Inconsistent reporting of life experiences: What people think and how they explain it?

Vanessa Azevedo* / Carla Martins** / Margarida Carvalho** / Ângela Maia**

* Universidade Fernando Pessoa, Porto, Portugal; ** Escola de Psicologia, Universidade do Minho, Braga, Portugal

Assessment of life experiences relies mainly on cross-sectional retrospective design, despite the concerns regarding inconsistent reports. Studies suggest that some individuals change their answers when asked repeatedly, but common opinions underlying this behaviour remain unknown. Our study explored personal perceptions regarding inconsistent reporting and identified associated reasons including individual, experiences-related, and design-related characteristics. Seventy-two individuals, enrolled in a longitudinal study about life experiences, answered a measure about general perceptions and involved reasons. Participants seemed to be aware that inconsistent reporting is a common behaviour, which highly impact on research. A cluster analysis revealed two clusters (i.e., variables involved *versus* not involved). Most disagreed that sociodemographic variables influence inconsistency, whereas memory, mood, valence, impact, mode of data collection, and interviewer features were pointed as key-variables. Our results suggest that inconsistent reporting is not straightforward and it is probably rooted in a varied and complex set of variables.

Key words: Adults, Life events, Reliability, Reporting practices, Retrospective assessment.

Across all fields of psychology and especially on psychology of crime, victimization, and justice, research and assessment about life experiences rely mostly on cross-sectional retrospective designs, through the application of lists, schedules or interviews (e.g., Hardt & Rutter, 2004; Kendall-Tackett & Becker-Blease, 2004; Paykel, 2001). However, concerns about the inconsistency of reports frequently overshadow findings and conclusions. At this point, to avoid misconceptions, it is crucial to clarify inconsistency. Briefly, Dube and colleagues (2004) distinguished validity that assesses the veracity of the report, from reliability, which means that a report is stable across time. Consequently, a report can be stable but not valid, while a valid report is necessarily stable. Overall, research about the validity is constrained by the (im)possibility of reports' verification; oppositely, reliability allows for different and easier designs, i.e., test-retest

using the same or different methods of data collection, which extends research opportunities and potential practical implications. Despite of the consensus around the construct, different labels (e.g., reliability, stability, inconsistency) are applied interchangeably. In line with more recent studies (e.g., Ayalon, 2017; Colman et al., 2016; Spinhoven et al., 2012) and to dissociate the phenomenon from the psychometric terminology, we referred to the phenomenon as inconsistency. More specifically, we defined inconsistency as the change on responses regarding a specific issue across time and inconsistent reporters (as labelled in Results section) are those people that change their answers about lived (or not) life experiences (e.g., a person that at T1 claimed that had a serious illness, while at T2 denied/omitted it).

Although figures of inconsistency are quite varied (e.g., Dube et al., 2004; Fergusson et al., 2000), there are always some people that change their answers when asked twice. This behaviour is far from being innocuous, especially on legal context, as it can have a serious impact on the loss of credibility of reports and on research findings due to misclassifications (Langeland et al., 2015). Consequently, it has been addressed in an effort to quantify, describe, and explain since the 1960's (e.g., Casey et al., 1967; Cleary, 1980; Paykel, 1983).

Despite decades of research, empirical data about reasons involved in inconsistent reporting is still lacking and inconclusive. Usually, researchers established one or more variables of interest (e.g., mood), collect data about them and, then, through statistical procedures, tested their hypotheses about factors that contribute to inconsistency. As a result, there are some cues about inconsistent reporting, but findings are neither comprehensive nor clarifying. Additionally, this approach lacks an ecological perspective and dismissed common people perceptions and meanings. Not surprisingly many variables (e.g., shame, protection of third parties), in the absence of empirical data, remain as a plausible conjecture. A second approach to address this issue consists in asking subjects about their reasons for inconsistent reporting, through quantitative or qualitative designs; but very few studies applied this strategy.

The earliest work by Sobell and colleagues (1990) was based on a sample of 69 college students, which were interviewed twice about adult life events, and then were asked to comment and explain their inconsistencies. Authors found that the re-evaluation of the events' importance was the main reason, followed by the incorrect dating of the events, both explaining 64.3% of the inconsistencies. Similarly, Langeland et al. (2015) studied the reasons involved in inconsistent reporting of child sexual abuse in a sample of 633 adults. After answering twice to an online survey, participants were confronted with their changes, being asked to choose their explanatory reasons from a predefined list. Authors concluded that misunderstanding of the questions, memory issues, feeling overwhelmed and avoidance, were the main reasons provided. Through a qualitative approach, Carvalho (2015) explored the perceptions and reasons based on the semi-structure interviews of 12 participants enrolled in a longitudinal study about the topic. These participants explained inconsistency through memory issues, mood, valence, shame, and interviewer's characteristics.

