Data collection with and about people with diverse SOGIESC in social protection and other humanitarian and development programming: Recommendations and considerations from a review of literature

Billy Tusker Haworth, Edge Effect, for the 'We See You' project

---

1. Introduction

> "*Cultural values shape and are shaped by the ways that data is imagined and put to use; embodied histories of power and oppression underlie assumptions about what data is"…*
> *We can see data as "a tool of control or a potential instrument of empowerment,"*
> (Ruberg and Ruelos, 2020, p.1-2).

The invisibility of people with diverse sexual orientations, gender identities and expressions, and sex characteristics (SOGIESC; aka lesbian, gay, bisexual, transgender, queer, intersex, asexual and other gender and sexual minority identities, or LGBTQIA+ people[1]) in data collection has direct and significant consequences for addressing their rights, needs and strengths in Social Protection and in broader development and humanitarian aid programs. When people with diverse SOGIESC are invisible in regularly collected social and economic data, in research on inequality or in program specific needs assessments, it is easy for them to be overlooked in the design of interventions. Not only do people with diverse SOGIESC miss out on support, but existing inequalities can also be reinforced.

While social and political invisibility perpetuates inequalities, invisibility can also provide privacy and safety from threats, and data collection and reporting must be cognisant of doing no harm when 'shining a light' (Daigle, 2021). Edenborg (2019) describes the "double edged sword" of visibility: while visibility 'demystifies' queer people and their issues, "it simultaneously creates an environment for states and other actors to stigmatise and discriminate against people based on sexual orientation and gender identity" (p.353), with important examples given from Tanzania, Uganda and Egypt. It doesn't mean that visibility is inherently bad, but we need to be cautious that we don't "problematise the notion of visibility as inherently beneficial and pay attention to its contradictory and differentiated consequences" (Edenborg, 2019, p.353). This points to the importance having methods, principles, policies and guidance for data collection and use that are safe, ethical and fit-for-purpose for the people and contexts being researched.

Development and humanitarian actors, including those implementing Social Protection programs funded by donors such as the Department of Foreign Affairs and Trade (DFAT), are unlikely to make progress on

[1] In this paper, "LGBTQIA+" and "people with diverse SOGIESC" are used interchangeably to refer collectively to all gender and sexual minority identities. However, it is recognized that there are many local and cultural concepts of sexual and gender diversity in the world, and the use of the term LGBTQIA+ and SOGIESC here are used without prejudice to the ways that people and groups identify themselves in their own lives and contexts.

diverse SOGIESC inclusion in the absence of improved research and data collection on the lived experience of people with diverse SOGIESC. The are several reasons for the current state-of-affairs in which people with diverse SOGIESC are not included in research, data and analysis, including:

- Absence of data about people with diverse SOGIESC in existing datasets and reports, contributing to a) a lack of awareness of issues faced by people with diverse SOGIESC, b) absence of people with diverse SOGIESC in research that relies upon secondary sources and c) challenges establishing baselines.
- Lack of diverse SOGIESC inclusion in thematic policy and practice guidance that sets expectations or provides processes for data collection in Social Protection and other development and humanitarian programs.
- Lack of adaptation of research tools (across ethics, data collection and data management) to make them fit-for-purpose for working with people with diverse SOGIESC.
- Limited use of analytical frameworks that support use of data about people with diverse SOGIESC to improve diverse SOGIESC inclusion in Social Protection and other development and humanitarian programs.

The purpose of this paper is to present the results of a literature review into current best practices for collecting and using data on LGBTQIA+ people from a range of related fields, including social protection, humanitarian aid, international development, and public health. The review was conducted to meet the objectives of a) determining general best practices that can be adopted for safe and effective LGBTQIA+ inclusive data collection, and b) identifying knowledge gaps and/or critical considerations that should be investigated further and addressed to ensure LGBTQIA+ data collection activities are fit-for-purpose in social protection and humanitarian programs in the regions and contexts that DFAT works. Section 2 of this paper briefly presents the literature review methods, section 3 presents the literature review findings summarised under key themes and issues, and section 4 presents critical considerations for application and translation of LGBTQIA+ data collection best practices into DFAT/social protection/humanitarian contexts.

