INSIGHTS

MONITORING DRUG USE IN THE DIGITAL AGE: STUDIES IN WEB SURVEYS

Quantitative and qualitative assessments of data quality within the European Web Survey on Drugs

Eleni Kalamara and Kateřina Škařupová

European Monitoring Centre for Drugs and Drug Addiction

Abstract: This paper assesses data quality in the first two rounds of the European Web Survey on Drugs (EWSD). Data quality has various dimensions, such as accuracy, consistency, validity and completeness, but can be defined as fitness of the data for the intended use. For the quantitative assessment, the paper considers issues such as respondent fatigue and impact of the item format on responses. In conjunction, the authors conducted a qualitative data quality assessment, within a broader project that aimed to assess the validity and reliability of the Czech version of the questionnaire used in the first round of the EWSD. The paper shows how these assessments were used to improve the design of the questionnaire used in the second round of the EWSD, which appear to have yielded more accurate responses to the survey. This underscores the value of conducting data quality assessments on an ongoing basis to ensure that the questionnaire continues to reflect current drug use practices, which are fluid and may rapidly change. Overall, the paper provides insight into the validity of the EWSD, while also underscoring the importance of data quality assessments in future web survey data collection on drug use and related behaviours.

Introduction

Web surveys, despite their many advantages and increasing use, have historically raised concerns about their representativeness, generalisability of results and data quality. Utilising both quantitative and qualitative approaches, this paper assesses data quality in the first two rounds (2016 and 2017-2018) of the European Web Survey on Drugs (EWSD), demonstrating how these data quality assessments were used to improve the design of the questionnaire between the two survey rounds.

Data quality is a broad term that relates to how much confidence we can have in the data and can be simply defined as the fitness of data for the intended use (Sidi et al., 2012). Behind this definition, data quality has various dimensions, including accuracy, consistency, validity and completeness (Sidi et al., 2012). A variety of quantitative indicators can be used to assess these different dimensions. For example, consistency (or equivalence of response) can be checked by assessing similarity in responses across different items asking for the same information in the same questionnaire or between modes of data collection (Dillman and Bowker, 2001). Response rates can be used to assess data accuracy, as the bigger the sample (and thus, the higher the response rate), the more probable it is that the data accurately estimates the true population parameters in representative surveys (Dillman and Bowker, 2001).

Data quality can be affected by various factors such as survey design and administration (e.g. questionnaire administration mode and item format), but also by factors on the respondents' side (e.g. fatigue and motivation). Beyond quantitative assessments of data quality, the validity of a questionnaire greatly depends on how respondents interpret and understand a survey's questions and the answer options. This understanding is, in turn, affected by the respondents' experience, social context and social identity (Miller et al., 2014). This process can be broken down into four distinct steps known as the question-response model, namely comprehension, retrieval from memory, judgement and response (Tourangeau et al., 2000). To better understand these issues, cognitive interviews have been introduced into survey design methodology as a tool for improving the quality of questionnaires, for example by providing a qualitative assessment of the performance and equivalence of survey questions in different contexts and across socio-cultural (or linguistic) groups (Miller et al., 2014).

The paper begins with an overview of the quantitative and qualitative methods used to assess data quality in the EWSD. This is followed by results from the quantitative analysis, focusing on respondent engagement and fatigue. Results from the qualitative assessment are then given, based on cognitive interviews with Czech respondents. Finally, a mixed methods approach is used to demonstrate how findings from the qualitative study informed changes to the EWSD questionnaire in its second round, focusing on frequency of use.

Methods

This paper studies data quality in two rounds (2016 and 2017-2018) of the EWSD, which collected data in 16 countries on purchases and patterns of use of the most commonly consumed illicit drugs. The EWSD methodology has been described elsewhere (Matias, 2022; Matias et al., 2019). As such, only the aspects of the questionnaire design and analytical approach relevant to our examination of EWSD data quality are discussed here, beginning with the quantitative analysis.

Quantitative sample and analysis

The quantitative analysis used a combined dataset for the two EWSD rounds (2016 and 2017-2018) and included all 16 countries participating in the survey. Respondent fatigue was examined across drug modules as the number of drug modules that respondents agreed to participate in stratified by the number of drugs consumed in the last year, that is, by the number of modules they were eligible to complete, and within

drug modules (as item non-response organised by order of appearance in the module).

As response rates and drop-outs are also important indicators of data quality, it is useful to consider how the EWSD dealt with these issues. To prevent drop-outs after having started the survey, respondents were given control over the questionnaire length: drug-specific modules were displayed in random order and only to those respondents who reported having used the respective drug in the past 12 months. At the beginning of each drug module, respondents could choose whether they wanted to answer more questions about a particular drug or to skip the entire module. As respondent fatigue was expected to increase with the number of modules they were eligible to answer, and as the use of some drugs (e.g. cannabis) would be more common than others, the EWSD randomised the order in which the drug modules were presented to respondents in order to try to balance the sample sizes for individual drug modules.

The impact of the item format was examined for questions on frequency of use in the last year and assessed by comparing the exact number of days the drug was consumed, with responses provided in categorical format (daily, almost daily, etc.).