In an effort to improve the current state of the art, we aimed to study what people think about inconsistent reporting, through a quantitative approach, introducing some novelties in the design and variables assessed. More specifically, participants were enrolled on a longitudinal study about positive and negative experiences, throughout the lifespan, being asked twice about them. Instead of confronting participants with their own inconsistencies, they were asked to focus on general behaviour; this strategy allied to the omission of their inconsistency status, seems to be a less threatening and a more generalizable approach (detailed information about how it was handle will be presented in Method – Procedures subsection). Regarding variables, besides individual factors, we also included design and experiences related features. A better knowledge of inconsistent reporting is critical not only to revise and re-evaluate past research, but also to design future enhanced studies and to improve assessments on legal contexts. Considering these benefits, this study had two main aims, namely: (1) To explore personal perceptions regarding inconsistent

reporting (e.g., frequency, pattern, impact for research, designation); (2) To identify reasons involved in inconsistent reporting, assessing individual's characteristics (e.g., gender, mental and physical health status, memory, secrecy), experiences' characteristics (e.g., valence, impact, severity, developmental stage), and design's characteristics (e.g., setting, time interval, mode of data collection, interviewer's characteristics).

Method

Participants

Participants were 72 adults from the community, mainly females (n=61; 84.7%), age ranging from 19 to 64 years of age (M=39.39, SD=13.25), enrolled in a longitudinal study about life experiences. Participants were recruited through schools from the North of Portugal, based upon three inclusion criteria: individuals older than 18 years, capable of speaking, reading and writing in Portuguese, and not planning migration in a short-medium term. Due to the Portuguese economic crisis (2010-2014), there was a strong increase in the unemployment rates, people changed home and phone numbers often and there was an intense migratory wave. These circumstantial conditions added extra complexities to longitudinal studies. In order to minimize these effects, data was collected mainly in schools that, in one hand, are less prone to mobility (at least, people used to stay for a school year) and, in the other, their populations are quite heterogeneous regarding sociodemographic variables. Therefore, participants can be school professionals (both teachers and non-teachers), parents (or their legal representatives), or adult students.

At T1, 171 participants were enrolled in the study and 123 were randomly selected to T2. Of these 32 were excluded because they did not return the questionnaire (n=9), refuse to participate (n=16), or were unable to contact (n=7). Lastly, 91 participants were enrolled at T2, and 2 were excluded due to incomplete assessment. Of these 89 participants, 80.90% were analysed here and the remaining were further analysed in other study (Carvalho, 2015).

The majority were married or cohabiting (59.7%) and 34.7% were single. Regarding employment status, most were employed (70.8%) and 23.6% were students; unemployed, retired, homemaker and other status represented 5.6% of the participants. The majority of the participants had a graduate degree (65.3%), 13.9% had an high school diploma, 12.5% had less nine or less years of education, and 8.3% had an undergraduate degree.

Measures

Data was collected through a closed-ended and self-report questionnaire entitled Perceptions and Reasons Involved in Inconsistent Reporting (PRIIR), which was specifically developed to gather personal perceptions and reasons involved in inconsistent reports of life experiences. PRIIR was especially design for this study and was based on a literature search, including not only variables that are empirically tested (e.g., mood), but also those that are traditionally speculative (e.g., shame). It was made an effort to guarantee that questions were nondirective and non-judgmental (e.g., mood affects the report *vs.* sad people provided inconsistent reports); additionally, PRIIR presents an overall perspective about the phenomenon asking about general people behaviour (*vs.* self).

PRIIR begins with an initial briefing about the topic, which clarifies the status quo of the topic and what are the purposes of the questionnaire. Then, participants were asked about the frequency of the behaviour based on a real scenario, i.e., Imagine that we asked ten people about their life experiences, in two distinct occasions; from zero to ten, how many people do you think that would change their reports, as well as the usual pattern, i.e., underreporting *vs.* overreporting *vs.* tie.

These general questions were followed by a list of potential reasons involved in inconsistent reporting, that was organized in three groups, i.e., related to experiences, related to design, and related to individuals. Reasons related to experiences included five variables, i.e., valence, importance, severity, developmental stage and number of life experiences. Setting, time interval between assessments, mode of data collection, the characteristics of the interviewer and the change of the interviewer between assessments were the variables assessed in reasons related to design. Finally, reasons related to individual included 16 variables, i.e., gender, age, education, marital status, employment status, economical status, mood, personality's characteristics, mental or physical health status, substance abuse, memory, secrecy, shame, protection of third parties, denial, and help seeking. Being participants from the community, these variables were presented in based on general and shared labels (instead of a formal scientific-way operationalization). For instance, personality's characteristics were assessed through the item "The characteristics of personality/temperament of the person" and if the participant agree should further clarify these characteristics, while mood was formulated as "The way the person feels (e.g., unhappy, happy, etc.). For each variable, participants rated their level of agreement on a 3 point-Likert scale (disagree *vs.* neither agree or disagree *vs.* agree); whether they agreed, they were asked to refine their answers, selecting among the options provided (e.g., gender matters → males/females tended to be inconsistent reporters). An open-ended question about other reasons finished this section. Based on these answers, and associated trends identified, we profiled inconsistent reporters.

A last set of questions assessed general perceptions about how inconsistency impacts on research (through a five-point Likert scale ranging from 0 – not at all to 4 – absolutely), the level of difficulty of the questions (through a five point-Likert scale ranging from 0 (not difficult) to 4 (very difficult), and previous thoughts about the topic (using a yes *vs.* no format). Finally, participants were invited to name the group of people that endorsed the behaviour.

The majority of items of PRIIR are similar to a checklist, being independent among themselves. Therefore, psychometric properties such as Cronbach alpha were not appropriate for this kind of measure.

