[2] However, not all factual details of the FBM-triggering event are necessarily encountered or learnt, especially on the day of the event, as information might be scarce or sometimes even contradictory. Detailed knowledge of the facts also depends on media exposure. For this reason, accuracy – and consequently, consistency – of EM can evolve depending on the extent of media coverage and the individual's attention to these types of media but independently to the strength of FBMs (e.g., Hirst & Phelps, 2016; Koppel et al., 2013; Tinti et al., 2014).

The Role of Social Identity (SI)

The second main aim of our research is to investigate the role of social identification in FBMs. Research in FBM has often emphasised emotional and cognitive predictors. For example, in their emotional-integrative model, Finkenauer et al. (1998) proposed two paths that could lead to the formation of FBMs: directly through the activation of the appraisal of novelty and the emotion of surprise or indirectly through the appraisal of importance of the event, which triggers an emotional responses that can then increase rehearsal and EM, finally leading to the formation of FBMs. However, Luminet and Curci (2009) showed that the route through which FBMs are formed depends on the activation of the individual's social identity.

[2] On a side note, we want to ascertain that all autobiographical memories, including FBMs, also rely on semantic knowledge about the self and personal scripts.

In her review on the role of social identity in the formation of FBMs, Berntsen (2018) argued that the social and cultural relevance of public events played a more important role than personal relevance. In accordance with this view, Talarico et al. (2019) found that Germans were more likely to have formed FBMs of the Fukushima disasters than were Dutch or Danish participants as the event was not congruent with their social group's beliefs about the dangers of nuclear energy. Nevertheless, research has generally examined social identity by using national identity (in contrast to other groups membership) as a variable and only in a dichotomous way (the person is either Danish or is not Danish). In other words, social identification has mainly been based on an external qualification and not on the personal evaluation of respondents' level of social identity with certain groups. Yet, the level of personal identification with one's country is not always a given. In certain parts of the world, other groups might be more relevant than the nation when evaluating social identity. For example, many Hong Kongese do not identify themselves as Chinese and many Catalans do not like to be seen as Spanish.

In two recent studies, Merck et al. (2020) proposed two different ways of examining the role of social identity in the world of sport. In their first study, they collected a measure of "baseball fandom" and correlated it with attributes of FBM in the context of the 2016 Major League Baseball World Series. They found that strong baseball fans were more likely to have consistent memories of the personal circumstances in which they learnt about the outcome of the game one year later than other participants. For their second study, they measured fandoms but for a particular team. They found that team-specific fandom positively correlated with confidence, but only when their team lost. The level of fandom was thus not relevant for a win, maybe because it was already consequential enough to matter, but it was for a loss. Overall, what their study highlights is the need to refine our measures of social identification when examining event and personal memories of social events by using continuous measures

and including different levels of social identity. Luckily for us, Belgium is a unique case study when it comes to examining social identity.

The 2016 Brussels Bombings – Belgium as a Unique Case Study

On Mars 22, 2016, we had an exceptional opportunity to collect data for a FBM study. The Brussels bombings happened early in the morning and that day, the second author war teaching a class of undergraduate psychology students. This combination meant that we could collect baseline data on an FBM questionnaire only hours after the event had occurred and before participants could forget details about how they learnt the news or consolidate their memory while sleeping, a feat that the second authors had always wanted to achieve. As mentioned earlier, this was a unique event as it was the most severe terrorist attack in Belgium that led to many people dying or being injured. Furthermore, it was very unexpected as it was only the second terrorist attack in Europe (Paris being the first) of the (sadly) long list of terrorist attacks that occurred in Europe around the same time. It had a major impact on people's life at the time and in the following years, with new rules and regulations appearing after that. The event was novel, surprising, consequential, and highly emotional.

Another reason why this research is quite unique lies in the Belgian cultural and political context. Belgium is an unusual case study when measures of social identity are considered. Due to its complex historical past and convoluted political structure, Belgium has always been a divided country and hardly fits the mould of global national pride (Luminet et al., 2012). When it comes to social identity, some Belgians might feel closer to their region or their city than to the Belgian nation (Klein et al., 2012). But Brussels is also often seen as the *de facto* capital of the European Union and many Belgians feel like being European is part of who they are, sometimes even more so than being Belgian (Duchesne & Frognier, 2007). Within this particular context, we thus have three level of social identity to consider: local – national – supra-national.

Furthermore, the intricacies of the Belgium system allow us to evaluate two types of social identification: attachment and pride. People can feel attached to a place that they do not belong to officially (e.g., a Walloon can still feel attached to Brussels and vice versa), whereas pride relates to one's official affiliation (I cannot be proud to be from Brussels if I am not from Brussels). These two modes should be positively correlated, however they should still be considered as distinct (Roccas et al., 2006).

In summary, we designed this test-retest study with two principal aims in mind: 1) to obtain an improved measure of FBM consistency over time by collecting baseline data on the day of the event and using an in-depth interview method to collect retest data, and 2) evaluate social identification at different level.

Method

Participants

We recruited two distinct groups of people for this study: one *Immediate Group,* tested at two points in time (T1 and T2), and one *Control Group,* tested only once (T2). For the *Immediate Group,* 20 undergraduate students completed the questionnaire at the end of their psychology class at the UCLouvain University in Belgium. One student did not participate in the second phase of the study and was thus removed from the sample. From the 19 participants left, 17 were Belgians and two from a different citizenship. We decided only to keep non-Belgian participants who indicated a strong attachment to the country by providing a score of minimum 4 out of 7 to the statement "I feel attached to Belgium". Following this criteria, one participant was removed. The final 18 participants were 14 females (78 %) and 4 males, aged between 20 and 26 years old (M = 21.33, SD = 1.68).

In order to evaluate whether our *Immediate Group* was representative of the student population, we collected data from another similar sample that served as a *Control Group* at T2. Sixty-two participants completed the questionnaire, of which 54 were Belgians. We used

the same process as for our immediate group and checked non-Belgian participants' attachment to Belgium. One participant was removed, leaving us with a sample composed of 42 females (70 %), 18 males and one other, aged between 18 and 26 years old (M = 22.08, SD = 1.82).

Procedure

During their psychology class, on March 22nd, 2016 around 4.15pm, students of the second author were invited to complete a pen and paper questionnaire on the events that had happened earlier in the morning (the first bombing started just before 8 a.m.). They were told the goal of the research was to better understand how psychological and emotional processes impacted the memory of public events and that their answers would remain anonymous. Students who did not wish to complete the questionnaire were free to do so. None of them selected this option.

In June 2017, so about one year and three months after the attacks, participants answered the same questionnaire in a face-to-face interview. These interviews were done by pairs of master students (three students helped us collect this data) and were recorded and transcribed. This mode of presentation allowed the interviewers to make sure participants properly understood all questions and measures, as well as provided some flexibility as participant could go back to previous questions if an answer had come back to mind. Importantly, the interviewers could probe for more details when necessary, especially with regards to the FBM attributes, in order to clarify answers which would, in turn, allow later on a better estimation of consistency over time.

Finally, in a parallel setting, an online version of the questionnaire was circulated via word of mouth, emails and social networks through students circles from mid-June to the beginning of August 2017.

