Soc Psychiatry Psychiatr Epidemiol
DOI 10.1007/s00127-014-0986-x

ORIGINAL PAPER

Inconsistent retrospective self-reports of childhood sexual abuse and their correlates in the general population

Willemien Langeland · Jan H. Smit · Harald Merckelbach · Gerard de Vries · Adriaan W. Hoogendoorn · Nel Draijer

Received: 11 October 2013 / Accepted: 10 November 2014

Abstract
Purpose Epidemiological research on childhood sexual abuse (CSA) and its consequences in adult life mainly relies on retrospective reports. This study explores their consistency and the correlates of inconsistent CSA self-reports in a random population sample.
Method A stratified subsample of 2,462 subjects (selected from a large-scale ($N = 34{,}267$) representative sample of Dutch adults aged 40 and beyond) participated in a two-phase online questionnaire survey on extra-familial CSA which was conducted on a four- to six-week interval. Subjects reporting CSA were overrepresented. Participants with consistent and inconsistent responses were compared with regard to demographics, family background, abuse severity, and clinical characteristics. Potential correlates of inconsistency were identified using logistic regression analysis. An additional questionnaire (Phase III) administered to inconsistent respondents explored possible reasons for their inconsistency.
Results Of the 1,992 respondents who had reported extra-familial CSA during Phase I, 707 (35.5 %) denied this in Phase II. Of the 2,462 respondents in Phase II, 727 (29.5 %; 9.2 % when considering sample stratification) gave a discrepant answer to the extra-familial sexual abuse item compared to their answers given in Phase I. Reports of less severe abuse, intra-familial CSA, and early parental separation predicted inconsistency. Reasons provided for inconsistency varied from misunderstanding (e.g., reporting intra-familial CSA rather than extra-familial CSA) to emotional motives (e.g., embarrassment, being overwhelmed) or practical considerations (e.g., lack of privacy while filling out the questionnaire).
Conclusions Inconsistent self-reports of extra-familial sexual abuse occur on a substantial scale and are associated with less severe forms of abuse (lack of salience) or classification difficulties (perpetrator being a family member or not). Consistency tests and probing for clarifications or corrections should be routinely conducted in order to increase the quality of CSA epidemiological research.

Keywords Childhood sexual abuse · Survey · Adults · Reporting practices · Consistency

Electronic supplementary material The online version of this article (doi:10.1007/s00127-014-0986-x) contains supplementary material, which is available to authorized users.

W. Langeland · J. H. Smit · N. Draijer (✉)
Department of Psychiatry and EMGO+ Institute, Vrije University Medical Center/GZZinGeest, A. J. Ernststraat 1187, 1081 HL Amsterdam, The Netherlands
e-mail: N.Draijer@ggzingeest.nl

J. H. Smit · A. W. Hoogendoorn · N. Draijer
Research Department GGZinGeest, Amsterdam, The Netherlands

H. Merckelbach
Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

G. de Vries
Scientific Council For Government Policy (WRR) The Hague, The Netherlands and Department of Philosophy, University of Amsterdam, Amsterdam, The Netherlands

Introduction

Childhood sexual abuse (CSA) represents an increased risk for psychopathology [1]. Therefore, research on risk factors often includes retrospective assessments of childhood adversities such as emotional, physical, and sexual abuse and neglect. The quality of retrospective self-reports,

Published online: 22 November 2014

however, can be compromised not only by recall errors such as omissions and biased retrieval [2], but also by a reluctance to report past abuse [3, 4]. These factors may lead to misclassification of research participants and ultimately distort the associations between CSA reports and adult psychopathology.

Limitations of retrospective CSA recall are illustrated by studies comparing documented CSA cases with adult self-reports. Typically, these studies find that a substantial proportion (a fifth to a third) of participants with a history of CSA fail to report such abuse when inquired in adult life [e.g., 5, 6]. Factors found to predict report failure were male gender, young age when the abuse happened, close relationship with the perpetrator, less severe abuse, and lack of maternal support [5–7].

Studies explicitly designed to test the stability of retrospective CSA reports provide information about the nature and determinants of reporting failures. Some relied on small clinical or forensic samples recruited from specific settings such as urban health care clinics, psychiatric services, and unaccompanied refugee minors [8–11]. Other studies looked at pre-post treatment changes in clinical samples [12, 13]. These clinical and forensic studies on the consistency of CSA reports may be skewed in the direction of relatively more severe abuse or pathological cases [14]. In addition, patients' memory performance and/or willingness to disclose CSA may not be representative of the general population. For these reasons, it is important to test the stability of retrospective CSA reports in general population samples [15–18]. Consistency studies among general population samples show discrepancies in self-reports of CSA on a non-trivial scale. In these samples, retrospectively obtained histories of CSA are often unstable over time, with kappa's ranging from 0.24 to 0.64, and percentages of agreement varying from 76 to 91 % [15–18]. Inconsistent reporting about autobiographical events has also been documented in general population samples for childhood physical abuse [14, 16], positive and negative childhood experiences [e.g., 19], life events [e.g., 20, 21], traumatic events [22], lifetime Axis I diagnoses [e.g., 23], and suicide attempts [e.g., 24]. This indicates that inconsistency in self-reports is not specific for the recall of sexual abuse, but rather occurs with various types of items and may reflect discrepancies over time in how respondents interpret questions and retrieve and label autobiographical events [25].