Procedures and data analysis

The study was reviewed and approved by the Institutional Review Board and the Portuguese Data Protection Authority. Participants were randomly selected from an initial pool of respondents, enrolled in a longitudinal study about life experiences. More specifically, participants were asked about 75 life experiences organized in eight domains (i.e., school, work, health, leisure, life conditions, adverse experiences, achievements, and people and relationships). These items included both positive and negative life experiences throughout the lifespan. For each item, participants were asked about its occurrence; when they answered positively, additional questions were asked, namely developmental stage, valence, and impact. This longitudinal study consisted of two assessments with a mean interval of 154.58 days. At T1, participants were asked to answer the Lifetime Experiences Scale (Azevedo et al., 2020) and the Brief-Symptom Inventory (BSI; Derogatis, 1993; we used the Portuguese version by Canavarro (2007), a short version of the Symptom Checklist-90-Revised (SCL-90-5); while at T2 participants were asked to answer the same measures plus PRIIR. Although in the second wave some participants were face-to-face interviewed, all answered PRIIR through self-report. A brief preamble preceded PRIIR questions, describing the current state of the art and the remaining questions. More specifically, the presentation and instructions provided to participants were as follow:

What do we know? Different studies shows that people can tell about their life experiences in different ways, attending to the context or timing, namely sometimes they tell more life experiences, while other tell few. For instance, sometimes a person tells that had a serious illness and other times the same person answers that did not lived such experience.

What we do not know yet? The reasons why this happens, neither what general people think about that. In this study, we are interested in your most sincerely opinion about this issue. Therefore, there are no wrong or correct answers.

Data analysis was performed using software IBM Statistical Package for Social Sciences (IBM SPSS; version 22 for Windows). According to our aims, univariate descriptive statistics were obtained for perceptions and reasons involved in inconsistent reporting. Additionally, a cluster analysis was performed to clarify associated variables.

Traditionally, cluster analysis is a technique to group people, but it can also be used to group variables, including categorical ones (e.g., Pereira et al., 2015). Considering that about 11% of the participants had one or more missing values and that cluster analysis deletes observations with missing data, the modal category for each of the variables was imputed. Although this is the simplest method on the treatment of missing values, according to a simulation performed by Acuña and Rodriguez (2004) comparing case deletion, mean imputation, median imputation and KNN imputation, authors concluded that in datasets with a low number of missings there were no significant differences between the methods. Additionally, in a review about missing data analysis (Graham, 2009, p. 562) claimed that "although some researchers believe that missing categorical data requires special missing data procedures for categorical data, this is not true in general. The proportion of people giving the "1" response for a two-level categorical variable coded "1" and "0" is the same as the mean for that variable". Regarding clustering procedures, it was conducted a hierarchical cluster analysis for variables, using Ward's method and Square Euclidean distance for binary measures. The number of clusters was based on a preliminary cluster analysis without specifying that parameter; then, two clusters were identified through the dendrogram and the analysis were re-runned defining the number of clusters and requesting SPSS to present the number of cluster associated with each variable (Field, 2000). As suggested by other authors (e.g., Ketchen & Hult, 2000), more than one technique was used to test the model stability, namely a random division of the study sample into two halves and the application of a different similarity measure (specifically, Jaccard coefficient). When cluster analyses were repeated under these conditions, the models obtained were similar to the original one.

Results

When asked about previous experience on the topic, most of the participants (67.6%) admitted they were thinking about it for the first time. Regarding the level of difficulty associated with the task, only 7% considered that it was very difficult, 19.7% that it was quite difficult, 32.4% somewhat difficult, 25.4% slightly difficult and 15.5% not difficult.

Regarding the frequency of inconsistent reporting, 4.3% (n=3) of the participants predicted that none participants would change the answers across time; oppositely, 8.3% (n=6) admitted that all people would endorse the behaviour. Most participants (70%; n=49) considered that half or less of the people would change the reports between occasions; additionally, respondents considered that, in mean, 4.83 (SD=2.48) people would change their answers. Moreover, the majority of participants (62.3%) claimed that people would report more experiences in the second occasion and 26.1% considered that people would report more experiences in the first occasion. Almost 79% of the participants recognized that inconsistent reporting had very or absolute impact on research, whereas only 2.8% considered it as harmless.

Focusing on individual characteristics, as shown in Table 1, participants seemed to disagree that gender, education, marital status, employment status and income were involved in inconsistent reporting. Age seemed to be a less consensual variable, since we observed a tie between the percentages of agreements and disagreements. According to agreement responses, participants described inconsistent reporters mainly as females, young, divorced or separated, and having some

education (i.e., basic education – third cycle or graduate/undergraduate education). Additionally, participants indicated that people unemployed and with low income seemed to be more prone to inconsistent reporting.