Qualitative sample and analysis

The qualitative assessment of data quality was conducted within a broader project that aimed to assess the validity and reliability of the Czech version of the questionnaire used in the first round of the EWSD (see the box 'Assessing reliability and validity in the EWSD'). Participants were recruited and interviewed in 2016 via chain referral techniques through drug services and the researcher's informal networks to represent a broad spectrum of people who use drugs.

Only participants aged 18 or older who had used at least one of the main substances covered in the first round of the EWSD (cannabis, ecstasy/MDMA, cocaine, amphetamines) in the last 12 months were interviewed. The sample included people whose patterns of drug use varied from occasional to regular and frequent use, including the use of different drugs. In total, 19 individuals aged between 19 and 45 years were interviewed, of whom 13 were males and 6 were females. About half of the sample lived in unstable conditions and would qualify as having high-risk patterns of use (EMCDDA, 2020). All 19 participants used more than one substance: herbal cannabis was the most frequently used (18), followed by amphetamines (16), cannabis resin (14), ecstasy/MDMA (10) and cocaine (8). Six respondents had participated in and completed the EWSD questionnaire online prior to the interview.

The interviews were structured as a combination of a think-aloud approach and a probing interview (Willis, 2005), with the

aim of understanding all the difficulties, misunderstandings and obstacles that participants experienced in providing the most accurate answers to the EWSD questions. During the interview, participants went through the questions and answer options on their own hard copy of the questionnaire and the responses and comments were noted by the interviewer.

All 19 interviewees completed all relevant drug modules (based on drugs they reported to have used in the last 12 months). Each difficulty or hesitant response prompted a contextual discussion about the habits and practices of the participant that may have affected their ability to provide a valid and reliable answer. After the questionnaire was completed, respondents were asked to elaborate on some aspects of their drug using habits. Subsequently, the items that appeared difficult or otherwise problematic were annotated and the narrative parts of the interviews were summarised to guide the development of the EWSD questionnaire for the second round. The main findings stemming from these cognitive interviews, and those relating to data quality specifically, are examined in turn in this paper.

Results from the quantitative analysis

Respondent engagement and fatigue

More than 84 000 individuals opened the EWSD link and accessed the questionnaire across the two rounds in 2016 and 2017-2018. Before proceeding to its content, respondents were asked to provide informed consent to participate in the

survey. The percentage of those who did so and who agreed to participate was similar in both rounds, with an average of around 70 %. Between countries, this figure ranged from 64 % (in Luxembourg) to 83 % (in Italy). Among those who agreed to participate in the survey, 48 % of respondents in the first round and 52 % in the second round were eligible for analysis: they were living in the country where they completed the questionnaire, were at least 18 years old and had consumed at least one of the drugs included in the EWSD in the last year. The percentage of eligible respondents also varied by country, ranging from 33 % in Estonia to 70 % in Italy. Around 17 % of those who agreed to participate were not eligible because they had not answered any of the drug questions. This ranged from 7 % of respondents in Italy to 24 % in Lithuania.

Module response, expressed as the percentage of those who had agreed to participate in each drug module among all eligible for it, is presented in Table 1, both overall and by the number of drugs used in the last year. Across both rounds, the highest response rate was registered for the cannabis module (78 %), followed by the MDMA (67 %) and cocaine (63 %) modules. Response rates were higher in the first EWSD round than in the second, the latter of which contained more drug modules, which potentially competed for respondents' time and attention. For instance, 76 % of eligible respondents agreed to fill in the cannabis module in the second round, whereas 84 % did so in the first round. These percentages rates varied between countries, ranging from 80 % in the Netherlands to 93 % in Czechia in the first round and from 70 % in Poland to 85 % in Cyprus in the second round.

Further, module response decreased with the number of modules to which respondents were potentially exposed. For

TABLE 1

Percentage of respondents who agreed to participate in the respective drug module among all those eligible to complete the module, overall and stratified by number of drugs used in the last year, EWSD 2016 and 2017-2018

	Number of drugs used in the last year						Overall
	1	2	3	4	5	6	
EWSD 2016							
Cannabis (n = 10 367)	88.7	84.9	81.9	74.4	–	–	84.1
MDMA (n = 5 120)	92.7	85.0	81.0	73.4	–	–	79.2
Cocaine (n = 3 773)	87.0	81.8	78.0	69.3	–	–	74.2
Amphetamines (n = 3 765)	75.0	77.1	70.7	67.1	–	–	69.5
EWSD 2017-2018							
Cannabis (n = 29 193)	85.2	79.4	71.9	65.9	59.0	47.4	75.9
MDMA (n = 13 656)	81.7	75.2	67.0	61.9	55.1	40.7	62.9
Cocaine (n = 9 846)	78.6	72.9	65.0	59.8	52.1	38.3	59.2
Amphetamine (n = 10 698)	78.0	69.0	62.3	56.9	50.9	39.4	52.3
Methamphetamine (n = 3 244)	50.0	50.5	44.7	37.7	35.1	30.0	35.7
New psychoactive substance (n = 8 398)	49.6	33.5	37.8	38.3	33.1	22.0	34.1

those who had consumed only cannabis, roughly 89 % and 85 % of respondents agreed to complete the module in the first and second rounds, respectively. However, this percentage was as low as 69 % of respondents who were eligible to answer all four drug modules in the first round and 47 % of respondents who were eligible to answer all six drug modules in the second round. This pattern was similar for almost all drug modules in both rounds (see Table 1).