2. Literature review methods

Material for this literature review was sourced through a multi-pronged approach taking into account both peer-reviewed academic literature and grey literature (e.g., policy material, NGO reports). The search and filtering methods are summarised in figure 1. Once a paper was found and deemed relevant for inclusion in the review it was added to an Excel Spreadsheet. Papers were then read in full and summarised into an annotated bibliography. This was used to identify key cross-cutting themes that reflected the scope of best practice guidance and issues/gaps expressed in the literature. Many of the results were from a public health/clinical setting (e.g., HIV data collection), but the results also represented other fields, with papers from humanitarian and disaster studies, for example. A majority of existing guidance on LGBTQIA+ data collection was from International or United States based organizations (and some Australian). An important limitation of the literature review methods is that only items published in English language could be included, owing to the researcher/author's limited proficiency for other/multiple languages. Some of these methodological considerations will be discussed further in reflecting on the findings in the sections below.

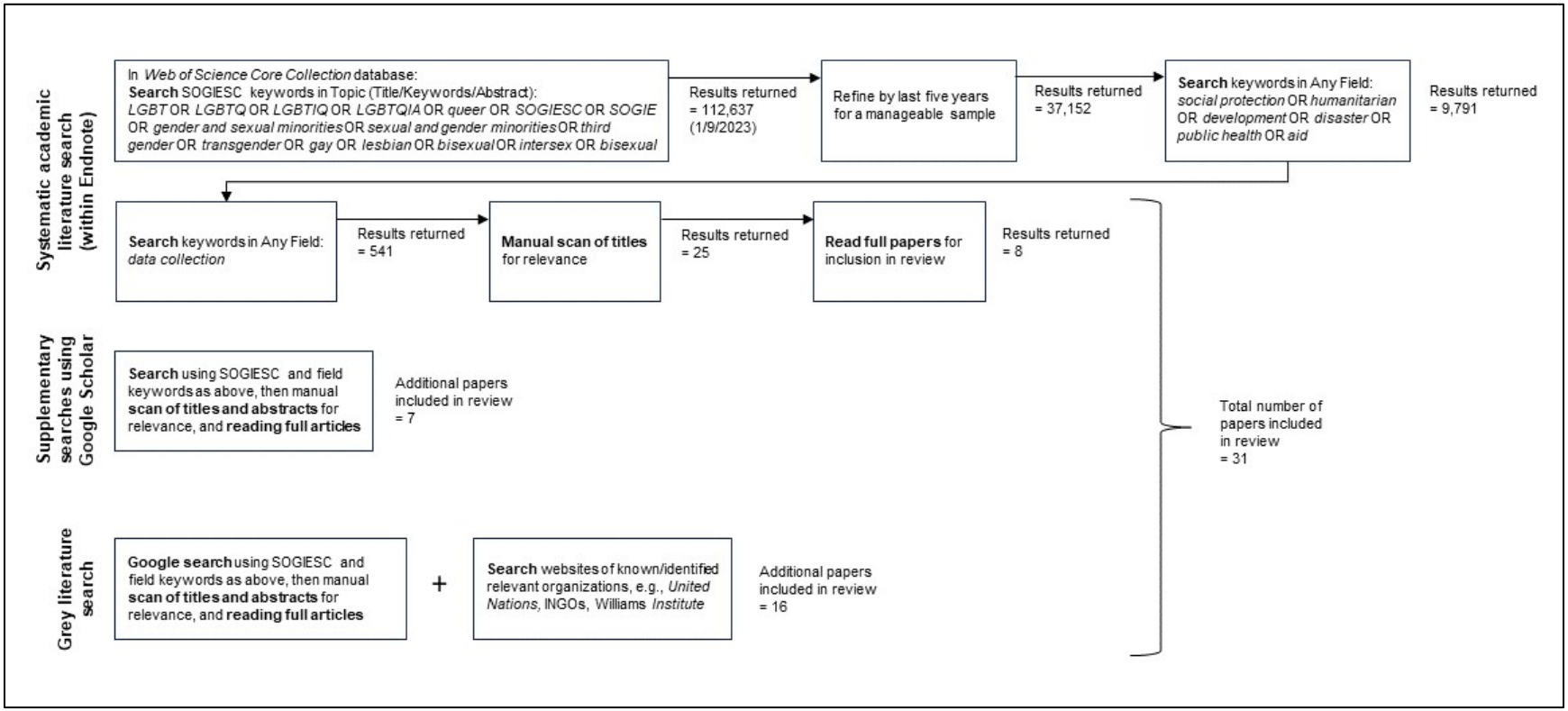

Figure 1. Literature search methods. In total 31 papers were sourced and included in this review (15 peer reviewed academic articles and 16 items from grey literature). Note: the *Web of Science Core Collection* "is the world's leading citation database" with "ten indexes containing information gathered from thousands of scholarly journals, books, book series, and conferences" (Clarivate, 2023).

3. Findings

3.1 Conduct research *with* diverse SOGIESC people

A key emphasis across the academic and grey literature on engaging with and collecting data from/about people with diverse SOGIESC was on the importance of working directly with diverse SOGIESC people themselves (Roth *et al.,* 2021; Tordoff *et al.,* 2022; del Río-González *et al.,* 2021; Haworth *et al.*, 2022). Gender and sexual minority communities should always be engaged in development of research and data collection about them, with appropriate means of consultation and recognition (Brown and Herman, 2020; Tordoff *et al.*, 2022), such as by beginning with listening to local communities and involving people with diverse SOGIESC as part of research teams and advisory boards (del Río-González *et al.,* 2021; Lanham *et al.*, 2019). This includes incorporating input at multiple stages of research design and implementation and centring participant experience in activities like survey design (Moseson *et al.*, 2020; Tordoff *et al.,* 2022). Taking a locally led, community-based or participatory and capacity-sharing approach is recommended (Tordoff *et al.,* 2022; del Río-González *et al.,* 2021; Roth *et al.,* 2021; van der Merwe *et al.*, 2023).