Material

At each point in time and for each group, we used an adapted version of the flashbulb memory questionnaire from Luminet and Curci (2009). However, the questionnaire was presented using different modalities, as described in the procedure.

Measures

The questionnaire was organised in 5 main sections: event memory, flashbulb memory, predictors of these two memory dimensions, background knowledge and social identity. However, due to our small sample size, we decided to only examine variables associated with our a priori research questions, and thus, for the sake of brevity, we only describe here these variables (i.e., EM, FBM and SI), in their presentation order.

Before we delve into our variables, it is important to note that we relied on a coding scheme for many of them. This coding scheme was based on previous studies (Curci & Luminet, 2006; Curci et al., 2001; Luminet & Curci, 2009) and adapted for this context. Criteria for attributing different scores were based on *a priori* decisions and on analysis of the frequency of the answers. We also decided to compute rather strict scores. Each coding scheme will be described here under and examples will be provided for clarity when necessary.

Event Memory. We computed three scores related to event memory: accuracy, consistency, and confidence. Six open-ended questions measured participants' memory of the two terrorist attacks, four about the airport attack: (a) the time of the day, (b) the exact location, (c) the number of explosions, (d) the number of people killed; and two about the metro attack: (e) the time, (f) the number of people killed.

EM accuracy was measured at both points in time and for both groups. Judges recorded the accuracy of each response by giving it a score between 0 and 2 for questions (b) and (c), and between 0 and 3 for the other questions as their range of potential responses was

wider (with the highest score indicating a highly accurate response). Scoring criteria were decided a priori based on the computation of journalistic sources in order to determine the most reliable answer to each question. It was applied in a systematic way and therefore, did not require inter-rater reliability to be computed. To illustrate, for the question "at what time precisely did the airport explosions occurred?" (the correct answer is at 7h58), participants received 3 points (high accuracy) if they provided a specific time between 7h50 and 8 a.m.; 2 points (moderate accuracy) if they provided a specific time between 7h30 and 7h49 or between 8h01 and 8h30; 1 point (low accuracy) if they provided a specific time between 6h30 and 7h29 or between 8h31 and 9h30, or gave a correct but approximate answer such as "between 7 a.m. and 8 a.m."; and 0 points (not accurate) for any other answer. To ease understanding, we then turned these scores into percentages to obtain *EM accuracy* score for each variable.

An *EM consistency* score was also obtained by comparing answers provided at T1 and at T2 for the immediate group. Three independent judges attributed either a score of 2 for responses that were almost identical or that became more specific at T2 (to account for the likelihood of providing more specific answer during an interview compared to a pen and paper questionnaire), a score of 1 for responses that were similar but expressed in a different manner or for a loss of information, or a score of 0 when there were no similarities (see also Luminet & Curci, 2009). For example, a participant received a consistency score of 1 as she indicated at T1 that the metro bombing happened at "8.30 a.m., maybe 9 a.m." but said at T2 "I think it was an hour after [the airport bombing], so I would say 9 a.m.". Inter-rater reliability was done over three complete participants and provided a reliable Fleiss Kappa of 0.825, which is considered an "almost perfect agreement" (Landis & Koch, 1977).

Finally, *EM confidence* was measured in the questionnaire presented in T2. At the end of the EM section, a single question asked participants to rate the level of confidence in their answers for this section on a scale from 1 (no at all confident) to 7 (completely confident).

Flashbulb Memory (FBM). We computed four scores for FBM: canonical consistency, details consistency, vividness, and confidence. Six open-ended questions corresponding to the canonical categories of FBM used in previous studies (Curci & Luminet, 2006; Curci et al., 2001; Finkenauer et al., 1998; Luminet et al., 2004) measured participant's recall of the circumstances in which they first learnt about the bombings: (a) the time of the day, (b) the exact location they were in, (c) the informant (TV, radio, newspaper, social media, another person), (d) the other people present, (e) their ongoing activity and (f) what they did after hearing the news.

FBM consistency was measured only for the immediate group. Three independent judges compared participants' responses given in T1 with the ones provided in T2 and provided a *FBM consistency* score between 0 and 2 following the same procedure as explained for the *EM consistency* score: a score of 2 for responses that were almost identical or became more specific at T2, a score of 1 for responses that were similar but expressed in a different manner or for a loss of information, or a score of 0 when there were no similarities. For example, to the question "where were you when you heard the news?", one participant got a score of 2 as she answered at T1 "I was alone" and at T2 "I was alone in my bedroom, my roommates were in their bedroom, but I did not see them that day before I heard the news". On the other hand, she got a score of 1 for the question "What were you doing when you heard the news?" as her answer at T1 was "I was turning on my computer to continue writing a homework" but said at T2 "I was working on a project for my bachelor 3 diploma that we had to hand in. I remember well. We had to give it back by the end of the year.". In this case, the response was similar, but she did not clearly state she was on her computer, so it

was expressed in a different manner. Inter-rater reliability was done over three complete participants and provided a Fleiss Kappa of 0.68, which is considered a substantial agreement (Landis & Koch, 1977). Disagreements between the judges were solved through discussion. All other questionnaires were then rated by individual judges.

FBM vividness was measured directly after their response on the canonical categories. Participants evaluated the vividness of their FBM by saying whether their answer relied on a "precise memory" (4), a "recall" (3), a "feeling of familiarity" (2) or a "guess" (1). The explanation of these terms was provided at the start of the questionnaire as followed:

> *"A precise memory*: you remember a detail with images or feelings. You hear or see again in your mind the recalled detail or you experience it–even though it might not be with the same intensity–the same way you did when you first heard the news, so during the original experience.
>
> *A recall*: you know that the memory you recall resembles that of the original experience.
>
> *A feeling of familiarity*: you have a feeling that your memory resembles in some ways the original experience.
>
> *A guess*: you simply guess your answer."

From these questions, an overall *FBM vividness* score was obtained by averaging participants' answers in order to get a score on a scale from 1 to 4.

FBM details consistency was also computed only for our immediate group. Participants were asked to provide up to five details about their environment or the situation they were in when they learnt about the attacks. We compared these details across the two time points and coded for similarities in order to get an *FBM details consistency* measure. Each T2 detail was thus coded as "repetition" (details that were described almost identically at T1 and T2), "similar" (details that were described in a similar but not identical way, for

example someone said it was sunny at T1 and explained the weather was nice at T2), "contradiction" (details on the same topic but that could not both be true, for example someone said remembering "looking down at the floor, and it was clean" at T1 but indicated remembering "looking down at the floor, and it was dirty" at T2). Two judges scored all participants and had a very high level of agreement, $\kappa = .814$. All disagreements were resolved through discussion.

Finally, *FBM confidence* was measured at the end of the section at T2 similarly to *EM confidence*. Participants rated the level of confidence in their answers for the FBM section on a scale from 1 (no at all confident) to 7 (completely confident).