Table 1

Percentages and frequencies of individual characteristics involved in inconsistent reporting

| | Agreement | | | |
Individual characteristics	Disagree n (%)	Neither agree or disagree n (%)	Agree n (%)	Profile of inconsistent reporters%
Gender[a]	**33 (46.5)**	18 (25.4)	20 (28.2)	Female: 55 Male: 45
Age[a]	24 (33.8)	23 (32.4)	24 (33.8)	18-30 years: 54.2 31-64 years: 29.2 65 or more years: 12.5
Education[a]	**33 (46.5)**	23 (32.4)	15 (21.1)	Basic education: 28.6 High education: 28.6 Elementary education: 21.4 None: 14.3S econdary education: 7.1
Marital status[a]	**34 (47.9)**	22 (31.0)	15 (21.1)	Divorced/separated: 33.3 Single: 20.0 Married/cohabitating: 20 Widowed: 20.0 Other answers: 6.7
Labour force status[a]	**32 (45.1)**	24 (33.8)	15 (21.1)	Unemployed: 40.0 Employed: 13.3 Retired: 13.3 Home maker: 13.3 Other answers: 20.1
Income[a]	**36 (50.7)**	23 (32.4)	12 (16.9)	Low income: 66.7 High income: 25.0 Medium income: 8.3
Mood[a]	6 (8.5)	12 (16.9)	**53 (74.6)**	Unhappy/low mood: 86.5 Happy/high mood: 7.7 Both: 5.8
Personality[b]	14 (20.3)	**31 (44.9)**	24 (34.8)	na
Physical and mental health atatus[c]	12 (17.1)	11 (15.7)	**47 (67.1)**	Mental health status: 70.2 Physical health status: 2.1 Both: 27.7
Substance use[a]	6 (8.5)	10 (14.1)	**55 (77.5)**	Drugs and alcohol use: 94.5 Drugs use: 3.6 Other answers: 1.8
Memory[a]	4 (5.6)	11 (15.5)	**56 (78.9)**	Amnesia: 57.1 Forgetfulness: 33.9 Both: 8.9
Secrecy[a]	9 (12.7)	17 (23.9)	**45 (63.4)**	na
Shame[a]	8 (11.3)	11 (15.5)	**52 (73.2)**	na
Third parties' protection[a]	6 (8.5)	12 (16.9)	**53 (74.6)**	na
Denial[a]	8 (11.3)	12 (16.9)	**51 (71.8)**	na
Attending professional help[a]	11 (15.5)	23 (32.4)	**37 (52.1)**	na

Note. Profile of inconsistent reporters is based on refinements made by participants that choose the category "agree"; y=years, na=not applicable; [a]N=71, [b]N=69, [c]N=70.

Moreover, most participants agreed that inconsistent reporting can be influenced by memory, substance use, mood, and physical and mental health condition. More specifically, people with amnesia, alcohol and drugs use, sadness/low mood, and a mental health condition tended to present an inconsistent report. Effort to protect third parties, shame, denial, and secrecy were also pointed as major reasons underlying inconsistent reporting. Help seeking was also reasoned by the majority of the participants as a variable involved in the behaviour. Participants did not present a clear position regarding the influence of personality characteristics, since the mode was neither agree or disagree (44.9%).

All variables devoted to experiences' characteristics, i.e., valence, impact, severity, developmental stage, and number of experiences, were presented as reasons for inconsistent reporting (Table 2). Specifically, experiences that are negative, that have high impact and severity and that occur during

childhood seemed to be more susceptible to inconsistent reporting. Besides, a high number of experiences was also considered a reason for the behaviour.

When participants were asked about design characteristics, as displayed in Table 3, most agreed that the mode of data collection, time interval between the first and second assessments (T1 and T2, respectively), interviewer's characteristics, and the change of the interviewer contribute to inconsistent reporting, which seemed to emerge when assessment occurs through interviews and with a time interval equal to 12 or more months.

Table 2

Percentages and frequencies of experiences characteristics involved in inconsistent reporting

	Level of agreement			
Experiences characteristics	Disagree n (%)	Neither agree or disagree n (%)	Agree n (%)	Inconsistent reporting happens when experiences are …%
Valence[a]	9 (12.5)	8 (11.1)	**55 (76.4)**	Negative: 69.1 Positive: 16.4 Neutral: 7.3
Impact[a]	8 (11.1)	10 (13.9)	**54 (75.0)**	High impact: 44.4 Low impact: 27.8 Medium impact: 25.9 Low and high impact: 1.9
Severity[b]	10 (14.1)	10 (14.1)	**51 (71.8)**	High severity: 51.0 Low severity: 27.5 Medium severity: 21.6
Developmental stage[b]	6 (8.5)	14 (19.7)	**51 (71.8)**	Childhood: 41.2 Adulthood: 39.2 Adolescence: 13.7 Childhood and adulthood: 3.9 All: 2.0
Number of experiences[c]	13 (18.6)	20 (28.6)	**37 (52.9)**	Many: 54.1 Few: 32.4 Some. 13.5

Note. Characteristics' specifiers are based on refinements made by participants that choose the category "agree"; [a]N=72, [b]N=71, [c]N=70.

Table 3

Percentages and frequencies of design characteristics involved in inconsistent reporting

	Level of agreement			
Design characteristics	Disagree n (%)	Neither agree or disagree n (%)	Agree n (%)	Inconsistent reporting happens …%
Setting[a]	17 (23.9)	22 (31.0)	**32 (45.1)**	Public place: 65.6 Private place: 34.4
Time interval between T1 and T2[a]	10 (14.1)	11 (15.5)	**50 (70.4)**	12 or more months: 42.0 From 6 to 12 months: 26.0 From 1 to 6 months: 18.0 6 or more months: 6.0 Few days: 4.0 From 2 weeks to a month: 2.0 From to 2 weeks to 6 months: 2.0
Mode of data collection[a]	7 (9.9)	11 (15.5)	**53 (74.6)**	Interview: 64.8 Self-report: 31.5 Both: 3.7
Characteristics of the interviewer[a]	9 (12.7)	20 (28.2)	**42 (59.2)**	Same sex[c]: 16.7 Different sex[c]: 26.2 Close in age[c]: 26.2 Different age[c]: 16.7 Acquaintance[c]: 52.4 Stranger[c]: 26.2 Empathic: 23.8 Not empathic: 50.0
Change of the interviewer[b]	11 (15.9)	22 (31.9)	**36 (52.2)**	Different interviewer: 97.2 Same interviewer: 2.8

Note. Experiences' specifiers are based on refinements made by participants that choose the category "agree"; T1=time 1, T2=time 2; [a]N=71, [b]N=69, [c]Dyad interviewer/respondent.