Item non-response

Respondent fatigue or disengagement was also assumed to be shown through higher item non-response rates as they progressed through individual modules. Using the cocaine module from the second round as an example, the percentage of missing values generally increased as respondents advanced through the module (Table 2). The first cocaine module question had only 0.5 % missing or invalid answers in the second round, while some questions appearing towards the end of the module were answered by about half of those who started the module. For instance, the question about

Assessing reliability and validity in the EWSD

The test-retest reliability of the core questionnaire items was tested by Škařupová et al. (2019) using the Czech part of the EWSD. About seven weeks after completing the questionnaire, a subset of Czech EWSD respondents were asked to answer the core question items again. Test-retest reliability was rated as moderate to good for most prevalence items. Items assessing frequency of use that asked for an exact number of days were typically more reliable than categorical items that assume a regular pattern of use (e.g. ranging from daily to less than once a month). Moreover, categorical items were interpreted differently by the respondents. The results of this study demonstrate that simplicity and unambiguity increase the reliability of responses and that tools measuring drug consumption need to consider the irregularity of drug use patterns.

TABLE 2

Distribution of missing and invalid answers in the cocaine module, EWSD 2017-2018

	Answered	Missing answer	Invalid answer	Missing or invalid (%)
During the past 12 months, how often have you used cocaine powder?	5 800	30	N/A	0.5
On how many days have you used cocaine powder during the last 30 days?	5 674	156	N/A	2.7
On how many days have you used cocaine powder during the past 12 months?	5 624	206	N/A	3.5
How much do you use on a typical day you use cocaine powder (in grams)?	4 781	629	420	18.0
How do you usually consume cocaine powder?	5 691	139		2.4
How do you usually get cocaine powder?	5 662	168	N/A	2.9
When you buy cocaine powder, how many grams do you usually buy?	*3 062*	*125*	*50*	*5.4*
How much does it usually cost (the whole amount of grams you buy at one time)?	*3 066*	*132*	*39*	*5.3*
[You give away...] After making your purchase, how much do you usually give away or sell to others? (percentage)	*1 864*	*1 373*	*N/A*	*42.4*
[You sell to others...] After making your purchase, how much do you usually give away or sell to others? (percentage)	*1 766*	*1 471*	*N/A*	*45.4*
[You share with others...] After making your purchase, how much do you usually give away or sell to others? (percentage)	*2 647*	*590*	*N/A*	*18.2*
Please think back to the last 30 days. How often have you bought cocaine powder during this period?	*3 166*	*71*	*N/A*	*2.2*
How much money have you spent on cocaine powder during the last 30 days?	*1 479*	*1 683*	*75*	*54.3*

For rows in italics, the denominator is only those who reported that they had used *and* bought the drug in the past 12 months.
N/A, not applicable: these questions required respondents to select from pre-defined categorical answers; as such, no invalid inputs were possible.

money spent on cocaine during the last 30 days was skipped by more than half of respondents. We note here that some of these questions were placed at the end of the module for their complexity and that the difficulties in answering them may have contributed to the proportion of missing answers. Invalid answers provided in these items would support the notion of the difficulty of these questions, indicating that simple item non-response may not be a good indicator of fatigue.

Results from the qualitative study

Recall of prevalence and patterns of use

The cognitive interviews undertaken as part of the qualitative study in Czechia provided a number of insights into how respondents understood questions about their drug use and how well they were able to answer them. Inaccurate answers compromise data quality and the interviews aimed to identify issues in the questionnaire that may contribute to that. The qualitative study found that respondents' ability to provide accurate answers to some questionnaire items, such as simple prevalence questions, varied by frequency and intensity of use.

Exceptional experiences were generally easier to recall. For instance, respondents who regularly injected methamphetamine were able to recall precisely occasions when they had used ecstasy/MDMA, either because of its distinctive effects or because of the unusual context in which they had used the drug. Similarly, recall was better when it related to substances that were less available, more expensive and had higher perceived status, such as cocaine or cannabis resin. Special life events, such as birthday celebrations, weddings, holidays and major festivals, were useful for placing sporadic use in the correct time frame. When such special memories were not available, events near the edges of the last 30 days and the last 12-month period caused some hesitation and required more effort to recall. Moreover, some respondents changed their answer to the prevalence question after they were invited to recall the most recent event at which they had used the respective substance: distant events were originally reported as more recent and vice versa. This unintentional effect is known as telescoping (Tourangeau et al., 2000) and affects survey responses regardless of the mode of data collection.

Respondents may also alter their responses intentionally for various reasons. In our study, one participant reported frequent heavy use of cocaine in the past 12 months, while in fact his use of the drug was rather rare during this period. In the interview, he explained that he referred to experiences that had occurred four years prior to the survey in another country

and that he 'did not want to cause a bias'. Such inaccurate reporting may occur when respondents are overly motivated to help and when they might misinterpret the purpose of the study. This is especially relevant when surveying self-selected online samples of people who use drugs, as their motivations to participate may influence their answers. Therefore, the purpose of the survey needs to be made as clear as possible to mitigate this problem.