Studies focusing on inconsistent self-reporting of general population respondents indicate that inconsistency rates are raised in male respondents and for events low in salience. Psychiatric symptoms do not seem to affect inconsistency rates [15–17]. Friedrich and colleagues [15] questioned a random sample of 610 adults about CSA twice over a period of approximately 20 months and found that 14.4 % of the respondents gave discrepant answers to the abuse items. More specifically, 27 % of the total CSA reports at initial testing were not reported at follow-up, whereas 28 % of the total CSA reports at follow-up displayed the reverse form of inconsistency. Men were more likely to provide inconsistent reports than women (odds ratio = 4.4, 95 % *CI* (2.1, 9.3) $p < 0.0001$). Less severe forms of sexual abuse (i.e., non-contact) were less consistently reported than severe forms.

Fergusson and colleagues [16] repeatedly surveyed a representative birth cohort of approximately 1,000 young adults about CSA experiences over a 3-year period. Of those reporting CSA events at the age of 18, about half failed to subsequently report these events at age 21, and vice versa. Of the 137 respondents who reported CSA at either wave, 33.8 % were consistent. Inconsistent CSA reports were not associated with demographic characteristics, family background, or respondents' psychiatric status. In a second study [17], the researchers developed a structural equation model to estimate the effects of reporting errors and bias in retrospective abuse reports in their birth cohort. For CSA, measurement unreliability accounted for between 46 and 52 % of reporting variance; and reporting bias accounted for less than 1 % of this variance, with the remainder due to true abuse exposure. In addition, an important finding of this second study is that although retrospective CSA reports are subject to substantial errors of measurement, these errors do not pose a significant threat to the validity of estimates of the associations between exposure to childhood abuse and adult mental health.

Hepp et al. [22] also failed to detect an association between current psychological symptoms and inconsistent trauma reports in their community-based cohort ($N = 342$).

In recent years, internet-based surveys have become increasingly popular as a cost-effective way of gathering epidemiological data [26, 27]. Internet research has shown to be useful in collecting data on sensitive topics such as trauma [e.g., 28]. Some studies even suggest that victims prefer answering questions about adverse childhood experiences online compared to face-to-face questioning [29]. Given the potential of internet surveys to screen general population samples for CSA, it is important to establish to what extent inconsistent self-reports occur with such surveys. Arguably, inconsistent self-reports will undermine the integrity of assignments to CSA and control groups [26] and it will complicate longitudinal comparisons. Apart from that, temporal instability in CSA reports and their correlates are important research topics in themselves.

To our knowledge, the consistency of CSA reports in an online access general population sample has never been systematically investigated. The aims of the present study were (1) to investigate how many of the self-reports of extra-familial CSA are inconsistent across two time points in an internet panel of adult respondents, and (2) to examine whether demographic factors, abuse severity, severity of psychiatric symptoms, and the tendency to over-endorse symptoms predict inconsistent reports.

Method

Sample

Participants were drawn from the general population survey conducted by the Commission of Inquiry into sexual abuse of minors in the Dutch Catholic Church [30]. This expert group collected relevant data in 2010 and 2011 with the help of survey agency TNS NIPO using web-based procedures. The general focus of the study was the prevalence, characteristics, and impact of CSA in the Dutch Roman Catholic Church among adults age 40 and beyond. These adults were children during the 1950s, 1960s, and 1970s, a period in which the Catholic Church played a prominent role in the daily lives of many (roughly 40 %) Dutch families, in particular in the South of The Netherlands. A peak in levels of self-reports of CSA by representatives of the Catholic Church had occurred in these cohorts in the voluntary reports made to the Dutch Commission of Inquiry [30, p. 60–61]. To make the age distribution as similar as possible to that of the sample of the volunteers, the general population sample consists of adults age 40 and beyond.

Participants completed online questionnaires addressing their demographic characteristics, psychiatric symptoms, and extra-familial sexual abuse experiences. The data for the current study focusing on inconsistencies in CSA reports were gathered in three subsequent questionnaires referred to as Phase I, II, and III. Phase I was conducted between March and April 2011, Phase II took place in May 2011, and Phase III took place in June 2011. Response rates were high and varied from 85 to 91 %.

In Phase I, a random sample of Dutch adults aged 40 and beyond ($N = 34{,}267$) was screened for a history of extra-familial sexual abuse, a history of childhood institutionalization, and a Roman Catholic background. In Phase II, a selected stratified sample of subgroups ($N = 2{,}462$) based on these three characteristics was questioned about the presence and nature of CSA, using similar questions as in Phase I. This time, psychiatric symptoms were assessed as well. Because the second phase intended to draw a subsample of selected target groups, including non-abused control subjects, the majority ($n = 1{,}992$) had reported extra-familial sexual abuse at Phase I, reflecting an over-representation of CSA reporting subjects. The selection process is described in detail elsewhere [30; see also: www.onderzoekrk.nl].

The current study is based on these 2,462 respondents (1,539 women; mean age = 56.3 years, SD = 10.2) who completed the survey questions at both phases. They came from households with a mean size of 2.4 persons (SD = 1.2). Social class varied from "professional" (415; 16.9 %) to "skilled" (1973; 80.1 %) and "semi and unskilled" (74; 3.0 %).

Inconsistent responders in this sample were questioned in a third phase about the reasons for their inconsistencies. This phase was introduced as follows: *"*You were invited for the study in May because of your answers in the monthly screening in March or April. It turned out that you answered differently at some questions in the screening on childhood experiences and mental health. (..) We are interested in the cause of these differences. (…) There might be various explanations for these differences. For our research project, it is very important to know which answer is the correct one in your case and what the reason was to answer differently. We hope that you understand the importance of collecting reliable data on this theme.*"*

Persons were included after indicating their voluntary participation. The survey met the ethical guidelines of the European Society of Opinion and Marketing Research (ESOMAR) World Research Codes.