Additionally, when interviewers change from T1 to T2, when the interviewer is not nice and is an acquaintance the report seemed to be more inconsistent. The setting was presented as an influential reason for 45.1% of the participants; especially inconsistent reports seemed to happen when assessment occurs in public places.

Nine participants added other reasons underlying inconsistent reporting, such as social desirability, participant's availability, honesty, self-protection, avoidance, malice and personal gains/benefits.

According to the cluster analysis (Table 4), reasons were grouped in two clusters. The first, that was labelled variables involved in inconsistent reporting, comprised valence, importance, severity, developmental stage, setting, time interval between T1 and T2, mode of administration, interviewers' characteristics, change of the interviewer, mood, personality characteristics, physical and mental health, substance use, memory, secrecy, shame, third parties protection, and help seeking. The second cluster, labelled as variables not involved in inconsistent reporting, included gender, age, education, marital status, employment, and income.

Table 4

Cluster analysis: Group membership

Case	2 Clusters
Valence	1
Importance	1
Severity	1
Developmental	1
Number experiences	1
Setting	1
Time interval (T1 and T2)	1
Mode of administration	1
Interviewers characteristics	1
Change of the interviewer	1
Gender	2
Age	2
Education	2
Marital status	2
Employment status	2
Income	2
Mood	1
Personality	1
Physical and mental health	1
Substance_use	1
Memory	1
Secrecy	1
Shame	1
Third parties	1
Denial	1
Help seeking	1

Lastly, participants were asked to name people with inconsistent reporting; almost 30% cannot provide any answer and 12% answered that they did not know. Almost all participants presented idiosyncratic labels, with few exceptions, i.e., indecisive (*n*=3), humans (*n*=3) and insecure (*n*=2). Other names suggested were: sources of (mis)information, chameleon, masked, opaque, grey zone, different group, variable, inconstant, contradictory group, uncertain, inconsistent, volatile, unstable, silence, liars, psychopaths, disturbed, requesting more attention, uninhibited, rewind, many tells, storytellers, recount of experiences, dual report, life stories, dreamers, survivors, transparency and private.

Discussion

This study arises as a contribution to understand perceptions and reasons involved in inconsistent reporting based on participants' opinions, who were asked about general behaviour (instead of being confronted with their own behaviour). Briefly, it had some interesting findings: participants seem to be aware that inconsistent reporting is a common behaviour, which highly impact on research, and that it is rooted in a varied set of reasons. Due to its novelty, it is difficult to compare our results to other similar researches, but other studies in the field allow the interpretation and discussion of the findings.

Literature about the frequency and the pattern of inconsistent reporting is far from being unambiguous. For instance, Hepp et al. (2006) concluded that inconsistent reporting of potentially traumatic events was around 64%, whereas Ayalon (2017) provided less worrying data, i.e., 20% of their participants reported inconsistently at least one negative early life event. The same applies to the pattern of responses: comparing reports from T1 to T2, Dube et al. (2004) observed mixed patterns (both under or overreporting) across experiences, although there is a more common overall trend toward underreporting (e.g., Hardt & Rutter, 2004). Results about participants' perceptions are similar to those conclusions, suggesting that the behaviour is neither simple, nor linear. Indeed, it can present different configurations: some studies or variables are more prone to inconsistent reporting than others or have a high risk for underreporting whereas others promote overreporting, that are not still well studied.

Combining these data with results about reasons, it is reasonable to suggest that those differences may be related to the variables pointed out by participants to explain inconsistent reporting. Langeland et al. (2015) stressed reasons such as misunderstanding of the questions and avoidance. However their study is limited to variables associated with individual factors. The present research went further and addressed other potential variables not only focused on individuals, but also on design and experiences' characteristics, allowing for a more comprehensive perspective. Based on our participants' opinions, two clusters of variables were identified: the first seems to group those variables involved in inconsistent reporting and included individual variables, design variables and experience variables, whereas the second seems to group those variables not involved in inconsistent reporting and it comprises only sociodemographic variables. Notably, these patterns match the empirical inferential data available: According to a previous review (Carvalho, 2015), there was a trend of nonsignificant results concerning sociodemographic variables (despite of being widely studied) and the underlying factors were quite heterogeneous, involving not only individual characteristics but also design and experiences' characteristics.