Multiple positive outcomes arise from approaches involving people with diverse SOGIESC, including: designing and conducting research that is more accurate, applicable, and contextually meaningful to the populations being studied (Tordoff *et al.*, 2022); increasing the likelihood of maintaining 'do no harm' principles through data collection and research (Mazurana *et al.*, 2023); potentially increasing research participation (see section 3.3 below); and supporting work that is not extractive but empowering and transformative (Dwyer, 2021). Haworth and others (2022) argue that increased reflection on research ethics broadly is needed, considering questions like who is conducting studies, how the voices of LGBTQIA+ people can be elevated, and how appropriate it is for more powerful groups to conduct research about people less powerful, suggesting that one approach to help minimise these issues is active involvement of LGBTQIA+ people in research processes, not just as research subjects.

An instructive example is a project in El Salvador, Trinidad and Tobago, Barbados and Haiti, where transwomen were involved in a study to inform policies and health programs for transwomen experiencing gender-based violence (Lanham *et al.,* 2019). The women were trained to interview other transwomen to facilitate more in-depth understandings of how the challenges faced and reported can be considered and addressed in development projects. Study design and interview guides were also informed by extant gender-based violence research, advisory groups, and pilot testing. The transwomen involved also participated in participant recruitment and data interpretation, helping to shape the research and results from their own perspectives, positionalities, and lived experiences (Lanham *et al.,* 2019).

This kind of engagement requires time, dedicated investment and resourcing, and prior research. Roth and colleagues (2021) consulted with a range of key informants on engaging with LGBTQIA+ people for humanitarian and development research and concluded that "international actors must invest the time and resources necessary to understand the context before any engagement with LGBTQI communities even begins. This understanding should include context analyses of national- and community-level dynamics, social norms, and attitudes; historical perspectives including the influence of colonialism on current laws and policies; the current political reality; and the linkage of these contextual factors to

conflict trajectories," (p.5). While actors should leverage existing LGBTQIA+ organisations and community groups in research processes (Mazurana *et al.,* 2023), they must remain attentive to the ways existing organisations and networks can both benefit and exclude individuals and identity groups within LGBTQIA+ (e.g., based on race, ability, class, ability, and other intersecting factors) (Haworth *et al.,* 2022).