Measurements

Demographics and family background

Demographics were assessed in Phase I. Current household size was defined as the number of persons in the household. Social class of respondents was classified using the standard scoring system of TNS NIPO based upon educational level and (the latest) occupation ranging from "professional" (A) to "semi and unskilled" (D).

Family background—assessed in Phase II—pertains to reports about childhood experiences: early separation from (a) parent(s), parental dysfunction, sexual abuse by a family member, physical abuse by a parent or other adult in the family household, and violence between parents before age 18. Five survey items flagged parental dysfunction (addressing whether a parent was depressed or anxious, had a psychiatric disorder, had a serious alcohol problem, had used drugs, and/or had attempted suicide). On the basis of these five items, a parental dysfunction severity index was constructed (range 0–5), with higher scores indicating more severe dysfunction reports.

Extra-familial child sexual abuse

Extra-familial CSA was defined as any sexual contact by non-familial adults with a child or adolescent under the age of 18.

The survey item about CSA was derived from the item used in the Netherlands Mental Health Survey and Incidence Study (NEMESIS) [31] and read as follows: "Before you were 18, were you approached sexually by an adult outside your family? This means: being touched sexually against your will, or being forced into sexual contact with someone else." Responses to these inquiries were coded as: never, once, sometimes, regularly, often, or very often. These responses were dichotomized to reflect the occurrence of at least once versus the absence (never) of sexual abuse by a non-family member.

The sexual abuse item used was similar in Phase I and II, but the introduction differed. In Phase I, the introduction to the abuse item was as follows: "The purpose of the following questions is to measure potentially disturbing youth experiences that Dutch people aged 40 and over had with people outside their family. Although some questions are personal, we would like to invite you to answer the question honestly. Your answers will be handled anonymously and your data will be held in strict confidence. The study is being performed on behalf of a non-profit organization. The questions are about experiences that might have happened to you before you were age 18. For instance, incidents involving people such as a teacher, a leader of a sports team or youth group, a nurse, a minister, a priest, a religious brother or nun, a verger, a doctor or other care provider, or some other adult from outside your family. If you find the questions difficult to answer, you have the option of skipping the question by selecting "Unwilling to participate."

In Phase II, the sexual abuse item was introduced as follows: "By 'approached sexually' we mean being touched sexually against your will, or being forced into sexual contact by someone else. By 'family member' we mean a blood relative or a relative by marriage, including foster or step parents, or a new partner of your father or mother. Thus, this question is about people who do not belong to this group." In this phase, items about the details of the abuse were added, asking respondents whether there had been one or more perpetrators, whether they had been victims of penetrative abuse or other types (e.g., unwanted sexual touching or sexual assault) of abuse, how often the abuse had taken place (once or more than once), whether they had been threatened by the perpetrator(s), and whether or not the abuse had spanned a longer period of time (for the items, see Supplement 1). Based on these items, a sexual abuse severity index (range 0–12) was construed, with higher scores reflecting more severe abuse reports.

In Phase III, inconsistent responders filled out a brief questionnaire so as to clarify which response to the pertinent sexual abuse item they wished to have recorded. They were also asked about the reasons for their inconsistency (for the items, see Table 4). Respondents, who reported extra-familial CSA in Phase I and Phase III, but not in Phase II, were asked to fill out similar questions about the details of the abuse as posed respondents in Phase II.

Current psychiatric symptoms and negative response bias

Current psychiatric symptoms and negative response bias were assessed in Phase II. In the current paper, we use these data to evaluate if inconsistent reporting is related to psychopathology or negative response bias. All 2,462 respondents filled in the Brief Symptom Inventory-Short Form (BSI-18) [32] assessing current psychiatric symptoms. This scale consists of 18 items rated on a 5-point scale (anchors: 0 = not at all; 4 = always; range 0–72) that tap into depression, anxiety, and somatization (Cronbach alpha = 0.92). Higher scores reflect higher levels of psychiatric symptoms.

To assess symptom over-endorsement (i.e., negative response bias), four items from the Wildman Symptom Checklist [33] were added, addressing improbable symptoms ("I have headaches that are so severe my feet hurt"; "The buzzing in my ears keeps switching from the left to the right"; "I notice that the color of objects around me keeps shifting"; and "I find myself frequently blacking out when I sit down."). These items were rated similar to the BSI-18 items (range 0–16) (Cronbach alpha = 0.58). We used a threshold score of 4 to identify respondents with raised levels of negative response bias. For further details, see [30].

Lifetime psychiatric symptoms and suicide attempts

In Phase II, respondents were given 13 items addressing psychiatric symptoms that have frequently been shown to be associated with sexual abuse or other traumatic experiences [34] and are included in the diagnostic criteria for (complex) posttraumatic stress disorders (DSM-5, ICD-10). Items rated lifetime sleep problems, nightmares, flashbacks, self-mutilative behavior, feelings of guilt or shame, anger outbursts, relational problems, and somatic complaints on the earlier mentioned 5-point scale (range 0–52) (Cronbach alpha = 0.90). A higher total score reflected higher levels of lifetime psychiatric problems. Additionally, respondents were given three questions addressing suicidal ideation and suicide attempts that were rated dichotomously (no/yes). These data were used to study whether inconsistent reporting was related to these symptoms.

To prevent interference, all participants were asked to rate the degree to which psychiatric symptoms bothered them. Subsequently, they were asked about the presence of childhood adversities.