Regarding individual's variables, memory was identified by our participants as a key-variable in inconsistent reporting, as noted also by Sobell et al. (1990) and Langeland et al. (2015). It was not a surprising result: usually participants claimed that they never thought about a certain experience before and then when asked about it they remember or they are not very confident about their memories. Currently, there are a couple of reviews about autobiographical memory (e.g., Fivush, 2011; Koriat et al., 2000); nevertheless, as Hardt and Rutter (2004) also noticed, it is still a major (devaluated) challenge regarding the reports of life experiences, considering that most evidence relies on experimental studies (e.g., Sarwar et al., 2014). Our participants also emphasized a quite common variable in this field of research, namely mood. According to our results, almost 75% agreed that mood, especially negative mood, impacts on inconsistent reporting, however inferential data is less straight: for instance, Schraedley and colleagues (2002) concluded that improvement, but not worsen, in depression status promotes inconsistency on childhood traumatic events. Other researchers (e.g., Fergusson et al., 2000; Paivio, 2001) claimed that mood was not an influential variable. Until now, a group of variables (e.g., shame, secrecy) remained merely speculative and, usually, they are proposed as explanatory reasons in discussion section.

Our participants seem to confirm these conjectures, suggesting that those variables are among the most important at the individual level.

Moreover, our results suggest that design and experiences' variables seem to play a major role in inconsistent reporting. Participants stressed particularly those variables associated with experiences, such as valence and impact, that as far as we know have been a little devaluated by researchers, with a few exceptions [e.g., Krinsley et al. (2003) studied the developmental stage; and McKinney and colleagues (2009) studied the total number of events]. The same applies to design variables: There are few studies and their conclusions do not corroborate our results. For instance, studies about time interval between T1 and T2 (Andresen et al., 2003; Assche, 2003) and interviewer features (Fry et al., 1996; Mills et al., 2007; Weinreb, 2006) did not achieved significant results, whereas the majority of our participants suggested that these were key variables. In line with our results, Pessalacia and colleagues (2013), using a cross-sectional design, asked participants to imagine they were enrolled on a research and, then, presented a predefined list of themes and situations asking them to rate the potential of embarrassment involved in each one. Authors concluded that some topics, i.e., betrayal, violence, and the death of a close one, are prone to cause embarrassment, but the conditions (i.e., dismissed information about themes under investigation; concerns about confidentiality and anonymity, or the capture of images or the audio-recording) were rated as more relevant. Taking in consideration all results, we can conclude that inconsistent reporting is a multi-dimensional phenomenon, involving many and complex interactions among variables, as suggested by Carvalho (2015).

Until now we did not know what the general opinion about the topic was. Curiously, participants not only recognized the behaviour, but also its impact. Naming the group of people with inconsistent reporting was a hard task – even among researchers there are no consensual term (e.g., inconsistency, unreliability, stability), but more important than this is the fact that suggested labels seem to denote two different approaches: for some participants inconsistent reporting is part of human-being and, consequently, it is an ordinary behaviour; oppositely , other participants seem to attribute some kind of malignancy and distrust to inconsistent reporters and, accordingly, inconsistent reporting occurs on purpose. This duality was also pointed out by Carvalho (2015) and according to McAdams (2001, p. 662) "it is likely that individual differences in the ways in which people narrate self-defining memories reflect both differences in the objective past and differences in the styles and manners in which people choose to make narrative sense of life".

Limitations and future research

Despite the strengths, some limitations should be stressed. First, results were based on a small sample, which limits the generalizability of the findings. Nevertheless, this sample should not be devaluated: participants were adults from the community; they were enrolled in a longitudinal study about life experiences, being familiar with the topic and the task; and they were asked about general behaviour. Additionally, they were recruited from schools and despite their heterogeneity, they may be especially sensitive for collaborating in research projects. Forthcoming studies should replicate and extend our results with other samples. Other potential limitation was the the measure applied, that was especially design for this study, in the absence of any known measure about the topic. Additionally, although parameters such as Cronbach alpha are not appropriate here, its reliability and validity should be analysed in detail using alternative strategies. Despite this limitation, due to its strengths (i.e., it can be easily adapted to a single or a set of life experiences, and applied to different target-groups) it can be a useful tool in future. Another limitation is related to cluster analysis' criticisms, i.e., different clustering methods and similarity measures can produce different clusters; results tend to be influenced by the order of variables; and clusters are quite sensitive to cases' dropouts (Field, 2000). Despite of them, since few statistical analyses allow for a more

complex outlook of categorical variables, we decided to compute cluster analysis, providing as many information as possible for evaluation and replication's purposes (Clatworthy et al., 2005).

Implications

This study has several implications that deserve a comment. As researchers we embrace total power and control toward the design and we make great efforts to maximize homogeneity (e.g., mode of data collection). Nonetheless, our results suggested that perhaps this imbalance is not the best approach: Design variables seem to influence inconsistent reporting of life experiences, and these potential effects can be minimized if power and control are shared with participants. Our study also stressed the influence of many variables that until now were merely speculative (e.g., shame); it has obvious implications not only for data collection or assessment, but also for data analysis. Attending to the link between research and practical applications in legal contexts, our results are even more informative and demanding. Lastly, inconsistent reporting may be a regular or a tricky behaviour, which occurs frequently, and should be addressed regularly in both research and practical contexts: for instance, individuals could be allowed to change previous answers and efforts should be made to understand changes.

In sum, inconsistent reporting of life experiences cannot be undervalued and the phenomenon is far from being deeply understood. This study represents one exploratory step for understanding it. We hope that it will serve as a prompt for more systematic research and, consequently, for the improving of confidence about retrospective life experiences reports.