3.2 Importance of *partnerships*

Linked to working with people with diverse SOGIESC is the importance of partnerships. Reaching and consulting with LGBTQIA+ populations ethically, appropriately, safely, and effectively while "doing no harm" requires working in partnership with local LGBTQIA+ organisations where possible (Daigle, 2021; Mazurana *et al.*, 2023). Relationships are particularly useful and important with community leaders and gatekeepers (van der Merwe *et al.,* 2023). Trust is an important factor for the success of many research methods and participant recruitment/involvement, and therefore community outsiders may have difficulty working with populations like LGBTQIA+ without forming and demonstrating genuine relationships and partnerships with relevant local people and organisations (Smith *et al.,* 2017; Wolfe *et al.,* 2023). To avoid perpetuating unintended harm with local communities, global actors need to foster long-term and non-extractive partnerships (Roth *et al.,* 2021). To achieve such partnerships, it is recommended that actors conduct "power analyses within the contexts they work to understand existing power dynamics as they relate to the financial, convening, and decision-making power that international agencies hold over local actors, policies, and practices. In addition, power analyses should engage local LGBTQI serving organizations to examine local dynamics of power and agency within the LGBTQI community within a specific context," (Roth *et al.,* 2021, p.4).

3.3 Consider *facilitators* of and *barriers* to participation

Best practice literature for engaging people with diverse SOGIESC in research and data collection activities highlights the importance of considering factors that either help to facilitate or contribute to barriers to participation. Table 1 presents a summary of the key topics in this area. In designing and implementing research and data collection the goal should be to enhance facilitators while mitigating barriers wherever possible. It won't always be possible to mitigate all barriers, especially given the diversity of people and experiences within diverse SOGIESC populations. LGBTQIA+ people should not only be engaged as research subjects or in implementation of recruitment and research activities, but they are also best placed to inform on what activities and practices will be most appropriate and effective for their own communities in terms of facilitating research and minimising barriers to participation. "International humanitarian organizations... must recognise and value the expertise that LGBTQI people have on their own realities, and therefore elevate their lived experiences of risk to their safety and wellbeing, as well as the barriers and facilitators to access within each context." (Roth *et al.,* 2021, p.4)

Table 1. *Factors important for facilitating or presenting barriers to participation of people with diverse SOGIESC in research and data collection activities.*

Facilitators	Barriers
Establishing confidentiality and privacy – during data collection and in storage and use of data (Harkness *et al.,* 2022; Barclay and Russell, 2017)	Privacy, anonymity, confidentiality and other ethical concerns, e.g., fear of being "outed" and potential consequences (e.g., violence, immigration, financial...), stigma (Brown and Herman, 2020; Wolfe *et al.,* 2023; Harkness *et al.,* 2022) - Individuals who have not (or not fully) disclosed their identities will be especially difficult to reach for research participation if there is perceived risk of disclosing their identities (even to researchers) (Bell, 2017)
Altruism and sense of helping their own community as motivation (Wolfe *et al.,* 2023; Harkness *et al.*, 2022)	Increasing visibility of diverse SOGIESC also aids actors wanting to stigmatise and discriminate, e.g., in Tanzania, Uganda, Egypt (Edenborg, 2019; Daigle, 2021)
Recruitment using tailored messaging and leveraging peers/networks to enhance trust (Wolfe *et al.,* 2023; Lanham *et al.*, 2019; Harkness *et al.,* 2022)	Distrust of researchers, government, institutions… (Wolfe *et al.,* 2023; Haworth *et al.,* 2022; Family Safety Victoria, 2019; Brown and Herman, 2020) - justifiable suspicion of "research" because historically labelled sexual minorities as "sick or deviant" (Bettinger, 2010)
Linked to the previous two, and for hard-to-reach populations (e.g., LGBTQIA+ in Nepal) snowballing is critical (Koehler and Menzies, 2023)	Don't relate to the need for study/intervention and therefore don't anticipate any benefits (Harkness *et al.,* 2022)
Accessibility of study locations or methods (e.g., online) (Wolfe *et al.,* 2023; Harkness *et al.,* 2022)	Inaccessibility or other requirements, e.g., transport, internet/smartphone (Wolfe *et al.,* 2023; Harkness *et al.,* 2022; Craig *et al.,* 2020; Smith *et al.*, 2017)
Participant experience: feeling welcome, respected, confident and trusting of investigators, questions resonate with	Feeling unseen or misunderstood (e.g., in an interview or survey) – including researchers not recognising heterogeneity, e.g.,