Data analysis

Contingent on their responses to the extra-familial sexual abuse items in Phases I and II, respondents were categorized into two groups: consistent and inconsistent respondents. In line with a previous population-based survey [14], both percentage of agreement and Cohen's kappa [35] were used to assess the consistency of extra-familial CSA reports. Links between inconsistent reporting and demographics, family background and abuse characteristics, as well as clinical characteristics (psychiatric problems, negative response bias, and suicidal attempts) were evaluated with Chi-square or t tests where appropriate, and effect sizes (Phi coefficient φ or Cramer's V, or Cohen's' d) were calculated. Logistic regression analysis was used to identify predictors of inconsistent CSA reporting. Variables significant in univariate comparisons were included in the regression model.

Finally, the prevalence of inconsistency of extra-familial CSA reports was estimated taking into consideration the effects of sample stratification using a relatively simple formula of selection probability of respondents to be included in Phase II and percentages of false positives and false negatives.

Results

In Phase I, 1,992 (80.9 %) of the 2,462 subjects reported sexual abuse before age 18, whereas in Phase II, only 1,305 (53.0 %) reported such abuse (Table 1). In total, 727 (29.5 %) of the 2,462 respondents gave answers to the extra-familial sexual abuse item in Phase II that were discrepant with their answers to this item in Phase I. More than two-thirds (70.5 %) of the 2,462 respondents did report consistently about the presence or absence of extra-familial CSA before age 18 at both questionnaires. Of the 2,012 respondents who reported extra-familial CSA at either phase, 63.9 % were consistent. The overlap between responses to the pertinent items in Phase I and II in terms of kappa was 'fair' [36]; Kappa = 0.39, 95 % $CI \pm$ (0.36, 0.42)].

In the large majority of cases (97.3 % out of the 727 inconsistent respondents), inconsistencies involved an affirmative answer in Phase I and a negative answer in Phase II ("no, not abused"). Because in this sample CSA cases (affirmative answer to the pertinent item in Phase I) were oversampled, these data need to be analyzed in a proportional way: The proportion of inconsistencies among all respondents who reported CSA in Phase I was 35.5 % ($n = 1{,}992$), while the proportion of inconsistencies among all respondents reporting not being abused in Phase I ($n = 470$) was 4.3 % (χ^2 (1) = 551.81, $p < 0.0001$). Thus, inconsistencies occurred far more often in cases where respondents had reported extra-familial CSA (i.e., "yes"–"no") than in cases where they gave a negative answer in the first phase (i.e., "no"–"yes").

Table 1 Consistency of childhood sexual abuse reports of adults ($N = 2{,}462$) in Phase I (T1) and Phase II (T2)

CSA reported at T1	CSA reported at T2		
	No *N* (%)	Yes *N* (%)	Total
No	450 18.3 (95.7 %)	20 0.8 (4.3 %)	470 (100 %)
Yes	707 28.7 (35.5 %)	1,285 52.2 (64.5 %)	1,992 (100 %)
Total	1,157 (47.0 %)	1,305 (53.0 %)	2,462 (100 %)

Prevalence estimate

When considering the effects of sample stratification, the prevalence estimate of reporting inconsistency of CSA reports in our sample decreases from 29.5 % (727/2,462 = 0.2953 not considering the sample stratification) to 9.2 %. For the calculation of this latter estimate, we used the following formula:

We have a stratified (sub) sample of $n = 2{,}462$ out of eight strata consisting of a total of $N = 34{,}267$ respondents, in which respondents have been selected out of five strata of respondents. Because the focus is on the consistency of extra-familial CSA reporting, we conflate the eight strata to two strata:

- in stratum 1 ($N_1 = 5{,}370$ respondents), persons reported CSA in Phase I
- in stratum 2 ($N_2 = 28{,}897$ respondents), persons did not report CSA in Phase I.

Because the focus was on CSA reports, the selection probability per stratum differed:

- in stratum 1, $n_1 = 1{,}992$ respondents were selected with a probability of 1,992/5,370 = 0.3709
- in stratum 2, $n_2 = 470$ respondents were selected with a probability of 470/28,897 = 0.0163.

The data in Table 1 show the inconsistency of extra-familial CSA reports:

- in stratum 1, the fraction of inconsistent reports (false positives) is $p_1 = 707/1{,}992 = 0.3549$

- in stratum 2, the fraction of inconsistent reports (false negatives) is $p_2 = 20/470 = 0.0426$.

Considering the sample stratification, we use the strata-weights $W_h = N_h/N$, *or* $W_1 = 5{,}370/34{,}267 = 0.1567$ for stratum 1 and $W_2 = 28{,}897/34{,}267 = 0.8433$ for stratum 2. We estimate the fraction of inconsistent extra-familial CSA reporting $W_1p_1 + W_2p_2 = 0.1567 \times 0.3549 + 0.8433 \times 0.0426 = 0.0915$, or 9.2 %.

Associated characteristics

Table 2 gives an overview of demographics, family background, and clinical characteristics of consistent and inconsistent respondents. Of the demographic variables, only gender was associated with inconsistent reports: Women were more often inconsistent in their extra-familial CSA reports than men (χ^2 (1) = 9.38, $p < 0.01$; Phi coefficient φ or Cramer's $V = 0.06$).

As for family background variables, inconsistent extra-familial CSA reports were associated with reports of having been sexually abused by a family member (χ^2 (1) = 31.85, $p < 0.001$; Phi coefficient φ or Cramer's $V = 0.11$).