In addition, survey respondents may tend to answer in line with how they see themselves (Tourangeau et al., 2000). This may impact on respondents' accuracy in reporting their drug use. In our study, participants recruited among the clients of drop-in centres appeared to systematically under-report their drug use, actively searching for periods of no use (for instance due to illness) and claiming that they were 'trying to quit' or 'taking a break'. While interviewees with problematic patterns of use tended to underestimate their consumption, occasional users and those using in recreational settings tended to over-report their use of some substances. Whether or not an individual self-identifies as a user of a specific drug and tends to conform to that identity may explain this behaviour (Schwarz and Bless, 1992; Vinopal, 2008). Use of cannabis appeared to be an exception to this rule, which might be because of its general acceptability and the absence of significantly positive or negative connotations among the interviewed sample.

The interviews also showed that drug use, even when habitual, frequent or due to addiction, often lacks the regularity implicit in the way survey questions are framed and formulated. In particular, when respondents were asked about their usual behaviours or had to select one category that would represent their drug use over longer stretches of time, a straightforward answer often proved challenging. For instance, the question 'When you buy [a drug], how many grams do you usually buy?' assumes that purchases follow a regular pattern, while they may be subject to variation according to the availability or quality of the drug, financial resources, or the reputation of a certain source or dealer, among other factors. Price fluctuations may also impact the quantity purchased, with a number of participants admitting to occasionally making an exceptionally large purchase when they had the opportunity. Questions on usual behaviours often prompted a relative answer ('It depends...') and resulted in a range rather than a single value. Moreover, when participants were asked to recall the last occasion of purchase or use of a substance, the value often differed from what they reported as the 'usual' behaviour. The EWSD questionnaire endeavoured to account for such fluctuations by including items on cumulative behaviours, such as the total amount purchased and money spent over a defined period of time. However, in less detailed surveys, questions about averages, ranges or the last instance may be considered in order to collect more accurate responses.

Difficulties with estimating quantities consumed

For estimates of quantities consumed, respondents to
the EWSD were first asked which of a range of routes and
modes of administration they had used in the past year and
then about the quantity of drugs they usually used for those
methods of administration. For instance, when respondents
reported smoking as a route of cannabis administration, they
were subsequently asked whether they smoke it in a joint, dry
pipe or water pipe, and how much of the drug they usually
put into each of these devices. Participants in the cognitive
interviews highlighted that alternative consumption methods
not originally covered by the questionnaire, such as vaping,
are also increasingly popular. This underscores the value of
regularly testing such questions to ensure that they are kept up
to date with current drug use practices.

The cognitive interviews demonstrated that drug use may be
irregular even for daily users as it is affected by a range of
contextual factors. For example, the interviewees highlighted
some of the factors that may affect the quantities of cannabis
used. As participants described, the sizes and numbers of
joints and pipes are not consistent and can differ at different
times of the day, between weekends and weekdays, and for
solo use or use in a group. The quantity of cannabis or resin
used in a pipe or joint depends on the shape of the pipe, the
quality and potency of the drug, the number of people it will be
shared with, and the person who prepares it. The quantity of
the drug that is available also plays a role: some participants
described consuming larger doses of cannabis resin just after
a purchase and using less towards the end of the stock. By
including separate questions in the EWSD about different
ways of using drugs, the survey sought to facilitate a correct
assessment of the quantities consumed. However, some
participants felt that this was overly complicated and stated
that providing a weight in grams consumed over a period
of time would be easier and more straightforward for them
to answer.

Use of visual aids to assist the assessment of cannabis quantities

To help improve the accuracy of the information obtained
on the quantities of herbal cannabis and cannabis resin
consumed and purchased, the EWSD questionnaire provided
visual aids to guide respondents. Four photographs were
shown with different quantities of cannabis (either resin or
herb, as appropriate) next to a ruler and a credit card for scale,
and respondents could choose which best represented the
amount they used (Figure 1).

Overall, the photographic clues were considered helpful,
although some respondents first stated the quantity and
then searched for the picture that represented that quantity.
It should also be noted that the aspect ratio of the images
changed when displayed on different devices: when printed for
the interview, the ruler and card appeared smaller than their
real size and seemed to confuse some participants. Those
who reported the use of cannabis resin often said that the
pictures are not illustrative of all types of the drug appearing
on the drug market. Types of resin that are difficult to crumble
(referred to as 'plastic' or home-made resin) were specifically
mentioned. Moreover, the largest quantity in the photographs
for the questions on amount of purchased herbal cannabis and
resin was too small for some of the participants.

Taking local context into account

Drug using and purchasing patterns vary considerably between
countries, and such local differences may have an impact on
whether the content of the questionnaire will reflect the local
situation.