participants' lived experiences - in turn aids quality and completeness of collected data (Moseson *et al.,* 2020; Harkness *et al.,* 2022; Smith *et al*., 2017; Craig *et al.,* 2020)	assuming the same cultural references for engagement are relatable to all LGBTQIA+ people (Harkness *et al.,* 2022)

3.4 *Qualitative* methods/data alongside quantitative

Often quantitative data and measures are the default to collect, as they can be *seen* to be representative or populations/groups and are useful (preferred?) for informing policy development and targeted interventions. However, Brown and Herman (2020) describe a tension between measures that are inclusive of diverse identities and limiting identities to allow for quantitative analyses. For example, some would err against capturing data that is overly nuanced and disaggregated (e.g., describing a population as "LGBTQIA+" rather than reporting data for lesbians, gay people, bisexual people etc, or collecting transgender men, transgender women, non-binary, gender diverse, and all other non-cisgender gender identities as a single "other" (to cisgender men or women)) as this reduces the sample sizes required for certain quantitative or statistical analyses.

Utilising qualitative research methods and data offers alternative and/or complementary approaches with a range of benefits. Quantitative data and data standardisation, particularly in clinical/health contexts, capture details like participant identities, but fundamentally miss details about the unequal contexts shaping the lived experiences of those identities and individuals, resulting in an inability to understand or address inequity despite possible expectations of change accompanying the collection of identity data (Cruz and Paine, 2021). In this way, qualitative social science approaches are strong and essential for understanding and addressing the complex and changing dynamics and contextual factors associated with social problems (Cruz and Paine, 2021; Haworth *et al*., 2022).

Specific techniques provide particular advantages. For instance, interviews can foster feelings of catharsis for the participant as they share their story (Craig *et al.,* 2020; Haworth *et al.,* 2023). Methods like photo elicitation privilege the voices and intersectional experiences of participants, such as sexual and gender minority youth in Craig and colleagues 2020 study. The participant's photos and stories become the centre of the interview, rather than the interviewer's questions, helping to redefine researcher-participant power dynamics (Craig *et al.,* 2020) and reduce power, class, and knowledge differentials (Smith *et al*., 2017). Other benefits of photo-elicitation methods, especially for disempowered or disenfranchised groups, include increased trust and rapport, enhancing participants' memory recall and informing more accurate data collection (Smith *et al*., 2017). There are important limitations to such methods as well, and some related to barriers to participation, such as the sometimes-time-consuming nature of methods like interviews (conducting, transcribing/processing, and analysing), or needing to travel or to possess technologies (e.g., camera or smartphone) to participate.

3.5 Guidance for *inclusive* and *effective* surveys

Acknowledging the benefits of qualitative methods outlined in the previous section, quantitative methods such as surveys continue to have value for diverse SOGIESC research and data collection. The literature

describes a range of practices to make surveys more inclusive and, relatedly, effective. Of particular importance is to recognise that while some broad "best practices" are summarised here, context is always important, and methods and language that are suitable for one person or group in one context may be inappropriate for a different person or group or in a different context. Methods and tools should always be careful reviewed for use in their applicable context before deployment.

> "*Concepts, definitions, and language used to talk about sex, gender, and transgender bodies and experiences are cultural, historical, and have varied over place and time, and are likely to continue to change*," (Tordoff *et al.,* 2022, p.6).