Social Identity (SI). At the second data collection time, six Likert-scale questions measured participants' social identification at three distinct levels: local (Brussels), national (Belgium) and supra-national (Europe). For each of these identification levels, participants indicated whether, on a scale from 1 (not at all) to 7 (completely), they felt (a) attached to Brussels/Belgium/Europe and whether (b) they were proud of being from Brussels/Belgian/European. However, after the data collection, we realised that many participants were not from Brussels as they indicated so during the interview when asked the question "are you proud of being from Brussels" (9 clearly said they were not, 3 gave unclear answers, such as a laugh after the question followed by a really low answer, and only 6 clearly indicated being from Brussels). Thus, we decided to remove this item from our analyses.

Results

Representability of our *Immediate Group*

The first important step is to examine whether our *Immediate Group* sample – which is relatively small – is representative of what we would get with a larger sample of similar participants, our *Control Group*. To do so, we decided to conduct Bayesian t-test analyses for

independent samples. Bayes factors allow us to quantify the evidence for the null hypothesis compared to the hypothesis that our two groups are different (Kruschke, 2013). We also decided to conduct these analyses across all measured variables and their composite scores collected at T2. Including all variables and not only the ones examined within the manuscript provides a stronger test of similarity across our two samples.

Overall, we ran 52 comparisons, using a default Cauchy prior width of r=0.707. We report the BH01 (the Bayes factor in favour of the null hypothesis) in Table 2. Overwhelmingly, our results provide anecdotal (BH01 = 1 to 3) to moderate (BH01 = 3 to 10) evidence that the data is more likely to be overserved under the null hypothesis (Jeffreys, 1998). In only two cases did we have evidence that the data is more likely to be observed under the hypothesis that the groups are different. Participants in the *Control Group* could tell more accurately at what time exactly the metro attack occurred. They also provided a stronger rating of vividness when recalling what they did after they learnt the news. Setting aside these minor differences, we have strong grounds to believe that our *Immediate Group* is representative of a broader similar population, however it cannot be generalised to other groups with different socio-demographic attributes (such as age or education levels).

Flashbulb Memory

We measured four operationalisations of FBM: canonical consistency, details consistency, vividness and confidence.

Participants' *canonical consistency* score for their answers about the canonical variables of FBM from T1 to T2 was relatively high, with a mean of 1.24 (*SD* = 0.48) out of 2 points. Most participants could provide many details about the reception contexts and 72.22% of the time, these details were either similar (score of 1) or identical (score of 2) to their original answer. In Table 3, we see that participants were particularly precise when describing

the exact place they were at the time and who else was present. On the other hand, they generally could not recall with exact precision the time they learnt about the bombings.

Participants had the opportunity to add up to five specific details about the environment they were in when they heard the news. They provided on average three to four details at both points in time (T1: $M = 3.33$, $SD = 1.68$; T2: $M = 3.61$, $SD = 1.65$). With regards to *details consistency*, 22.19% of details provided at T1 were described identically at T2 while 25.94% were described similarly at both point in time. Combined, this means that 48.13% of details freely provided at T1 were also spontaneously provided at T2 in a consistent way. Participants contradicted themselves in only 11.04% of cases. The other instances were news details (40.81%). In other words, if we only consider details that were present at both time points, we see that participants were consistent in 85.42% of cases.

The *vividness* of participants' answers to the FBM attributes was very high at T1 ($M = 3.88$, $SD = 0.23$, on a scale from 1 to 4) and remained high at T2 ($M = 3.58$, $SD = 0.57$). There was no significant difference between the two points in time (when we maintain a conservative alpha of .01 to avoid type I errors), $t(17) = 2.60$, $p = .019$, 95% CI [0.06, 0.54], $d_z = 0.61$. It is worth noting that at T1, for almost 90% of their memories about the reception context, participants indicated they felt they were mentally experiencing it the same way they did when they first heard the news, and that was still the case for more than 73% of memories at T2.

Finally, participants indicated feeling relatively *confident* in their answers for the FBM canonical attributes (M = 5.44, SD = 1.50) on a scale from 1 (not very confident) to 7 (extremely confident).

Relationship Between Operationalisations of Flashbulb Memory

To increase our understanding of the relationships between the different FBM attributes, we examined the inter-correlation matrix of our four measures of FBMs: *canonical*

consistency (percentage), *details consistency* (combined percentage of same or similar details calculated on the details presented at both T1 and T2)[3], *vividness* at T2 (score ranging from 1 to 4) and *confidence* at T2 (score ranging from 1-7). In Table 4, we see that most operationalisations correlate with one-another except for *confidence*, which does not correlate with any other FBM attribute. We should thus be careful when examining other variables in relation to FBM confidence ~~our~~scores. Furthermore, the correlations between *vividness* and *canonical consistency* are not particularly high for such a small sample, indicating that they do not represent a single construct of FBM and should be examined as separate entities.

Event Memory

Participants' knowledge of the details of the two terrorist attacks was low with an *accuracy* average score of 70.25% at T1 and 49.85% at T2, indicating an overall significant decrease between the two time points, $t(17) = 3.93$, $p < .001$, 95% CI [9.45, 31.35], $d_z = 0.93$. However, this result highly depends on the type and difficulty of the selected questions.[4]

The overall *consistency* of their answers, measured on a scale from 0 (not consistent) to 2 (very consistent) was also relatively low between the two time points ($M = 0.75$, $SD = 0.30$). Participants were the most consistent about the number of explosions at the airport ($M = 1.5$) or the precise location of the attack at the airport ($M = 1.28$) but were the least consistent about the death toll at the airport ($M = 0.17$) or at the metro ($M = 0.41$), as well as the exact time of the metro attack ($M = 0.39$). The consistency of their answers for the exact time of the airport bombing was slightly higher though ($M = 0.72$), though slightly easier to

[3] Alternatively, we could have used the percentage of same or similar details calculated on the *total* number of details provided at T2 (so including new details that were not discussed at T1). However, we decided that it was a less accurate measure of consistency (as new details are neither consistent nor inconsistent). Furthermore, the correlations between this version of the detail consistency measure and other FBM operationalization were similar, although slightly less strong.

[4] We also examined each variable separately. We found that this decrease in accuracy is mainly due to lower accuracy at T2 for questions related to the number of people killed at the airport (18.52%) or in the metro (38.89%) and the time of the attacks in the metro (29.63%). With regards to the number of people killed, it is worth noting that on the day of the event, media outlets reported death tolls separately for the two attacks, whereas a few days after, official reports – and thus news media – reported only a combine death toll for both attacks.

remember as it was almost at 8 a.m. on the dot. Ultimately, it is worth noting that consistency was lower for EM than FBMs; $t(17) = 3.57$, $p < .001$, 95% CI [0.20, 0.80], $d_z = 0.84$.

Finally, at T2, participants indicated on a scale from 1 (not very confident) to 7 (extremely confident) feeling not too *confident* overall in the answers they gave in the event memory section ($M = 3.44$, $SD = 1.29$), and this level of confidence was significantly lower than their confidence in their answers for the FBM attributes at T2; $t(17) = 3.70$, $p < .001$, 95% CI [0.17, 1.83], $d_z = 0.87$.