Interviews with Czech participants shed light on some
common and unique aspects of the local context, an
understanding of which helped to improve the EWSD

FIGURE 1
Example of visual clues for herbal cannabis
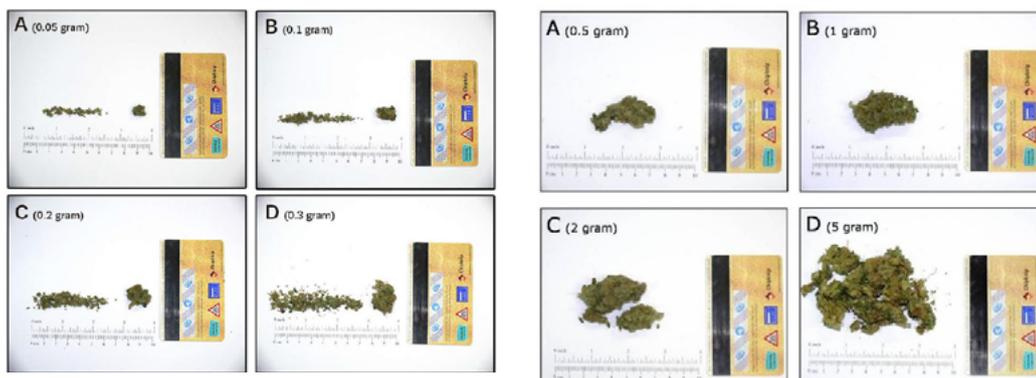
Source: Van Laar et al., 2013. Images reproduced with permission of the authors.

questions and their wording. For example, some participants stressed that the quantity of cannabis they produce or buy is not always the amount they consume, depending on the proportion of stems and leaves with a low cannabinoid content (referred to as 'garbage'). The larger the proportion of low-THC material, the greater the discrepancy between the amount of cannabis used and the quantity produced or purchased. Others, however, reported using whole plants by processing the less potent parts into cannabis butter or ointment.

Beyond cannabis, the interviews highlighted some unique aspects of methamphetamine use and supply in Czechia. Methamphetamine is historically the most common substance used by people with high-risk patterns of use in the country and is overall perceived to be more available than amphetamine, which is generally limited to nightlife settings (Mravčík et al., 2020). Participants who had used both of these substances highlighted how the drugs differ in their potency, intensity and duration of intoxication, and also in their after-effects and withdrawal symptoms. Compared with amphetamine, the effect of methamphetamine is more intense, lasts longer, results in more severe exhaustion and is perceived as a less controllable and predictable substance. In the first EWSD round, the questionnaire referred jointly to 'amphetamines', while only one item asked whether they had used amphetamine, methamphetamine or both in the last 12 months. This was perceived as problematic by respondents who used both drugs as it was not clear in relation to which substance they should be providing answers. Based on these interviews, the amphetamines module in the second round of the EWSD was separated into two distinct modules, one for amphetamine and one for methamphetamine.

Additionally, interview participants said that the traditionally home-based production of methamphetamine structures the drug scene into networks clustered around individual producers who are often also users. The drug is commonly traded for other services or goods such as precursors (mostly pharmaceuticals containing pseudoephedrine). These arrangements affect selling and purchasing practices and related slang, which makes it impossible to capture the complexity of methamphetamine markets in the one-size-fits-all questionnaire. It is not known how widespread the practice of exchanging drugs for other goods is in the European context, but it is likely that it is limited to groups of people with problematic patterns of use, who represent a hard-to-reach minority in web surveys. Although no changes were made in the questionnaire based on this finding, it is important to consider it when interpreting the data.

A mixed methods approach: using qualitative findings to inform survey design and quantitative analysis of results

Impact of item format on assessments of frequency of use

The qualitative interviews shed light on several issues that helped inform the item format of questions in the second round of the EWSD. Among the main changes were the inclusion of alternative methods of administration for cannabis, creating separate modules for amphetamine and methamphetamine, adapting the answer options in frequency questions to accommodate users experimenting with drugs, and changing the answer format in the numerical frequency of use items from an exact number of days to a range of days.

Some of these changes to the second round of the EWSD are analysed further here, namely those affecting the items measuring frequency of use. In the first round of the survey, questions on frequency of use were asked in different ways. First, respondents were presented with a categorical question in which they were asked to select a verbal description of their frequency of use ([1]). They were then asked to state the exact number of days they had used each drug through a free-text numerical response. When asked to answer the categorical question during the cognitive interviews, respondents struggled to select just one of the categories. Those who had only experimented with the drug in question lacked the option of 'once or twice', which in their experience suited them better than 'less than once a month'.

Even regular users reported periods of temporarily switching to another drug, abstinence attempts and times when the substance was not available or when they were in treatment. For many of the participants, their substance use followed seasonal fluctuations, with peaks during summer and lows in winter. The question on the number of days of use in the past 12 months allowed for such irregularity in drug use described by the respondents. The numerical frequency questions were also found to be more reliable in the test-retest reliability element of the study (Škařupová et al., 2019). However, calculating the number of days of use in the past year, when use has been irregular, may pose other cognitive challenges. This leads to a tendency to provide a quick, rounded answer (e.g. 5, 10, 20, 50 days), which could score highly on test-retest reliability while not necessarily being more valid (Škařupová et al., 2019).

---

[1] From 'daily' to 'less than once a month'. See Table 3 for details of categories in the first and second rounds of the survey.