Table 2 Demographics, family background, abuse severity, and clinical characteristics of consistent and inconsistent respondents ($N = 2{,}462$)

Variable[a]	Respondents	
	Consistent ($n = 1{,}735$)	Inconsistent ($n = 727$)
Demographics		
Age (SD)	56.43 (10.19)	56.45 (10.50)
% Women	60.6	67.1**
Number of persons current household(*SD*)	2.4 (1.2)	2.5 (1.2)
% Highest social class (A)	16.7	17.3
Family background		
Early parental separation, % 'yes'	16.3	19.8*
Intra-familial CSA, % 'yes'	16.4	26.3***
Physical abuse, % 'yes'	24.2	21.9
Inter-parental violence, % 'yes'	25.3	22.7
Parental dysfunction (SD)	0.47 (0.85)**	0.37 (0.74)
Severity of extra-familial sexual abuse		
Abuse severity index (SD)	3.16 (2.87)***	1.33 (1.88)
Psychological symptoms		
Current symptoms (BSI) (SD)	8.17 (9.41)	8.21 (9.21)
Lifetime symptoms (SD)	21.60 (8.38)**	20.55 (7.71)
Lifetime suicide attempts, % 'yes'	8.8*	6.2
Negative response index (SD)	0.66 (1.42)	0.64 (1.47)

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$ (Chi-square analysis or *t* test)

[a] All variables except demographics were assessed in Phase II

In addition, compared to the consistent group, the inconsistent respondents more often reported early parental separation (χ^2 (1) = 4.37, $p < 0.05$; φ or $V = 0.04$), and less severe parental dysfunction [t (2,460) = 2.89, $p < 0.01$; Cohen's $d = 0.11$], with the 95 % confidence interval (*CI*) of the difference ranging between 0.03 and 0.18. Furthermore, the inconsistent group reported less severe forms of extra-familial CSA than the consistent group: t (2,460) = 15.82, $p < 0.001$; $d = 0.64$, with the 95 % CI of the difference ranging between 1.60 and 2.05.

Inconsistent and consistent groups did not differ with regard to levels of current psychiatric problems or over-endorsement of symptoms (all t's < 0.36; all p's > 0.72). However, the inconsistent group did report less lifetime symptoms [t (2,460) = 2.79, $p < 0.01$; $d = 0.11$, 95 % CI of the difference ranging between 0.34 and 1.76] and less suicide attempts (χ^2(1) = 4.79, $p < 0.05$; φ or $V = 0.04$) than did consistent respondents.

After testing for multicollinearity among independent variables, a logistic regression was performed that included the seven variables on which the two groups significantly differed. The full model was able to distinguish between the two groups ($p < 0.001$) and explained between 13.5 % (Cox and Snell *R square*) and 19.1 % (Nagelkerke *R square*) of the variance, correctly classifying 74.3 % of the cases. Only three of the independent variables made a unique and statistically significant contribution to the model, namely abuse severity, intra-familial CSA, and early parental separation (Table 3). The strongest predictor of consistency was intra-familial CSA, with an odds ratio of 2.60 (95 % CI 2.02–3.35).

Thus, respondents who reported intra-familial CSA were over two and half time more likely to report inconsistently about extra-familial abuse than those who did not report intra-familial CSA, controlling for all other factors in the model. The odds ratio's of 0.70 for both abuse severity and early parental separation indicated that respondents who reported less severe abuse or early separations from parents were slightly more likely to report inconsistently about extra-familial CSA, controlling for all other factors in the model.

Reasons for inconsistent reports

Of the 707 participants who changed their CSA reports from "yes" to "no," 614 (86.8 %) participated in the third phase as well as 19 (95 %) of the 20 subjects who changed their CSA reports from "no" to "yes".

Participants were provided with possible explanations (options listed in Table 4) for inconsistent CSA reports. The most frequently endorsed explanation was a misunderstanding or misreading of the pertinent sexual

Table 3 Logistic regression predicting likelihood of inconsistent reports about extra-familial child sexual abuse ($N = 2{,}462$)

	B	SE	Wald	*Df*	*P*	Odds ratio	95 % CI	
							Lower	Upper
Abuse severity index	−0.36	0.03	210.76	1	0.000	0.70	0.67	0.73
Intra-familial CSA	0.96	0.13	55.34	1	0.000	2.60	2.08	3.35
Early parental separation	−0.37	0.13	8.32	1	0.004	0.70	0.54	0.89
Gender	0.17	0.10	2.77	1	0.10	1.19	0.97	1.45
Parental dysfunction	−0.10	0.07	2.41	1	0.12	0.90	0.79	1.03
Lifetime suicide attempts	−0.15	0.20	0.53	1	0.47	0.86	0.58	1.29
Lifetime psychological symptoms	0.00	0.01	0.006	1	0.94	1.00	0.99	1.01
Constant	−0.05	0.19	0.06	1	0.80	0.96		

$R^2 = 0.14$ (Cox and Snell), 0.19 (Nagelkerke)

Model χ^2 (7, $N = 2{,}462$) = 355.26, $p < 0.001$

abuse question or the instructions, mostly during testing in Phase I (Table 4). Reasons given include reporting on intra-familial abuse rather than extra-familial abuse, reporting on sexual harassment or ambiguous experiences rather than 'hands on' sexual abuse, but also reporting on sexual abuse by a peer rather than by an adult or reporting on sexual abuse experienced after age 18 rather than experienced before age 18.

The second most frequently endorsed items pertained to mistakes during testing in Phase I or II. For example, respondents could not remember having endorsed or rejected the pertinent CSA item in Phase I or Phase II.