3.5.1 Guidance for capturing gender(diversity)

A range of guidance exists for collecting accurate data on gender in a safe way. Firstly, it is advised that all parties seeking to collect data make informed assessments about what information is necessary to collect for particular purposes/services (Barclay and Russell, 2017). It may not be necessary at all for one's analysis and reporting to ask particular questions that may be sensitive to a participant, such as asking a transgender or non-binary person about their sex assigned at birth (Ruberg and Ruelos, 2020; Family Safety Victoria, 2019), and thus such questions could be removed from a survey. Asking gender or sex questions when not necessary can have negative implications for both participants and the research/data collection, including by offending participants, deterring respondents, misrepresenting transgender people, and/or implying sex is more important than current gender (especially if surveying in non-clinical settings) (Medina and Mahowald, 2022).

If collecting data on gender (diversity) is deemed necessary, some consensus is present in the literature for adopting multi-step or two-step methods in surveys (del Río-González *et al.,* 2021; Tordoff *et al.,* 2022; Park, 2016; Brown and Herman, 2020; Bauer *et al*., 2017; Family Safe Victoria, 2019; Medina and Mahowald, 2022). The recommendation is to ask multiple gender self-ID questions, such as "What is your current gender? 1. Male 2. Female 3. Non-binary, genderqueer, gender-nonconforming, or agender 4. Two-spirit or other culturally specific terms 5. Another gender [free-text option] 6. Decline to answer," and a separate question like "What sex were you assigned at birth, meaning on your original birth certificate? 1. Male 2. Female 3. Decline to answer" (Medina and Mahowald, 2022). It is noted again, as above, that asking about sex assigned at birth can still be sensitive (i.e., asking trans and non-binary folks; Ruberg and Ruelos, 2020), and that while two-step methods are often touted as the least invasive and most likely to capture transgender identities, this only applies in Western settings where strong binary distinctions exist between male/female, man/woman, and sexual orientation/gender identity – so, one must consider what terms are appropriate in other places (Koehler and Menzies, 2023). Medina and Mahowald (2022) suggestion another approach is to add a transgender status question, such as "What is your current gender? ..." and then "Do you consider yourself to be transgender? 1. Yes 2. No 3. I don't know 4. Prefer not to respond...".

Other best practices reported include asking gender questions over time at all data collection points for longitudinal studies (del Río-González *et al.,* 2021), reporting disaggregated data (i.e., do not report trans and non-binary people as a single group, and disaggregate for trans men, trans women, non-binary, etc... regardless of sample size; these groups have distinct and diverse experiences, issues, and barriers to report; Tordoff *et al.,* 2022), and avoid the use of "other" as an alternative for diverse gender options

(Family Safety Victoria, 2019). Guyan (2022) describes examples of four types of people that might describe themselves as "other", therefore rendering the question ineffective: 1. those who don't identity as LGBTQIA+ but don't like terms like "hetero" or "straight", 2. those who reject traditional gender or sexual orientation terms, 3. transgender respondents, 4. those experimenting or questioning, so one cannot assume what "other" means when applied to all LGBTQIA+ people.

3.5.2 Capturing intersex variations

For data collection about people with intersex variations, Family Safety Victoria (2019) inform that adding an "intersex" option to questions regarding sex or gender is not appropriate as most intersex people identify with a binary sex of male or female. Instead, they recommend including a separate question to gender/sex specifically on intersex. In addition to more accurate data collection, this approach avoids misgendering people with intersex variations and inadvertently including people who confuse inte rsex for a gender identity. They also recommend including a short definition of intersex to before or with related questions (Family Safety Victoria, 2019).

Medina and Mahowald (2022, p.29) provide the following recommendation for an intersex status survey question: "Some people are born with chromosomes, hormones, internal sex anatomy, or genitalia that do not fit within typical notions of "male" or 'female" sex. This is known as intersex. Many learn they are intersex from a medical diagnosis. This is different from being transgender. Are you intersex? 1. Yes 2. No 3. I'm not sure 4. I don't understand the question."

3.5.3 Capturing sexual orientation

Sexual orientation is distinct from and should never be conflated with gender. Park (2016, p.4) describes three dimensions of sexual orientation that can each be measured in a survey:

- Self-identification: how does one identify one's own sexual orientation
- Sexual behaviour: the sex of one's partners (i.e., same sex, different sex, or multiple sexes)
- Sexual attraction: the sex or gender of people that one feels attraction towards.