Peaks in use were often related to binges occurring over short periods of time. Three scenarios appeared repeatedly in participants' narratives. First, frequent or daily users of herbal cannabis stated that they had bought or received cannabis resin only a few times a year, as it was not easily available in Czechia, and that they smoked it daily for a week or two until their supply was finished. Second, a typical pattern of heavy methamphetamine use, including injecting use, was described by several respondents. This involved binges lasting from two to four or more days followed by several days of exhaustion and sleep. These respondents often failed to state the number of days they used methamphetamine as they would use more than once on some days of the binge while being 'only' intoxicated on some days in between due to the lasting effects of the drug. Third, festival seasons and holidays are characterised by nightlife drug use (typically of ecstasy/MDMA or cocaine) during a week or two once or twice a year. These respondents would fall into the 'less than once a month' category for the majority of the year, although the number of days of use in the past 12 months would qualify them as more frequent users (use more than 12 days in a year), resulting in a data discrepancy that nevertheless correctly represents reality.

Results of quantitative analysis on frequency of use

Based on these findings from the qualitative interviews, the categorical ('daily', 'almost daily', etc.) and numerical use frequency questions were altered in the second round of the EWSD. The categorical item in the second round included the new option 'once or twice', while the numerical question provided ranges of days (1-5, 6-10, 11-20 days, etc.), instead of asking for an exact numerical input. To quantitatively examine how question format impacts on the results of frequency of use, we compared the answers to these categorical and numerical frequency questions in the first and second rounds of the survey. Table 3 uses the example of cocaine to illustrate the degree of consistency of answers to these two questions.

In the second round, the largest group of users (44 %) reported that they had consumed cocaine once or twice, considerably more than those who reported consuming it less than once per month (27 %). In sum, the total proportion of respondents who reported cocaine use less than once per month in the second round was 70 %. This was slightly higher than the 65 % recorded in the first round, where only one category represented a low frequency of use (once per month or less).

In Table 3, the numerical answers on frequency of use from the first round of the EWSD have been recoded to match the ranges provided for this question in the second round of the survey. In the table, the numerical responses are compared with the categorical ones. However, the options provided in the

categorical question do not exactly match the ranges provided in the numerical question. For example, the category 'less than once a month' could correspond to anything between 0 and 11 days, which overlaps with three numerical ranges: 1-5, 6-10, and, importantly, 11-20 days. As such, in Table 3, the responses are placed in one of three categories depending on the accordance between the numerical and categorical questions: those that are definitely discordant, those that may be both correct and discordant and those that are correct and not discordant.

Another issue that presents itself is that for the higher-frequency categorical response option, a decision has to be taken on the corresponding numerical responses that are deemed valid. For example, for the categories 'almost daily' and 'not daily but more than once a week', it is not entirely clear where the corresponding correct numerical ranges begin and end. The corresponding numerical ranges deemed likely to be correct can be seen in Table 3, although some caution should be exercised in interpreting these results.

In both EWSD rounds, about 80 % (78 % and 79 % for the first and second rounds, respectively) of those completing the cocaine module selected matching responses for the categorical and numerical questions on frequency of use. The agreement between these responses was highest among infrequent cocaine users. For instance, in the second round, 97 % of those who reported using cocaine once or twice in the past 12 months answered the numerical range question by stating that they had used the drug on 1 to 5 days in the same period.

Similarly, 92 % and 83 %, respectively, of those who reported using cocaine less than once per month in the first and second rounds of the EWSD also reported doing so on 1-5 or 6-10 days in the past 12 months in the numerical range question. The agreement between the numerical and categorical questions was lower for more frequent users. For example, agreement between the category 'less than once a week, but at least once a month' and the corresponding numerical range of 11-80 days per year was 58 % in both rounds. This means that 42 % of the numerical responses lie outside the range of the categorical response. Among those who reported using the drug once a week or more often, around half reported a numerical range outside the 'plausible' range for this category in both rounds. Disagreement between the numerical and categorical answers seems generally to relate to respondents providing lower numerical answers corresponding to the category of frequency initially selected (the numerical responses were generally skewed towards smaller values).

TABLE 3

Consistency of answers to categorical and numerical questions on frequency of use of cocaine in the last 12 months, EWSD 2016 and 2017-2018

		Daily	Almost daily	Not daily but more than once a week	Once a week	Less than once a week, but at least once a month	Less than once a month	Once or twice	Total
EWSD 2016	1-5 days	2	1	10	6	85	1 411		**1 515**
	6-10 days	0	3	6	19	125	270		**423**
	11-20 days	0	2	14	19	163	105		**303**
	21-50 days	1	4	42	72	140	30		**289**
	51-80 days	3	3	16	21	19	1		**63**
	81-100 days	0	8	32	8	14	1		**63**
	101-150 days	0	7	17	6	5	3		**38**
	151-200 days	2	8	18	2	1	1		**32**
	201-250 days	2	8	5	0	0	1		**16**
	251-300 days	4	5	5	0	0	1		**15**
	301-350 days	1	6	0	0	0	0		**7**
	More than 350 days	24	5	0	0	0	0		**29**
	Total	**39**	**60**	**165**	**153**	**552**	**1 824**	**0**	**2 793**
EWSD 2017-2018	1-5 days	3	3	17	48	192	790	2 374	**3 427**
	6-10 days	1	4	18	39	175	447	48	**732**
	11-20 days	1	6	29	50	317	182	15	**600**
	21-50 days	3	10	53	102	205	50	2	**425**
	51-80 days	1	6	54	50	41	4	2	**158**
	81-100 days	1	8	39	18	12	8	1	**87**
	101-150 days	1	9	29	8	3	2	3	**55**
	151-200 days	6	14	17	2	3	2	1	**45**
	201-250 days	4	9	12	5	2	1	0	**33**
	251-300 days	3	9	4	1	0	1	0	**18**
	301-350 days	9	3	1	0	0	1	2	**16**
	More than 350 days	9	2	1	0	0	1	4	**17**
	Total	**42**	**83**	**274**	**323**	**950**	**1 489**	**2 452**	**5 613**

Note: Responses that are definitely discordant between the numerical and categorical questions are shaded red, answers that may be both correct and discordant are shaded yellow, while those that are correct are left unshaded. Category 'once or twice' was only introduced in the EWSD 2017-2018.