Social Identity

Participants' attachment to Belgium was slightly stronger than to Europe ($p = .021$) and much stronger than to Brussels ($p = .001$; see means and standard deviations in Table 5). They also were slightly prouder of being Belgium than being European ($p = .031$). The inter-correlations between all these items are medium to high (ranging from .261 to .748) with only the correlation between one's attachment to Belgium and to Europe bellow .300. However, even with high inter-correlations, we will examine attachment and pride separately as we have no measure of pride related to Brussels.

The second main objective of this study was to investigate how the different levels of social identity (local: Brussels – national: Belgium – supranational: Europe) are associated with the attributes of FBMs and EMs. We can see in Table 6 that participants' attachment to Brussels correlates (>.300) with all FBM attributes. Attachment to Europe correlates with most attributes at the exception of how confident participants were in their answer. However, participants' attachment to Belgium does not correlate with any FBM attribute. The pattern of responses for how proud participants were of being Belgian or European is similar to that of attachment. Overall, 10 out of the 12 correlations between Brussels or Europe SI and FBM attributes are above .300 while none of the 8 correlations between Belgium SI and FBM

attributes are. Finally, our results do not reveal any strong links between social identity and EM.

Discussion

The study of FBM is an unusual field of research. Even when researchers have particular questions they would like to investigate, in most cases, data collection cannot be planned. And when an event of national or international significance occurs, researchers need to act fast. The wheels are set in motion: designing a questionnaire, finding participants, collecting data; everything is done as quickly as possible to preserve the memory of the first few moments. This issue of timing is central; the accuracy of baseline data – needed for the estimation of consistency over time – depends on it. This is why most studies of FBM tend to collect data within the first few days of the event, yet no one has ever managed to have data collected *on the day* of a highly emotional, surprising and important event, before memory is consolidated by sleep. In this study, this is precisely what we achieved.

When the 2016 Brussels bombings occurred, we collected data that very day, measuring multiple attributes of FBM and EM. We decided to use students from one of the author's class as participants since they were readily available. Ironically, because of the nature of the event (and in particular because of the second bomb attack on the Brussels metro), public transport was disrupted and many students could not make it to class, which resulted in a relatively small sample of participants. Our results should thus be seen as exploratory. Nevertheless, we believe some of the issues brought forward here are worthy of discussion and provide avenues for future research.

Improved Consistency Measure of FBMs

Our first aim was to obtain an improved measure of FBM consistency using a test-retest design and contrast it with other operationalisations of FBMs. We used two means to obtain this improved measure of consistency. First, we collected baseline data on the day of

the terrorist attacks to reduce forgetting and consolidations effects. Second, participants completed the retest in a face-to-face interview to allow more thorough descriptions of their FBMs.

Unsurprisingly, most participants were able to provide in-depth details about the canonical features of FBM when asked about how they learnt about the bombings and reported perceiving these details with high levels of vividness at both time points. They also felt relatively confident about the accuracy of their responses on the FBM attributes section more than a year after the event. What is worth investigating, however, is whether this belief in high accuracy corresponds with good consistency (used as a proxy for accuracy) over time. This seems to be the case. We found that almost three quarters of responses seemed consistent in that they shared enough similarities with the original answer given by participant on the day of the bombings and did not contradict themselves.

Yet, perfect consistency highly depends on the type of questions and the strictness of the coding system. In our study, participants could precisely recall the exact place they were in when they heard about the bombings and the other person/people present at the time, but they usually were more imprecise or did not know the exact time they learnt the news. These results fit with the idea that FBMs are reconstructed with almost perceptual clarity (Brown & Kulik, 1977). The exact time is not a perceptual element. One might recall with perceptual clarity that it was sometime in the morning as they remember the morning sun shining through their windows, yet they might not be able to specify whether it was 8.10 or 9.30. It is also harder to be consistent in the way one reports one's activity, as already seen in other studies (Curci & Luminet, 2009; Curci et al., 2001). There are many ways to express an activity. Someone might say "I was studying at might desk" at T1 and then explain at T2 that they were "working on their computer on an assignment". These responses are not consistent, yet they could both be accurate. We believe to have reduced this issue as much as possible by

using an interview method to collect data at T2 and prompt for further explanation when necessary.

Besides methodological effects, it is also important to consider the role of time in the evaluation of consistency. As discussed in the introduction, sleep tends to consolidate memories by reconstructing them in a way that fits our pre-existing schemas (Landmann et al., 2014; Stickgold & Walker, 2013), thus increasing the likelihood of consistency after this consolidation phase occurs even if accuracy is reduced. Furthermore, once errors (additions or modifications) are introduced, these inconsistencies tend to be included in further reconstruction of the memory and repeated over time (Coluccia et al., 2006). In a long-term study of 9/11, Hirst et al. (2015) found that FBMs only became fully consistent after a year. In the long run, what participants recalled were not their original memories but instead, they recalled the memories that had been reconstructed later on, whether accurate or not. These findings indicate that consistency is not always comparable to accuracy, especially when baseline data is collected later on.

Nevertheless, we believe that our study provides good indication of high consistency of FBM after more than a year. Indeed, for perceptual elements, participants were generally able to provide an accurate and consistent answer on the canonical questions. Notably, consistency for FBMs was better than for EMs, a result not always found in other research (e.g., Talarico & Rubin, 2003). We also compared the details they freely provided about their environment or the situation they were in when they learnt about the attacks at the two time points. We found very few contradictions overall, even if they also provided many new details at T2 that were not discussed at T1. As argued by Muzzulini et al. (2020), studies should examine free recall of the reception context more often.

In order to better understand the conceptualisation of FBM, we also examined the relationship between the different FBM attributes assessed. We found some moderate

correlations between all indicators except confidence. This finding was unexpected. Talarico and Rubin (2003, 2007) have argued that confidence should be one of the main attributes of FBM and as a result, it has been one of the most measured indicators within the last ten years. We suspect our results are due to a measurement issue. Confidence was assessed as a single question at the end of each section instead of measuring participants' confidence for each detail provided.

The correlation between consistency and vividness was also relatively moderate, especially within a small sample. This means that perceptual clarity as well as belief in accuracy do not predict consistency over time. Thus, in the absence of a clear theoretical conceptualisation of the way FBM should be operationalised, we should strive to empirically measure and compare them as often as possible. It is only through regular reporting of correlations that a future meta-analysis could finally provide some much-needed empirical clarification on these central components of FBM.