A considerable minority of inconsistent responders mentioned emotional reasons for discrepant reporting. That is, they felt overwhelmed, were reluctant to answer more questions about the topic, or did not want other persons present when responding to the item to know that they had been sexually abused.

Lastly, a variety of explanations for the discrepancies were given as "other reason(s)," including reevaluating the importance of the experience (respondents reporting that they thought it was not worth reporting the abuse because it was a minor incident or because it had no negative psychological consequences), not knowing whether abuse by a stepfather, an adoptive father or a brother in law should be categorized as intra- or extra-familial sexual abuse, embarrassment, or the wish to forget about the abuse.

Discussion

This study examines the occurrence and associated characteristics of inconsistent extra-familial CSA reports in a test–retest setup spanning several weeks in a large online community-based sample. First, we found modest levels of consistency (70.5 %) in reports of extra-familial CSA in a selective subsample in which sexual abuse reports were overrepresented. When we considered the effects of sample stratification, the percentage of inconsistent extra-familial CSA reports (29.5 %) dropped to an estimate of 9.2 % in the population sample. This finding replicates observations of other researchers who polled population-based samples about CSA experiences using different assessment methods (self-report questionnaire with multiple behavioral specific questions, face-to-face interview) [15, 16]. For example Friedrich et al. [15] found 14.4 %. Apparently, inconsistent CSA reports are quite common and not related to method of data gathering. However, previous research has illustrated that inconsistent reports are not specific for CSA experiences and might occur for, for example, suicidal thoughts [24]. Also, inconsistent reports on CSA are not specific for general population samples, but have also been documented in clinical and forensic samples [9–11].

The patterns we found for inconsistent reporting are in line with previous general population studies, namely reports of not having been sexually abused were more consistent than reports of having been sexually abused [15, 16; see for a similar pattern of reports about rape: 37].

Secondly, we found that inconsistent reporting was associated with a distinct set of factors, indicating that inconsistent recall of extra-familial CSA was more likely to arise when respondents also reported intra-familial CSA and/or early parental separation and when extra-familial CSA experiences had been less severe. This might also be the main reason why less severe parental dysfunction and less suicide attempts were reported by the inconsistent group of respondents—as these variables are related to less severe abuse. At univariate level, our study seems to contradict Friedrich et al. [15], who reported that men were more likely to provide inconsistent reports relative to women. But gender disappeared as a predictor of inconsistencies, controlling for other factors in the model. This could be related to the overrepresentation of women as CSA reporting subjects in our stratified sample. Still, the overall result pattern fits well with previous research, suggesting that respondents might find it difficult to classify CSA experiences—which may generate reclassification and re-interpretation on re-testing [25] —but that the intensity of the traumatic events is positively related to

Table 4 Frequencies and percentages of reasons given for inconsistent CSA reports provided by inconsistent responders ($n = 633$) (more answers possible)

Reported reasons	CSA report change			
	Yes to no ($n = 614$)		No to yes ($n = 19$)	
	N	%	*N*	%
I have understood this question in both questionnaires in different ways	122	19.9	6	31.6
I made a mistake in the initial questionnaire	54	8.8	4	21.1
I made a mistake in the second questionnaire	40	6.5	2	10.5
The initial and second questionnaire have not been filled out by the same person in my household	2	0.3	0	0
I was approached sexually in my childhood, but not by adults who were not related to me	30	4.9	NA	
I had negative experiences in my childhood, but at second thought, these were not of a sexual nature	20	3.2	NA	
I did not fill out both questionnaires seriously	3	0.5	0	0
I did not want to get and answer more questions about this subject	48	7.8	NA	
I felt overwhelmed, the subject stirred up too much memories	53	8.6	NA	
I felt overwhelmed by the subject in the initial questionnaire, but not in the second one	NA		2	10.5
I was not alone when filling out the second questionnaire and preferred the(se) others would not learn that I have been sexually abused in childhood by an adult person	20	3.3	NA	
I cannot remember	81	13.2	1	5.3
Other reason, namely	187	30.5	5	26.3

NA Not applicable

consistency [38]: more severe abuse being more consistently reported.

This is further underlined by the observation that inconsistent responders often said that misreading or misinterpretation of the introduction or the pertinent question had been the primary reason for their discrepant reports. For example, responders might say that they had had an abusing stepfather who was not yet a family member at the time the abuse took place and that they had doubted whether he should be considered a family member or not, or they wondered whether kissing is a sexual approach and worthwhile to mention. It is this type of ambiguity that seems to have fostered inconsistent reporting. To lesser extent, errors or emotional factors contributed to inconsistent reporting: 8 % of the inconsistent respondents said that they felt overwhelmed by the pertinent CSA items, a percentage that underlines the observation that people may be reluctant to classify themselves as an abuse victim. Overall, our data on self-reported reasons for inconsistent CSA reporting are in line with those reported by Cornelius et al. [26] concerning the online assessment of interpersonal violence.

The implication of our results for future studies is that longitudinal designs with repeated childhood abuse assessments should be concerned about the possibility of inconsistent reporting [21, 39]. Temporal misplacement of events or misunderstanding of items was given as the most important reason for inconsistent responses in our study, possibly due to the limitation of the investigation to a specific type (extra-familial) of CSA. Thus, one important way to reduce inconsistencies is to improve the phrasing of sexual abuse questions and provide more precise definitions and temporal contexts of the concepts used [40].