In Australia, people may not be used to being asked about sexual orientation in surveys (as government data collection mechanisms like Census have historically not, and still do not, ask such questions), but research from elsewhere suggests willingness of individuals to disclose sexual orientation, for example over 90% of people in the UK said they would be happy to answer sexual orientation questions in surveys (Family Safe Victoria, 2019). This point more generally, however, should be considered when designing surveys for different contexts, where it is a) more or less acceptable to have diverse sexual orientations, and consequently b) more or less acceptable to disclose or discuss sexual orientation. Different question wording variations, such as those above, may be more or less appropriate in different data collection settings.

For questions asking participants to self-identify their sexual orientation, Medina and Mahowald (2022, p.23) give the following recommended format, which aims to capture a full diversity of potential sexual orientations while not forcing respondents to put themselves into a category they do not align with (I.e., free text and opt-out options are included): "What is your sexual orientation? 1. Gay or lesbian 2. Straight 3. Bisexual 4. Pansexual 5. Queer 6. Asexual 7. Two-Spirit [or other culturally specific terms] 8. No specific label, but not straight 9. Something else [free-text option] 10. Prefer not to respond".

3.5.4 Using inclusive language and data collection processes

> "*Respecting how people identify themselves, and understanding their place in wider society, is key to pursuing locally owned, locally led and meaningfully inclusive humanitarian responses. Using appropriate language and cultural knowledge can be the basis for better ways of working. This can inform what needs are assessed and how, who gets counted or registered for assistance, and how that assistance is delivered,*" (Daigle, 2021).

Adopting inclusive language has the dual benefits of making people feel included when participating and contributing to increased data accuracy (Family Safety Victoria, 2019). The overall data population engagement and data collection approach should be conducted in respectful and inclusive ways. This means inclusive of people with diverse SOGIESC, first, and second, acknowledging and mitigating discrimination and marginalisation (Barclay and Russell, 2017). Barclay and Russell (2017) inform that terminology used should be empowering, consistent, and inclusive, researchers and data collection tools should be accommodating (e.g., of differences between details like sex on legal documents and person's actual/current gender), and processes should also be made inclusive, for instance by avoiding making assumptions about people based on appearance, voice, names and addresses, and by providing options for anonymous or pseudonymous data collection where possible.

We need to move beyond normative framings of gender and sexuality in research into people with diverse SOGIESC (Haworth *et al.,* 2022). For example, trans-inclusive language should be used that does not conflate sex assigned at birth with gender, such as not referring to people assigned female at birth as women but using *cisgender women* so as to not further the normative assumption that women implicitly means non-transgender (Tordoff *et al.,* 2022).

It is important to include culturally relevant and specific terms and translate SOGIESC terminology into local languages appropriately and accurately (Medina and Mahowald, 2022). Using culturally specific terminology and methods is important for increasing data collection effectiveness (Park, 2016). For example, using "Methi" and "Kothi" is more accurate than simply adding "Third gender" in the Nepal census. Brown and Herman (2020) describe how *travesti* in Brazil is culturally used and accepted as referring to people assigned male at birth with femme gender identities, but in Colombia the word is associated with sex work and is rejected by SOGIESC communities as it remains a slur that has not been reclaimed.

The fluidity of terminology and identities should also be considered. One should give definitions for concepts used in questions and responses, as terminology used by LGBTQIA+ populations evolves rapidly, paying attention to differences for different demographics and contexts (Medina and Mahowald, 2022). Moreover, sexual and gender identities, especially for people with diverse SOGIESC, are not static or fixed, but rather are sites of complex temporality and intersectional multiplicity (Ruberg and Ruelos (2020). Requiring participants to identify as a singular category of sexuality or gender identity along a continuum does not acknowledge people with fluid or multiple identities or identities present outside the binaries of heterosexual/homosexual and man/woman (Suen *et al.,* 2020). Each sexual orientation and gender

identity has a rich historical, cultural and political context (Suen *et al.,* 2020). . Further, individuals may change their responses about themselves over time as their gender or sexual identity evolves, and surveys need to accommodate this fluidity and create as many opportunities as possible to capture these dynamics and ensure data is reliable and usable (Medina and Mahowald, 2022). Some approaches to help facilitate the capturing of fluidity include allowing multiple answers with the prompt "select all that apply" in question stems, and including write-in options to acknowledge fluidity, complexity and individual's own understandings and preferred terminologies about themselves (Suen *et al.,* 2020)