The Role of Social Identification

Our second main objective was related to social identity. We assessed three levels (local, national and supra-national) of social identification on continuous scales measuring pride and attachment. As both pride and attachment were strongly correlated and provided similar results, we interpret them here together. We found that social identification with Brussels and Europe were related to most measures of FBM. Surprisingly, however, participants' attachment to Belgium and their pride of being Belgian did not correlate with any FBM attributes. Up until now, studies who considered social identification did so mostly by including participants' citizenship (e.g., Curci & Luminet, 2006; Curci et al., 2001; Talarico et al., 2019). What our results show is that citizenship might not be the only meaningful measure of social identification in all contexts. In our study, many factors may have affected the results. First, Merck et al. (2020) suggested that the reason they found a low

correlation between level of baseball fandom and confidence in FBM for a win was that it was already consequential enough to matter, regardless of how much a fan participant were. In our case, it is possible that regardless of the level of social identification with Belgium, all participants were equally surprised and impacted by the event. Yet, we can also find other interpretations within Belgium's culture and history. As discussed in the introduction, the long-standing conflict between the two major linguistic groups of Belgium has had a negative impact on national pride (Luminet et al., 2012). Some Belgians might thus identify more closely with their city, their region or even with Europe than with their country. The second factor is the perceived threat of the bombings. For some, the so-called "Brussels bombings" were an attack on the city itself, a city often associated with influential terrorist cells such has the one responsible for the November 2015 Paris attacks, but also a city that is generally seen as the *de facto* capital of the European Union. As such, many people across the globe have also interpreted the bombings as one more attack after many others (such as the ones in France, Spain or the UK) that embodies the global terrorist threat of the Islamic State on Europe and European values (Rathore, 2016). It is thus unsurprising that social identification with Brussels or Europe is associated with indicators of FBM.

What was more unexpected was the absence of a clear relationship between social identity and EM. Prior studies found that citizenship has an impact on the accuracy or the consistency of EM: Americans were more accurate about 9/11 than European (Curci & Luminet, 2006) and French exhibited a better memory of their President Mitterrand's death than Belgians (Curci et al., 2001). Nevertheless, these comparisons were between participants of different citizenships, on events that most likely impacted some countries more than others. What our results indicate is that within the population of Belgium, participants' level of identification with Brussels, their region or they country did not impact their knowledge about the event.

To conclude, social identity undoubtedly plays an important role in the formation of FBMs (Berntsen, 2018) by providing cultural cohesion in a shared collective past and feelings of belonging to the groups (Cameron, 2004; Merck et al., 2020; Sahdra & Ross, 2007; Talarico et al., 2019). However, we are yet to fully understand the processes by which social identity influences FBM. What we have demonstrated here is the need to consider the context of the event, examine a wider range of potential groups with which one could identify and evaluate the level of social identification with these groups. We also strongly recommend the use of scales to evaluate social identification instead of placing individuals into a priori categories such as citizenship.

Limits and Recommendations for Future Studies

There were some limits to our study. As discussed previously, we were unable to collect a large sample of participants on the day of the event. As this was one of the central aims of our research, we collected an online sample of participants in order to compare their responses at T2 with the responses from our original sample. Even if they did not differ on any important measures, the small size of our original sample still affected our statistical analyses and did not allow us to examine more in-depth all collected variables. Thus, the goal of this study is not to provide strong empirical data that could test predictive models but to broaden the debate on important points regarding the study of FBMs.

Another limit resides in the questionnaire design. Upon examination of the data, we realised that some measures were not ideal and some would have to be completely removed. For instance, questions relating to EM were too difficult for participants and led to a floor effect. Appropriate evaluation of EM is at best difficult and at worst unattainable as it cannot rely on standardised questions and thus will always be contingent on the choice of questions. Canonical categories of FBMs were also selected based on prior studies. However other categories, maybe better suited, could also have been added for a better understanding of

FBM, as discussed by Kızılöz and Tekcan (2013). The phrasing of the question related to vividness could also be improved, as it used terms from the remember/know/guess paradigm but the description of the different points of the scale related more to the vividness and the feeling of re-experiencing the event. We also should have measured confidence after each item and not only at the end of each section.

Finally, another limit resides within our measures of social identification and the fact that they were only obtained during the second collection time. Politics in Belgium are constantly evolving, but this was particularly the case in 2016 after the terrorist attacks discussed here. As a response, many new security measures were instated within the country and they were not always well received by the population. Other political problems destabilised the government and several politicians resigned from their positions that year. This means that participants' social identification with Belgium, Europe and Brussels measured in June 2017 might not have been representative of how they felt back in March 2016.

More importantly, we have provided some recommendations for future studies throughout this article. One of the main points we have highlighted is the need to use a broad range of variables – whether to measure EM and FBM attributes – and report correlations between them. FBM literature has recently struggled to agree on many essential points, in particular the way FBM should be operationalised (Luminet, 2018; Talarico & Rubin, 2018). Studies that attempt to measure more than two types of attributes are rare and ones that provide correlations between them are even rarer. However, understanding the correlational patterns across multiple studies could provide important empirical data to clarify the ways in which FBM differs from autobiographical memory. Similarly, studies should also endeavour to measure a broad range of predictors whenever possible and provide the correlational matrix.

Another recommendation is the importance of careful design combined with early data collection. The best way forward would be to work towards the creation of a standardised questionnaire that could be easily adapted for many different types of events. The benefit of such a questionnaire would be twofold: 1) it would speed up the setting up process and would increase the possibility to be operational to test on the day of the event; 2) it would allow direct comparison of data across studies and – ultimately – would provide the opportunity to run meta-analyses. We also recommend the use of face-to-face interviews for the second time of data collection, in order to improve measures of consistency.

The status of FBM has intrigued researcher for over forty years, yet the difficulty to find agreements on the way we conceptualise and operationalise them is forever ongoing. In this context, comparing results of studies that use different attributes of FBM and various methods to measure them is at best impractical. We believe the field would benefit greatly from working towards standardised questionnaires that use a multidimensional evaluation approach by including a broad range of measures and automatic reporting of correlational patterns. Furthermore, researchers should strive for early data collection to allow for more comprehensive evaluation of accuracy, but also to obtain measures of predictors as experienced on the day and not only as remembered days, weeks or even months later. Finally, future research should investigate the role of data collection method at retest in the evaluation of consistency over time.

Acknowledgments

This section has been removed to ensure anonymity

Disclosure of Interest Statement

"The authors declare that there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. Furthermore, the authors declare that this manuscript is an original work and have not been published before, not is it being considered for publication elsewhere either in printed or electronic form

"The authors confirm that the manuscript adheres to ethical guidelines specified in the APA Code of Conduct as well as authors' national ethics guidelines."

"Upon acceptance of the manuscript, data will be shared on OSF on request."

References

Berntsen, D. (2018). Flashbulb memories and social identity. In O. Luminet & A. Curci (Eds.), *Flashbulb Memories: New challenges and future perspectives* (pp. 182-200). Psychology Press.

Berntsen, D., & Thomsen, D. K. (2005, May). Personal memories for remote historical events: accuracy and clarity of flashbulb memories related to World War II. *Journal of Experimental Psychology: General, 134*(2), 242-257. https://doi.org/10.1037/0096-3445.134.2.242

Bohannon, J. N. I., & Symons, V. L. (1992). Flashbulb memories: Confidence, consistency, and quantity. In E. Winograd & U. Neisser (Eds.), *Emory symposia in cognition, 4. Affect and accuracy in recall: Studies of "flashbulb" memories* (pp. 65-91). Cambridge University Press.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5*(1), 73-99.