Our finding that consistency was greater for the most serious abusive and traumatic events is congruent with previous large-scale surveys on life events [20] as well as with research on highly traumatic childhood events among adolescents [9]. Taken together, the results imply that personally significant or more salient life events—traumatic or not—are reported more consistently over time. We also found that inconsistent CSA self-reports are not associated with poor mental health or symptom overendorsement. Similar results have been reported in other population-based studies [15, 16].

In the 1980s, underreporting of sexual abuse was the main problem in surveys, and the use of multiple behavioral questions was recommended to increase disclosure of CSA experiences [see for example, 4]. However, overreporting of CSA experiences might also occur. In our survey using a single item, most inconsistencies were of the "yes" to "no" type. In addition, although some authors have recommended indirect approaches such as the randomized

response technique [41] so as to counteract underreporting about sensitive topics, these techniques might not result in more realistic estimates of childhood sexual abuse victimization because of other reporting errors than non-disclosure due to embarrassment or social influence. The type of interpretational problems that our respondents believed to have created inconsistencies will likely also plague these indirect approaches. Other questionnaire design techniques might be helpful to produce reliable information in surveys on childhood sexual abuse. In our online CSA survey, consistency tests and probing for clarifications or corrections have been conducted in order to increase the quality of the collected data [39]. After correcting for inconsistencies in reports, we calculated prevalence estimates and analyzed associations between self-reports of CSA and psychiatric symptoms. Such a quality check may be useful for future (online) research on CSA. However, as researchers we have to live with the fact that retrospective assessment of traumatic events such as childhood abuse will remain open to a degree of imprecision due to the fallibility of memory and other (emotional) factors that affect abuse reporting.

Several limitations deserve comment. We did not obtain independent corroboration of CSA reports. Thus, our results are silent as to the validity of these reports. Our self-identified abuse victims may not represent abuse victims who are unwilling to volunteer for (online) research studies. In addition, comparisons between the first two stages of our survey are complicated by the fact that in Phase 1, the definition of a non-family member in the introduction to the abuse item was slightly less precisely worded than the definition used in Phase II. This might have led some respondents to react more accurately to the pertinent item in Phase II ('sorry, not non-familial abuse, but familial abuse') than in Phase I. Overall, our findings are in line with those of previous studies using samples of minors [e.g., 9] or young adults [15, 16]. So we do not think that there was an important effect of the age specification (adults aged 40 and beyond) in our sample on the outcomes of the study.

To sum up, inconsistent self-reports of non-familial sexual abuse occur on a non-trivial scale and are associated with less severe forms of abuse (lack of salience), ambiguity of questions, or classification difficulties (perpetrator being a family member or not). Consistency tests and probing for clarifications or corrections should be routinely conducted in order to increase the quality of CSA surveys.

Acknowledgments We are grateful to the respondents who participated in the survey and to the Commission of Inquiry into sexual abuse of minors in the Catholic Church in the Netherlands that made the database available for secondary analysis. We thank ZONMW (The Netherlands Organisation for Health Research and Development, The Hague) for providing the research grant under contract 16564.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Gilbert R, Widom CS, Browne K, Fergusson D, Webb E, Janson S (2009) Burden and consequences of child maltreatment in high-income countries. Lancet 373:68–81. doi:10.1016/S0140-6736(08)61706-7
2. Schwarz N, Sudman S (1994) Autobiographical memory and the validity of retrospective reports. Springer Verlag, New York
3. Della Femina D, Yeager CA, Lewis DO (1990) Child abuse: adolescent records vs. adult recall. Child Abuse Negl 14:227–231
4. Draijer N (1990) Seksuele traumatisering van meisjes door verwanten. [Intrafamilial sexual abuse of girls. A national survey]. SUA, Amsterdam
5. Williams LM (1994) Recall of childhood trauma: a prospective study of women's memories of child sexual abuse. J Consult Clin Psychol 62:1167–1176. doi:10.1037/0022-006X.62.6.1167
6. Goodman GS, Ghetti S, Quas JA, Edelstein RS, Alexander KW, Redlich AD et al (2003) A prospective study of memory for child sexual abuse: new findings relevant to the repressed-memory controversy. Psychol Sci 14:113–118
7. Widom CS, Morris S (1997) Accuracy of adult recollections of childhood victimization: part 2: childhood sexual abuse. Psychol Assessm 9:34–46
8. Ghetti S, Edelstein RS, Goodman GS, Cordon IM, Quas JA, Alexander KW, Redlich AD, Jones DPH (2006) What can subjective forgetting tell us about memory for childhood trauma? Memory Cogn 34:1011–1025
9. Aalsma MC, Aalsma MC, Zimet GD, Fortenberry JD, Blythe M, Orr DP (2002) Reports of childhood sexual abuse by adolescents and young adults: stability over time. J Sex Res 39:259–263
10. Dube SR, Williamsson DF, Thompson T, Felitti VJ, Anda RF (2004) Assessing the reliability of retrospective reports of adverse childhood experiences among adult HMP members attending a primary care clinic. Child Abuse Negl 28:729–737
11. Spinhoven P, Bean T, Eurelings-Bontekoe L (2006) Inconsistencies in the self-report of traumatic experiences by unaccompanied refugee minors. J Trauma Stress 19:663–673. doi:10.1002/jts.20152
12. Paivio SC (2001) Stability of retrospective self-reports of child abuse and neglect before and after therapy for child abuse issues. Child Abuse Negl 25:1053–1068
13. Kremers IP, Van Giezen AE, Van der Does AJ, Van Dyck R, Spinhoven P (2007) Memory of childhood trauma before and after long-term psychological treatment of borderline personality disorders. J Behav Ther Exp Psychiat 38:1–10
14. McKinney CM, Harris TR, Caetano R (2009) Reliability of self-reported childhood physical abuse by adults and factors predictive of inconsistent reporting. Viol Victims 24:653–668
15. Friedrich WN, Talley NJ, Panser L, Fett S, Zinsmeister AR (1997) Concordance of reports of childhood abuse by adults. Child Maltreat 2:164–171
16. Fergusson DM, Horwood LJ, Woodward LJ (2000) The stability of child abuse reports: a longitudinal study of the reporting behavior of young adults. Psychol Med 30:529–544
17. Fergusson DM, Horwood LJ, Boden JM (2011) Structural equation modeling of repeated retrospective reports of childhood