Discussion

This paper has shown how both quantitative and qualitative assessments of data quality in web surveys can be undertaken to improve their design and format. These methods can be repeated for other web surveys to help improve data quality.

As the quantitative analysis showed, the number of people who had agreed to participate in each drug module as a percentage of all those eligible for the module was the highest for cannabis, followed by MDMA and cocaine. Module response was overall higher in the first round than in the second, probably because there were more modules available in the second round, which may have competed for respondents' time and attention. Another indicator of respondent fatigue was shown in the significant decrease in module responses with the number of modules that respondents were exposed to (based on the number of substances they reported having consumed in the past 12 months). The fact that the cannabis module had the highest completion rate across both survey rounds may also be linked to the fact that compared with other drugs, cannabis had the largest proportion of respondents who did not report using any other drugs covered in the EWSD and who were thus eligible to complete only one module. Respondent fatigue also manifested as a higher item non-response as respondents passed through individual modules,

with a large proportion of missing responses for questions located towards the end of the cocaine module, for example. Factors that may have influenced missing responses also include the complexity of the questions located towards the end of the module, in addition to the order in which the modules were displayed to participants.

As shown by the qualitative analysis, insights into how members of the target population read and respond to a questionnaire can help make sense of data that have already been collected through a survey and further improve the data collection tools. In particular, when researching stigmatised populations such as some groups of people who use drugs, qualitative data can provide a contextual and multifaceted understanding of their behaviours, habits and living circumstances (Sifaneck and Neaigus, 2001), which may be difficult to capture through a questionnaire. These behaviours and habits, including shared language and concepts, are not constant and tend to adapt to contextual changes. New terms may appear, old words shift meaning and new substances are referred to in different ways as labelling practices change. As such, subtle alterations in the question wording may have a large impact on survey responses. For instance, reported prevalence of MDMA use increased sharply after the term 'molly' was added to the text of the ecstasy/MDMA questions in a nationally representative school survey in the United States (Johnston et al., 2017).

The qualitative interviews with Czech respondents highlighted several important issues in the context of improving participation in the EWSD and the data quality of subsequent survey rounds. These included the inclusion of the 'once or twice' category in relation to frequency of use in the second survey round. As shown by the quantitative analysis, the inclusion of this category appeared to better capture the experience of a large proportion of respondents. This could be seen in the high level of consistency among those reporting having used cocaine 'once or twice' and the numerical answers they provided.

In relation to frequency of use, another important change to the EWSD questionnaire was the addition of numerical ranges, rather than exact numbers of days, when asking about frequency of use in the last year. This seems to have reduced the number of respondents providing invalid or implausible answers in the second round of the survey. The quantitative analysis for the frequency of cocaine use appears to show that using a range of days improved the agreement between categorical and numerical frequency of use questions. Overall, in both EWSD rounds, the cross-tabulation between the two 'frequency of last year use' variables (numerical and categorical) showed higher agreement among the less frequent users, while answers by those who had used more were often less consistent. An important caveat is that the

numerical ranges corresponding to the categorical responses were not always clear-cut at the higher frequencies of use (e.g. in relation to the categories 'almost daily' and 'not daily but more than once a week').

While using a range of days in combination with categorical questions on frequency of use seems to provide more accurate answers, the purpose of the data collection exercise has to be considered and weighed against this choice. For example, an exact number of days might be needed for some research questions, such as when estimating the annual amounts of a drug used (EMCDDA, 2019), in which case an exact number of days may provide better information to support such analysis. Nevertheless, a range of days can still support the development of rough estimates on annual consumption.

Conclusion

This paper describes both quantitative and qualitative assessments of data quality in the EWSD. It shows the importance of such assessments for improving data quality in surveys. Further, it underscores the value of conducting data quality assessments on an ongoing basis to ensure that the questionnaire continues to reflect current drug use practices, which are fluid and may change rapidly.

The quantitative analysis revealed some discordance between respondents' answers to survey questions based on their wording and response choices. Meanwhile, the qualitative analysis demonstrated how cognitive interpretations of questions probably drive these observed differences. For example, what may seem to be respondent fatigue may be due to the idiosyncrasies of how individuals interpret particular questions, such as with regard to quantities of drugs used or frequency of use.

Combining qualitative and quantitative assessments of data quality can help validate the findings of both. It also allows for more reliable interpretation of the survey findings and for a better understanding of its limitations. Overall, these exercises provide considerable insight into the validity of the EWSD, while also informing future web survey data collection on drug use and related behaviours.

References

Dillman, D. A. and Bowker, D. (2001), 'The web questionnaire challenge to survey methodologists', in Reips, U. D. and Bosnjak, M. (editors), *Dimensions of Internet Science*, Pabst Science Publishers, Lengerich, pp. 159-178.