References

Acuña, E., & Rodriguez, C. (2004). The treatment of missing values and its effect on classifier accuracy. In D. Banks, F. R. McMorris, P. Arabie, & W. Gaul (Eds.), *Classification, clustering, and data mining applications* (pp. 639-647). Springer.

Andresen, E. M., Catlin, T. K., Wyrwich, K. W., & Jackson-Thompson, J. (2003). Retest reliability of surveillance questions on health related quality of life. *Journal of Epidemiology and Community Health*, *57*, 339-343. https://doi.org/10.1136/jech.57.5.339

Assche, S. B. van. (2003). Are we measuring what we want to measure? Individual consistency in survey response in rural Malawi. *Demographic Research*, *9*(1), 77-108. https://doi.org/10.4054/DemRes.2003.S1.3

Ayalon, L. (2017). Retrospective reports of negative early life events over a 4-year period: A test of measurement invariance and response consistency. *The Journals of Gerontology: Series B, 72*(5), 901-912. https://doi.org/10.1093/geronb/gbv087

Azevedo, V., Martins, C., & Maia, A. (2020). Lifetime Experiences Scale (LIFES): Development and validation in community samples. *Psicologia, 34*(1), 174-195. https://doi.org/10.17575/psicologia.v34i1.1456

Canavarro, M. C. (2007). Inventário de sintomas psicopatológicos (BSI): Uma revisão crítica dos estudos realizados em Portugal. In M. R. Simões, C. Machado, M. M. Gonçalves, & L. S. Almeida (Eds.), *Avaliação psicológica: Instrumentos validados para a população portuguesa* (pp. 305-330). Quarteto Editora.

Carvalho, M. (2015). *Perceções e motivos das inconsistências no autorrelato de experiências de vida* [Tese de mestrado não publicada]. Universidade do Minho, Braga.

Casey, R. L., Masuda, M., & Holmes, T. H. (1967). Quantitative study of recall of life events. *Journal of Psychosomatic Research*, *11*(2), 239-247. https://doi.org/10.1016/0022-3999(67)90013-X

Clatworthy, J., Buick, D., Hankins, M., Weinman, J., & Horne, R. (2005). The use and reporting of cluster analysis in health psychology: A review. *British Journal of Health Psychology*, *10*, 329-358. https://doi.org/10.1348/135910705x25697

Cleary, P. J. (1980). A checklist for life event research. *Journal of Psychosomatic Research*, *24*(3-4), 199-207. https://doi.org/10.1016/0022-3999(80)90042-2

Colman, I., Kingsbury, M., Garad, Y., Zeng, Y., Naicker, K., Patten, S., Jones, P. B., Wild, T. C., & Thompson, A. H. (2016). Consistency in adult reporting of adverse childhood experiences. *Psychological Medicine*, *46*(3), 543-549. https://doi.org/10.1017/S0033291715002032

Dube, S. R., Williamson, D. F., Thompson, T., Felitti, V. J., & Anda, R. F. (2004). Assessing the reliability of retrospective reports of adverse childhood experiences among adult HMO members attending a primary care clinic. *Child Abuse & Neglect*, *28*(7), 729-737. https://doi.org/10.1016/j.chiabu.2003.08.009

Fergusson, D. M., Horwood, L. J., & Woodward, L. J. (2000). The stability of child abuse reports: A longitudinal study of the reporting behaviour of young adults. *Psychological Medicine*, *30*(3), 529-544. https://doi.org/10.1017/S0033291799002111

Field, A. (2000). *Cluster analysis.* http://www.statisticshell.com/docs/cluster.pdf

Fivush, R. (2011). The development of autobiographical memory. *Annual Review of Psychology*, *62*, 559-582. https://doi.org/10.1146/annurev.psych.121208.131702

Fry, R. P. W., Rozewicz, L. M., & Crisp, A. H. (1996). Interviewing for sexual abuse: Reliability and effect of interviewer gender. *Child Abuse and Neglect*, *20*(8), 725-729. https://doi.org/10.1016/0145-2134(96)00060-9

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*(1), 549-576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Hardt, J., & Rutter, M. (2004). Validity of adult retrospective reports of adverse childhood experiences: Review of the evidence. *Journal of Child Psychology and Psychiatry*, *45*(2), 260-273. https://doi.org/10.1111/j.1469-7610.2004.00218.x

Hepp, U., Gamma, A., Milos, G., Eich, D., Ajdacic-Gross, V., Rössler, W., Angst, J., & Schnyder, U. (2006). Inconsistency in reporting potentially traumatic events. *British Journal of Psychiatry*, *188*(3), 278-283. https://doi.org/10.1192/bjp.bp.104.008102

Kendall-Tackett, K., & Becker-Blease, K. (2004). The importance of retrospective findings in child maltreatment research. *Child Abuse and Neglect*, *28*(7), 723-727. https://doi.org/10.1016/j.chiabu.2004.02.002

Ketchen, D. J. Jr., & Hult, G. T. (2000). Validating cluster assignments. *Psychological Reports*, *87*(3), 1057-1058. https://doi.org/10.2466/pr0.2000.87.3f.1057

Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology*, *51*, 481-537. https://doi.org/10.1146/annurev.psych.51.1.481

Krinsley, K. E., Gallagher, J. G., Weathers, F. W., Kutter, C. J., & Kaloupek, D. G. (2003). Consistency of retrospective reporting about exposure to traumatic events. *Journal of Traumatic Stress*, *16*(4), 399-409. https://doi.org/10.1023/A:1024474204233

Langeland, W., Smit, J. H., Merckelbach, H., de Vries, G., Hoogendoorn, A. W., & Draijer, N. (2015). Inconsistent retrospective self-reports of childhood sexual abuse and their correlates in the general population. *Social Psychiatry and Psychiatric Epidemiology*, *50*(4), 603-612. https://doi.org/10.1007/s00127-014-0986-x

McAdams, D. P. (2001). *The person: An integrated introduction to personality psychology*. Harcourt College Publishers.