Cameron, J. E. (2004). A three-factor model of social identity. *Self and identity, 3*(3), 239-262. https://doi.org/https://doi.org/10.1080/13576500444000047

Coluccia, E., Bianco, C., & Brandimonte, M. A. (2006, May). Dissociating veridicality, consistency, and confidence in autobiographical and event memories for the Columbia shuttle disaster. *Memory, 14*(4), 452-470. https://doi.org/10.1080/09658210500477709

Curci, A., & Luminet, O. (2006, Apr). Follow-up of a cross-national comparison on flashbulb and event memory for the September 11th attacks. *Memory, 14*(3), 329-344. https://doi.org/10.1080/09658210500340816

Curci, A., & Luminet, O. (2009). Flashbulb memories for expected events: A test of the emotional-integrative model. *Applied Cognitive Psychology, 23*(1), 98-114. https://doi.org/10.1002/acp.1444

Curci, A., Luminet, O., Finkenauer, C., & Gisle, L. (2001, Mar). Flashbulb memories in social groups: A comparative test-retest study of the memory of French President Mitterrand's death in a French and a Belgian group. *Memory, 9*(2), 81-101. https://doi.org/10.1080/09658210042000120

Diekelmann, S., & Born, J. (2010, Feb). The memory function of sleep. *Nature Reviews Neuroscience, 11*(2), 114-126. https://doi.org/10.1038/nrn2762

Duchesne, S., & Frognier, A.-P. (2007). Why is it so difficult to know if national pride leads the way to European identity or prevents it? *Cahiers Europeens, 3*, 1-18. https://doi.org/http://dx.doi.org/10.2139/ssrn.1522904

Dumas, R., & Luminet, O. (2016). Emotional context, rehearsal and memories: The mutual contributions and possible integration of flashbulb memory and eyewitness identification research. In C. B. Stone & L. Bietti (Eds.), *Contextualizing human memory: An interdisciplinary approach to understanding how individuals and groups remember the past* (pp. 37-65). Routledge.

Finkenauer, C., Luminet, O., Gisle, L., El-Ahmadi, A., Van der Linden, M., & Philippot, P. (1998). Flashbulb memories and the underlying mechanisms of their formation: Towards an emotional-integrative model. *Memory & Cognition, 26*(3), 516-531. https://doi.org/https://doi.org/10.3758/BF03201160

Greenberg, D. L. (2004). President Bush's false [flashbulb] memory of 9/11/01. *Applied Cognitive Psychology, 18*, 363-370. https://doi.org/10.1002/acp.1016

Hirst, W., & Phelps, E. A. (2016). Flashbulb Memories. *Current Directions in Psychological Science, 25*(1), 36-41. https://doi.org/10.1177/0963721415622487

Hirst, W., Phelps, E. A., Meksin, R., Vaidya, C. J., Johnson, M. K., Mitchell, K. J., Buckner, R. L., Budson, A. E., Gabrieli, J. D. E., & Lustig, C. (2015). A ten-year follow-up of a study of memory for the attack of September 11, 2001: Flashbulb memories and memories for flashbulb events. *Journal of Experimental Psychology: General, 144*(3), 604-623.

Hu, P., Stylos-Allan, M., & Walker, M. P. (2006). Sleep facilitates consolidation of emotional declarative memory. *Psychological Science, 17*, 891-898. https://doi.org/https://doi.org/10.1111/j.1467-9280.2006.01799.x

Jeffreys, H. (1998). *The theory of probability*. Oxford University Press.

Kızılöz, B. K., & Tekcan, A. I. (2013). Canonical categories in flashbulb memories. *Applied Cognitive Psychology, 27*(3), 352-359. https://doi.org/10.1002/acp.2913

Klein, O., Licata, L., Van der Linden, N., Mercy, A., & Luminet, O. (2012). A waffle-shaped model for how realistic dimensions of the Belgian conflict structure collective memories and stereotypes. *Memory Studies, 5*(1), 16-31. https://doi.org/https://doi.org/10.1177/1750698011424028

Koppel, J., Brown, A. D., Stone, C. B., Coman, A., & Hirst, W. (2013). Remembering President Barack Obama's inauguration and the landing of US Airways Flight 1549: A comparison of the predictors of autobiographical and event memory. *Memory, 21*(7), 798-806. https://doi.org/10.1080/09658211.2012.756040

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General, 142*, 573-603. https://doi.org/https://doi.org/10.1037/a0029146

Landmann, N., Kuhn, M., Piosczyk, H., Feige, B., Baglioni, C., Spiegelhalder, K., Frase, L., Riemann, D., Sterr, A., & Nissen, C. (2014). The reorganisation of memory during sleep. *Sleep medicine reviews, 18*(6), 531-541. https://doi.org/https://doi.org/10.1016/j.smrv.2014.03.005

Lipinska, G., Stuart, B., Thomas, K. G. F., Baldwin, D. S., & Bolinger, E. (2019). Preferential consolidation of emotional memory during sleep: A meta-analysis. *Frontiers in Psychology, 10*, 1014-1031. https://doi.org/10.3389/fpsyg.2019.01014

Luminet, O. (2018). Using Structural Equation Modelling approaches to better understand the formation of Flashbulb memories. In O. Luminet & A. Curci (Eds.), *Flashbulb Memories: New challenges and future perspectives* (pp. 48-72). Psychology Press.

Luminet, O., & Curci, A. (2009, Oct). The 9/11 attacks inside and outside the US: Testing four models of flashbulb memory formation across groups and the specific effects of social identity. *Memory, 17*(7), 742-759. https://doi.org/10.1080/09658210903081827

Luminet, O., Curci, A., Marsh, E. J., Wessel, I., Constantin, T., Gencoz, F., & Yogo, M. (2004). The cognitive emotional and social impacts of the September 11 attacks: Group differences in memory for the reception content and the determinants of flashbulb memory. *Journal of General Psychology, 131*(3), 197-224. https://doi.org/https://doi.org/10.3200/GENP.131.3.197-224

Luminet, O., Licata, L., Klein, O., Rosoux, V., Heenen-Wolff, S., Van Ypersele, L., & Stone, C. B. (2012). The interplay between collective memory and the erosion of nation states–the paradigmatic case of Belgium: Introduction to the special issue. *Memory Studies, 5*(1), 3-15. https://doi.org/10.1177/1750698011424027

Merck, C., Yamashiro, J. K., & Hirst, W. (2020, Jun 26). Remembering the big game: Social identity and memory for media events. *Memory*, 1-20. https://doi.org/10.1080/09658211.2020.1784232

Muzzulini, B., Tinti, C., Conway, M. A., Testa, S., & Schmidt, S. (2020, Jun 17). Flashbulb memory: Referring back to Brown and Kulik's definition. *Memory*, 1-17. https://doi.org/10.1080/09658211.2020.1778035

Payne, J. D., Stickgold, R., Swanberg, K., & Kensinger, A. (2008). Sleep preferentially enhances memory for emotional components of scenes. *Psychological Science, 19*, 781-788. https://doi.org/https://doi.org/10.1111/j.1467-9280.2008.02157.x

Pezdek, K. (2003). Event memory and autobiographical memory for the events of September 11, 2001. *Applied Cognitive Psychology, 17*(9), 1033-1045. https://doi.org/10.1002/acp.984

Rathore, S. A. (2016). Brussels attack: Lessons learned. *Counter Terrorist Trends and Analyses, 8*(3), 4-8.

Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. (2003). Open-ended vs. close-ended questions in web questionnaires. In A. Ferligoj & A. Mrvar (Eds.), *Developments in applied statistics* (pp. 159-177). FDV.

Roccas, S., Klar, Y., & Liviatan, I. (2006). The paradox of group-based guilt: Modes of national identification, conflict vehelence, and reactions to the In-Group's Moral Violation. *Journal of Personality and Social Psychology, 91*(4), 698-711. https://doi.org/10.1037/0022-3514.91.4.698

Sahdra, B., & Ross, M. (2007, Mar). Group identification and historical memory. *Personality and Social Psychology Bulletin, 33*(3), 384-395. https://doi.org/10.1177/0146167206296103

Schonauer, M., Gratsch, M., & Gais, S. (2015, Feb). Evidence for two distinct sleep-related long-term memory consolidation processes. *Cortex, 63*, 68-78. https://doi.org/10.1016/j.cortex.2014.08.005

Stickgold, R., & Walker, M. P. (2013). Sleep-dependent memory triage: Evolving generalization through selective processing. *Nature Neuroscience, 16*(2), 139-145. https://doi.org/10.1038/nn.3303

Talarico, J. M., Bohn, A., & Wessel, I. (2019). The role of event relevance and congruence to social groups in flashbulb memory formation. *Memory, 27*(7), 985-997. https://doi.org/10.1080/09658211.2019.1616097

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

Talarico, J. M., & Rubin, D. C. (2007). Flashbulb memories are special after all; in phenomenology, not accuracy. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 21*(5), 557-578. https://doi.org/https://doi.org/10.1002/acp.1293

Talarico, J. M., & Rubin, D. C. (2018). Ordinary memory processes shape flashbulb memories of extraordinary events: A review of 40 years of research. In O. Luminet & A. Curci (Eds.), *Flashbulb memories: New challenges and future perspectives* (pp. 89-111). Psychology Press.

Temler, M., Barnier, A. J., Sutton, J., & McIlwain, D. J. F. (2020). Contamination or Natural Variation? A Comparison of Contradictions from Suggested Contagion and Intrinsic Variation in Repeated Autobiographical Accounts. *Journal of Applied Research in Memory and Cognition, 9*(1), 108-117. https://doi.org/10.1016/j.jarmac.2019.12.001

Tinti, C., Schmidt, S., Testa, S., & Levine, L. J. (2014, May). Distinct processes shape flashbulb and event memories. *Memory & Cognition, 42*(4), 539-551. https://doi.org/10.3758/s13421-013-0383-9

Tuckey, M. R., & Brewer, N. (2003). The influence of schemas, stimulus ambiguity, and interview schedule on eyewitness memory over time. *Journal of Experimental Psychology: Applied, 9*(2), 101-118. https://doi.org/https://doi.org/10.1037/1076-898X.9.2.101

Winningham, R. G., Hyman, I. E., Jr., & Dinnel, D. L. (2000, Jul). Flashbulb memories? The effects of when the initial memory report was obtained. *Memory, 8*(4), 209-216. https://doi.org/10.1080/096582100406775

Table 1
Overview of the Selected Study for the Review and their Types of FBM Operationalisations

Year	Authors	Test-retest	Confidence	Consistency	Vividness	Specificity	TOTAL
2010	Kvavilashvili, Mirani, Schlagman, Erskine, & Kornbrot	Yes	1	1	1	1	4
2010	Coluccia, Bianco, & Brandimonte	Yes	1	1	0	1	3
2010	Denver, Lane, & Cherry	No	1	0	0	1	2
2011	Kulkovsky, Wang, Conway, Hou, Aydin, Mueller-Johnson, & Williams	No	0	0	0	0	0
2011	Schaefer, Halldorson, & Dizon-Reynante	Yes	0	1	0	1	2
2012	Talarico, & Moore	Yes	1	1	1	0	3
2012	Rimmele, Davachi, & Phelps	Yes	1	1	0	0	2
2012	Laciano, & Curci	Yes	0	1	0	0	1
2013	Kizilöz, & Tekcan	No	0	0	0	1	1
2013	Lanciano, Curci, Mastandrea, & Sartoi	Yes	1	1	0	1	3
2013	Metternich, Wagner, Schulze-Bonhage, Buschmann, & McCarthy	Yes	1	1	0	0	2
2013	Lanciano, Curci, & Soleti	Yes	0	1	0	0	1
2014	Tinti, Schmidt, Testa, & Levine	No	1	0	1	1	3
2014	Waring, Seiger, Solomon, Budson, & Kensinger	Yes	1	1	0	0	2
2014	Day, & Ross	Yes	1	1	0	0	2
2014	Kraha, Talarico, & Boals	Yes	1	1	1	0	3
2014	Kraha, & Boals	Yes	1	1	1	0	3
2015	Hirst, Phelps, Meksin, Vaidya, Johnson, Mitchell, Buckner, Budson, Gabrieli, Lustig, Mather, Ochsner, Schacter, Simons, Lyle, Cuc, & Olsson	Yes	1	1	0	0	2
2015	Curci, Lanciano, Maddalena, Mastandrea, & Sartori	Yes	1	1	0	1	3
2015	Demiray, & Freund	No	0	0	1	1	2
2015	Stone, Luminet, & Takahashi	No	1	0	0	0	1
2017	Valet, Manzanero, Aroztegui, & Garcia Rudo	No	0	0	1	0	1
2017	Gandolphe, & El Haj	No	0	0	1	1	2
2017	Tamayo-Agudelo, Castano, & Cardenas	No	0	0	0	1	1
2018	Halpern, Talarico, Gouda, & Williamsn	No	1	0	1	0	2
2018	Spanhel, Wagner, Geiger, Schulze, Bonhage, & Metternich	Yes	0	1	0	0	1
2018	Ford, DiBiase, Ryu, & Kensinger	Yes	0	0	0	0	0
2018	Lanciano, Curci, Matera, & Sartoni	No	1	0	0	1	2
2019	Lanciano, Curci, & Basile	No	1	0	0	1	2
2019	Talarico, Kraha, Self, & Boals	Yes	1	1	1	0	3
2019	Krackow, Deming, Longo, & DiSciullo	Yes	1	1	0	0	2
2019	Talarico, Bohn, & Wessel	No	1	0	1	1	3

2019	Tamayo-Agudelo, Cardenas, & Castano	No	0	0	0	1	1
2020	May, Dein, & Ford	No	1	0	1	1	3
2020	Muzzulini, Tinti, Conway, Testa, & Schmidt	Yes	0	1	1	1	3
2020	Merck, Yamashiro, & Hirst	Yes	1	1	1	0	3
2020	Cassol, Bonin, Bastin, Puttaert, Charland-Verville, Laureys, & Martial	No	0	0	1	1	2
2020	Schmidt, & Qiao	No	0	0	1	1	2

Table 2
Descriptive Data and Bayes Factors (BH01) for all Measured Variables and Composite Scores