EMCDDA (2019), *Estimating the size of the main illicit retail drug markets in Europe: an update,* Technical report, Publications Office of the European Union, Luxembourg (https://www.emcdda.europa.eu/publications/technical-reports/european-drug-markets-size-estimate_en).

EMCDDA (2020), *Statistical Bulletin 2020: methods and definitions for problem drug use*, https://www.emcdda.europa.eu/data/stats2020/methods/pdu.

Johnston, L. D., O'Malley, P. M., Miech, R. A., Bachman, J. G. and Schulenberg, J. E. (2017), *Monitoring the future national survey results on drug use, 1975-2016: overview, key findings on adolescent drug use,* National Institute on Drug Abuse, US Department of Health and Human Services, National Institutes of Health.

Matias, J. (2022), 'Development and implementation of the European Web Survey on Drugs', in *Monitoring drug use in the digital age: studies in web surveys*, EMCDDA Insights (https://www.emcdda.europa.eu/publications/insights/european-web-survey-drugs-overview-project_en).

Matias, J., Kalamara, E., Mathis, F., Skarupova, K., Noor, A. and Singleton, N. (2019), 'The use of multi-national web surveys for comparative analysis: lessons from the European Web Survey on Drugs', *International Journal of Drug Policy* 73, pp. 235-244, doi:10.1016/j.drugpo.2019.03.014.

Miller, K., Chepp, V., Willson, S. and Padilla, J. L. (editors). (2014), *Cognitive interviewing methodology*, John Wiley and Sons, New York.

Mravčík, V., Chomynová, P., Grohmannová, K., Janíková, B., Černíková, T. and Fidesová, H. (2020), *Drug situation in the Czech Republic in 2018*, Office of the Government of the Czech Republic, Prague.

Schwarz, N. and Bless, H. (1992), 'Constructing reality and its alternatives: assimilation and contrast effects in social judgment', in Martin, L. L. and Tesser, A. (editors), *The construction of social judgment*, Erlbaum, Hillsdale, NJ, pp. 217–245.

Sidi, F., Shariat Panahy, P. H., Affendey, L. S., Jabar, M. A., Ibrahim, H. and Mustapha, A. (2012), 'Data quality: a survey of data quality dimensions', *2012 International Conference on Information Retrieval & Knowledge Management*, pp. 300-304, doi:10.1109/InfRKM.2012.6204995.

Sifaneck, S. J. and Neaigus, A. (2001), 'The ethnographic accessing, sampling and screening of hidden populations: heroin sniffers in New York City', *Addiction Research & Theory* 9(6), pp. 519-543.

Škařupová, K., Singleton, N., Matias, J. and Mravčík, V. (2019), 'Surveying drug consumption: assessing reliability and validity of the European Web Survey on Drugs questionnaire', *International Journal of Drug Policy* 73, pp. 228-234.

Tourangeau, R., Rips, L. J. and Rasinski, K. (2000), *The psychology of survey response*, Cambridge University Press, Cambridge.

Van Laar, M., Frijns, T., Trautmann, F. and Lombi, L. (2013), 'Sizing the cannabis market: a demand-side and user specific approach in seven European countries', *Current Drug Abuse Reviews* 6, pp. 152-64, doi:10.2174/1874473706666131205152835.

Vinopal, J. (2008), *Kognitivní přístupy v metodologii výzkumných šetření: metoda okamžité validizace*, Sociologický ústav AV ČR, Praha, v.v.i.

Willis, G. B. (2005), *Analysis of the cognitive* interview *in questionnaire design,* Oxford University Press, New York.

About the EMCDDA

The European Monitoring Centre for Drugs and Drug Addiction (EMCDDA) is the central source and confirmed authority on drug-related issues in Europe. For over 25 years, it has been collecting, analysing and disseminating scientifically sound information on drugs and drug addiction and their consequences, providing its audiences with an evidence-based picture of the drug phenomenon at European level. Based in Lisbon, the EMCDDA is one of the decentralised agencies of the European Union.

About this series

EMCDDA Insights are topic-based reports that bring together current research and study findings on a particular issue in the drugs field. This paper is published as part of *Monitoring Drug Use in the Digital Age: Studies in Web Surveys*, an EMCDDA Insights that provides an overview of current knowledge and the latest developments in the field of web surveys on drug topics. The Insights contains in-depth reports on the methodology of web surveys, the available studies being carried out in different drug topics and analyses of the European Web Survey on Drugs. The Insights will be of interest to researchers and scientists, people who use drugs, policymakers and their advisors, specialists and practitioners, and all those concerned with the issue of drugs and innovative methods.

EMCDDA project group: João Matias, Alexander Soderholm, Katerina Skarupova, André Noor and Jane Mounteney

Recommended citation: Kalamara, E. and Škařupová, K. (2022), 'Quantitative and qualitative assessments of data quality within the European Web Survey on Drugs', Monitoring drug use in the digital age: studies in web surveys, EMCDDA Insights (https://www.emcdda.europa.eu/publications/insights/quantitative-qualitative-assessments-data-quality-european-web-survey-drugs_en).

This publication is only available in electronic format.

Publications Office
of the European Union