maltreatment. Int J Methods Psychiatric Res 20:93–104. doi:10.1002/mpr.337
18. Goldberg LR, Freyd JJ (2006) Self-reports of potentially traumatic experiences in an adult community sample: gender differences and test-retest stabilities of the items in a brief betrayal-trauma survey. J Trauma Dissociation 7:39–63. doi:10.1300/1229v07n03_04
19. Yancura LA, Aldwin CM (2009) Stability and change in retrospective reports of childhood experiences over a 5-year period: findings from the Davis Longitudinal Study. Psychol Aging 24:715–721. doi:10.1037/a0016203
20. Glickman L, Hubbard M, Liveright T, Valciukas JA (1990) Fall-off in reporting life events: effects of life change, desirability, and anticipation. Behav Med 16:31–38
21. Pachana NA, Brilleman SL, Dobson AJ (2011) Reporting life events over time: methodological issues in a longitudinal sample of women. Psychol Assessm 23:277–281. doi:10.1037/a0021337
22. Hepp U, Gamma A, Milos G, Eich D, Ajdacic-Gross V, Rössler W, Angst J, Schnyder U (2006) Inconsistency in reporting potentially traumatic events. Brit J Psychiat 188:278–283. doi:10.1192/bjp.bp.104.008102
23. Aneshensel CS, Estrada AL, Hansell MJ, Clark VA (1987) Social psychological aspects of reporting behavior: lifetime depressive episode reports. J Health Soc Behav 28:232–246
24. Christl B, Wittchen HU, Pfister H, Lieb R, Bronisch T (2006) The accuracy of prevalence estimations for suicide attempts. How reliably do adolescents and young adults report their suicide attempts? Arch Suicide Res 10:253–263
25. Loftus E, Joslyn S, Polage D (1998) Repression: a mistaken impression? Dev Psychopathol 10:781–792
26. Cornelius TL, Truba N, Bell KM (2011) Using the internet to prescreen participants for research on interpersonal violence: experimental designs considerations. Viol Victims 26:319–328
27. Easton SD (2012) Understanding adverse childhood experiences (ACE) and their relationship to adult stress among male survivors of childhood sexual abuse. J Prev Interv Community 40:291–303. doi:10.1080/10852352.2012707446
28. Read JR, Farrow SM, Jaanimägi U, Ouimette P (2009) Assessing trauma and traumatic stress via the internet: measurement equivalence and participant reactions. Traumatol 15:94–102. doi:10.1177/1534765608325121
29. DiLillo D, DeGue S, Kras A, Di Loreto-Colgan AR, Nash C (2006) Participant responses to retrospective surveys of child maltreatment: does mode of assessment matter? Viol Victims 21:410–424
30. Deetman W, Draijer N, Kalbfleisch P, Merckelbach H, Monteiro M, de Vries G (2011) Seksueel misbruik van minderjarigen in de Rooms-Katholieke Kerk. Uitgebreide versie, Deel 1 Het onderzoek. [Sexual abuse of minors in the Roman Catholic Church]. Amsterdam: Balans (http://www.onderzoekrk.nl/home.html)
31. Cuijpers P, Smit F, Unger F, Stikkelbroek Y, ten Have M, de Graaf R (2011) The disease burden of childhood adversities in adults: a population-based study. Child Abuse Negl 35:937–945. doi:10.1016/j.chiabu.2011.06.005
32. De Beurs E (2011) Brief Symptom Inventory (BSI) en BSI 18. Handleiding 2011 [Brief Symptom Inventory and BSI 18. Manual 2011]. Leiden: PITS
33. Wildman RW, Wildman RW (1999) The detection of malingering. Psychol Rep 84:386–388
34. Neumann DA, Houskamp BM, Pollock VE, Briere J (1996) The long-term sequelae of childhood sexual abuse in women: a meta-analytic review. Child Maltreat 1:6–16. doi:10.1177/1077559596001001002
35. Fleiss JL (1981) Statistical methods for rates and proportions, 2nd edn. Wiley, New York
36. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174
37. Scott KD, Aneshensel CS (1997) An examination of the reliability of sexual assault reports. J Interpers Viol 12:361–374
38. Krinsley KE, Gallagher JG, Weathers FW, Kutter CJ, Kaloupek DG (2003) Consistency of retrospective reporting about exposure to traumatic events. J Trauma Stress 16:399–409
39. Pereira da Silva SS, Da Costa Maia Â (2013) The stability of self-reported adverse experiences in childhood: a longitudinal study on obesity. J Interpers Viol 28:1989–2004. doi:10.1177/0886260512471077
40. Fricker AE, Smith DW, Davis JL, Hanson RF (2003) Effects of context and question type on endorsement of childhood sexual abuse. J Trauma Stress 16:265–268
41. Warner SL (1965) Randomized response. A survey technique for eliminating evasive answer bias. J Am Statistical Assoc 60:63–69