Measures	Range	*M (SD) / ratio*		BH01
		Immediate Group	Control Group	
Gender	M – F	4 M – 14 F	18 M – 42 F	3.117
Age	18 – 26	21.33 (1.68)	22.08 (1.82)	1.357
Event Memory				
EM accuracy % total	0 – 100 %	49.85 (17.16)	55.28 (22.22)	2.529
Time airport accuracy	0 – 100 %	64.82 (33.28)	56.83 (34.07)	2.683
Location airport accuracy	0 – 100 %	75.00 (35.35)	67.21 (38.60)	2.896
Explosions airport accuracy	0 – 100 %	72.22 (42.78)	75.40 (40.43)	3.566
Deaths airport accuracy	0 – 100 %	18.52 (32.78)	34.42 (40.80)	1.433
Time metro accuracy	0 – 100 %	**38.89 (28.58)**	**62.84 (36.05)**	**0.240**
Death metro accuracy	0 – 100 %	29.63 (35.95)	34.97 (39.17)	3.305
EM Confidence	1 – 7	3.44 (1.29)	3.31 (1.44)	2.529
Flashbulb memory				
FBM Confidence	1 – 7	5.44 (1.50)	5.45 (1.19)	3.690
FBM Vividness total	1 – 4	3.58 (0.57)	3.79 (0.33)	0.737
Time vividness	1 – 4	3.00 (0.97)	3.24 (0.94)	2.508
How vividness	1 – 4	3.72 (0.58)	3.91 (0.37)	1.121
Where vividness	1 – 4	3.89 (0.47)	3.90 (0.39)	3.673
With whom vividness	1 – 4	3.89 (0.32)	3.91 (0.37)	3.561
Ongoing activity vividness	1 – 4	3.61 (0.98)	3.86 (0.49)	1.441
Activity after vividness	1 – 4	**3.39 (0.92)**	**3.91 (0.33)**	**0.012**
FBM Number of details provided	0 – 5	3.61 (1.65)	3.72 (1.76)	3.609
Predictors				
Novelty total	1 – 7	5.36 (1.05)	4.73 (1.28)	0.860
Common	1 – 7	5.72 (1.71)	5.70 (2.02)	3.685
Unusual	1 – 7	4.50 (1.98)	3.37 (2.16)	0.764
Special	1 – 7	5.12 (1.58)	4.76 (1.95)	2.986
Singular	1 – 7	4.72 (1.84)	3.59 (2.29)	0.835
Ordinary	1 – 7	6.67 (0.69)	6.31 (1.18)	2.037
Unexpectedness	1 – 7	5.17 (1.79)	5.22 (1.59)	3.676
Unbelievability	1 – 7	4.56 (2.23)	4.13 (1.88)	2.823
Surprise emotion	1 – 7	5.22 (1.86)	4.18 (2.09)	0.847
Importance – consequentiality total	1 – 7	4.74 (0.95)	4.68 (1.11)	3.623
Consequence	1 – 7	4.72 (1.56)	3.93 (1.70)	1.035
Importance for self	1 – 7	3.83 (1.50)	3.90 (1.58)	3.653
Importance for friends/family	1 – 7	4.06 (1.35)	4.16 (1.47)	3.554
Importance for Belgium	1 – 7	5.94 (1.35)	5.95 (1.10)	3.693
International importance	1 – 7	5.17 (1.20)	5.26 (1.44)	3.595
Emotional intensity total	1 – 7	5.43 (1.25)	5.21 (1.46)	3.222
Upset	1 – 7	5.17 (1.20)	4.81 (1.50)	3.627
Indifferent	1 – 7	4.82 (1.42)	6.16 (1.69)	3.085
Affected	1 – 7	6.44 (1.04)	5.18 (1.92)	2.910
Shaken	1 – 7	5.56 (1.34)	4.72 (2.01)	3.279
Rehearsal through media total	1 – 5	3.41 (0.97)	3.03 (0.97)	1.617
TV	1 – 5	3.06 (1.39)	2.66 (1.20)	2.104
Internet	1 – 5	3.33 (1.41)	3.20 (1.24)	3.472
Social networks	1 – 5	3.83 (1.29)	3.32 (1.20)	1.402
Rehearsal through discussion total	1 – 5	3.86 (0.82)	3.90 (1.19)	2.716
How often	1 – 5	4.05 (0.80)	3.85 (1.15)	3.001
With how many people	1 – 5	3.64 (0.99)	3.90 (1.43)	2.941
Social identity				
Attachment Brussels	1 – 7	4.06 (1.98)	4.62 (1.72)	2.064
Pride Brussels	1 – 7	2.67 (2.33)	3.67 (2.31)	1.253
Attachment Belgium	1 – 7	5.89 (1.18)	5.77 (1.52)	3.554
Pride Belgium	1 – 7	4.89 (2.11)	5.39 (1.74)	2.386
Attachment Europe	1 – 7	4.83 (1.65)	4.86 (1.85)	3.685
Pride Europe	1 – 7	4.06 (2.13)	4.75 (1.94)	1.815

Table 3
Contingency Table, Means and Standard Deviations of FBM Canonical Consistency

Variable	Canonical consistency score			*M*(*SD*)
	0 (different)	1 (similar)	2 (identical)	
Other people present	2 (11%)	1 (6%)	15 (83%)	1.72 (0.67)
Exact location	3 (17%)	1 (6%)	14 (78%)	1.61 (0.78)
Informant	4 (22%)	4 (22%)	10 (56%)	1.33 (0.84)
What they did after	4 (22%)	5 (28%)	9 (50%)	1.27 (0.83)
Ongoing activity	6 (33%)	5 (28%)	7 (39%)	1.06 (0.87)
Time of day	11 (61%)	6 (33%)	1 (6%)	0.44 (0.62)

Table 4
Inter-Correlation Matrix for FBM Attributes

Attributes	Canonical consistency	Details consistency	Vividness T2	Confidence T2
Canonical consistency	1			
Details consistency	**.230**	1		
Vividness T2	**.552**	**.639**	1	
Confidence	-.062	.051	-.115	1

Note. Correlation >.300 are in bold.

Table 5
Means, Standard Deviations and Correlations of Social Identity Items

Items	**Attachment** ***M* (*SD*)**	**Pride** ***M* (*SD*)**	**t test**	d_z	***r***
Brussels	4.06 (1.98)	-	-	-	-
Belgium	5.89 (1.18)	4.89 (2.11)	$t(17)=2.38$, $p=.029$	0.56	.537
Europe	4.83 (1.65)	4.06 (2.13)	$t(17)=2.12$, $p=.049$	0.50	.688

Note. Range of response from 1 to 7

Table 6
Correlation Matrix Between SI and Attributes of FBMs and EMs.

	Attachment Brussels	Attachment Belgium	Attachment Europe	Pride Belgian	Pride European
FBM consistency	**.390**	.015	**.415**	.116	**.441**
FBM details consistency	**.545**	.180	**.438**	.018	.277
FBM vividness T2	**.519**	.146	**.717**	.270	**.549**
FBM confidence T2	**.405**	-.037	-.134	.091	-.063
EM consistency	.133	**.364**	.034	-.043	.029
EM confidence	.013	.265	-.018	-.110	.033
EM accuracy at T2	.197	.144	.160	.266	**.314**

Note. Correlation >.300 are in bold.