McKinney, C., Harris, T., & Caetano, R. (2009). Reliability of self-reported childhood physical abuse by adults and factors predictive of inconsistent reporting. *Violence and Victims*, *24*(5), 653-668. https://doi.org/10.1891/0886-6708.24.5.653

Mills, K., Teesson, M., Darke, S., & Ross, J. (2007). Reliability of self-reported trauma exposure among people with heroin dependence: A longitudinal investigation. *Journal of Traumatic Stress*, *20*(3), 313-323. https://doi.org/10.1002/jts.20219

Paivio, S. C. (2001). Stability of retrospective self-reports of child abuse and neglect before and after therapy for child abuse issues. *Child Abuse & Neglect*, *25*(8), 1053-1068. https://doi.org/10.1016/S0145-2134(01)00256-3

Paykel, E. S. (1983). Methodological aspects of life events research. *Journal of Psychosomatic Research*, *27*(5), 341-352. https://doi.org/10.1016/0022-3999(83)90065-X

Paykel, E. S. (2001). The evolution of life events research in psychiatry. *Journal of Affective Disorders*, *62*(3), 141-149. https://doi.org/10.1016/S0165-0327(00)00174-9

Pereira, F., Matos, M., Sherinda, L., & Scott, A. J. (2015). Perceptions and personal experiences of unwanted attention among Portuguese male students. *Psychology, Crime, & Law*, *21*(4), 398-411. https://doi.org/10.1080/1068316X.2014.989167

Pessalacia, J. D. R., Ribeiro, C. R. de O., & Massuia, D. (2013). Themes and situations that cause embarrassment among participants in research in which questionnaires or interviews are used. *Investigación y Educación en Enfermería*, *31*(1), 70-77.

Sarwar, F., Allwood, C. M., & Innes-Ker, Å. (2014). Effects of different types of forensic information on eyewitness' memory and confidence accuracy. *The European Journal of Psychology Applied to Legal Context*, *6*, 17-27. https://doi.org/10.5093/ejpalc2014a3

Schraedley, P. K., Turner, R. J., & Gotlib, I. H. (2002). Stability of retrospective reports in depression: Traumatic events, past depressive episodes, and parental psychopathology. *Journal of Health and Social Behavior*, *43*(3), 307-316. https://doi.org/10.2307/3090206

Sobell, L. C., Toneatto, T., Sobell, M. B., Schuller, R., & Maxwell, M. (1990). A procedure for reducing errors in reports of life events. *Journal of Psychosomatic Research*, *34*(2), 163-170. https://doi.org/10.1016/0022-3999(90)90050-E

Spinhoven, P., Bamelis, L., Haringsma, R., Molendijk, M., & Arntz, A. (2012). Consistency of reporting sexual and physical abuse during psychological treatment of personality disorder: An explorative study. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*(1), S43-S50. https://doi.org/10.1016/j.jbtep.2011.02.005

Weinreb, A. A. (2006). The limitations of stranger-interviewers in Rural Kenya. *American Sociological Review*, *71*(6), 1014-1039. https://doi.org/10.1177/000312240607100607

Relato inconsistente das experiências de vida: O que é que as pessoas pensam e como o explicam?

A avaliação de experiências de vida tem-se baseado principalmente no design retrospetivo transversal, apesar das preocupações com as inconsistências no relato. A investigação tem sugerido que alguns indivíduos alteram as suas respostas quando questionados repetidamente sobre as suas experiências de vida, embora as perceções do senso-comum sobre esse comportamento permaneçam desconhecidas. Este estudo procurou explorar as perceções individuais sobre o relato inconsistente e identificar os motivos associados a este comportamento, incluindo características individuais, da experiência e do design. Setenta e dois indivíduos, a participarem num estudo longitudinal sobre experiências de vida, responderam a um questionário sobre perceções gerais e motivos envolvidos. Os participantes pareceram estar cientes de que o relato inconsistente é um comportamento comum, com grande impacto na investigação. A análise de clusters indicou existirem dois clusters: variáveis envolvidas *vs.* não envolvidas no relato inconsistente. A maioria dos participantes discordou que as variáveis sociodemográficas influenciam a inconsistência do relato, enquanto a memória, o humor, a valência, o impacto, a estratégia de recolha de dados e as caraterísticas do/a entrevistador/a foram identificadas como variáveis-chave. Os resultados sugerem que o relato inconsistente não configura um fenómeno simples e provavelmente está enraizado num conjunto variado e complexo de variáveis.

Palavras-chave: Adulto, Acontecimentos de vida, Fiabilidade, Práticas de relato, Avaliação retrospetiva.