3.6 Developing *organisational capacity*

> "*Humanitarian organizations must undertake ongoing learning and capacity-sharing processes to ensure that they have inclusive and supportive policies in place to protect and support LGBTQI people, including staff, partners, clients, and research subjects*," (Roth *et al.,* 2021, p.5).

The final key theme on best practices for research engagement and data collection with people with diverse SOGIESC identified through the literature review is the need to developing organisational capacity, as the above quote from Roth and colleagues highlights. This includes developing, maintaining, and implementing LGBTQIA+ inclusive organisational policies, strategies, staff training and monitoring/reporting measures across whole organisations, not just among diversity and inclusion departments or interested individuals, for instance. Mazurana *et al.* (2023) emphasises the need for adequate funding for additional data collection and management, including staff training, if diverse SOGIESC research is something new to an organisation (which is likely), and for continuous funding for upkeep of data repositories and publicly available dashboards, allowing for secondary analyses and peer learning. They suggest data findings should be brought back to affected people and local partners, and that organisations should document and publicly share the impacts of more inclusive data to ensure learning across agencies and fields. Where organisational capacity is new or lacking, organisations should work with key expert/specialist stakeholders to fill gaps and increase capacity (Mazurana *et al.,* 2023). Organisational capacity for LGBTQIA+ inclusive should also include learning and capacity for understanding indigenous cultural understandings of sexuality and colonial constructions, if relevant to the contexts of working (Picq, 2019), and intersectionality as a key theme for action research for LGBTQIA+ inclusion (Badgett and Crehan, 2017).

Capacity is also required in the research and data collection, analysis and research communication methods being employed, for instance to understand and mitigate safety concerns through security of data, researchers and participants (e.g., issues of hostile governments, or research in public spaces) (Brown and Herman, 2020). Research and data collection literacy will also ensure organisations and their staff understand and can critique or appraise different approaches, data, and findings in more useful ways, for example recognising the costs of data standardisation in erasing the richness and uniqueness of individual lives (Koehler and Menzies, 2023) or understanding participant sampling techniques to be able to develop strategies that account for intersectionality and diversity among diverse SOGIESC populations (Roth *et al.,* 2021).

4. Critical considerations

The above guidance drawn a wide range of extant literature has been and will continue to be useful, especially with the current dearth of quality LGBTQIA+ inclusive data collection practices currently being adopted in the sector. But there are important questions to consider and areas to build upon in relation to this guidance. This section lists critical considerations for developing more tailored guidance for effective and inclusive LGBTQIA+ data collection for the specific range of applicable social protection, humanitarian and other related contexts of interest to DFAT.

Further consideration is needed for which of the above guidance needs to change for different contexts and why. The limited guidance available on data collection involving people with diverse SOGIESC comes from high-income countries in the 'global north' where there are often (though not always) higher levels of social acceptance of people with diverse SOGIESC. Translation and applicability of best practices/methods in different contexts should not be assumed – we may have data collection best practices for Australia, for example, but what about the Solomon Islands, Ethiopia, and other vastly different geographical, cultural, and political contexts? What factors determine the applicability (or not) of data collection practices?

Moreover, many of the guidance in the literature come from health or clinical settings (especially best practices for collecting LGBQTIA+ information through surveys) – which guidance does or does not translate to social protection, humanitarian aid and other relevant contexts? Methods that are suitable in one area (e.g., general public health) may be inappropriate or dangerous in others (e.g., emergency humanitarian aid). There is a need for LGBTQIA+ research and data usage guidance that is relevant within the varied contexts in which development and humanitarian organizations work, that aligns with the way data is collected and used within development and humanitarian programs, that is consistent with sector commitments including local participation, accountability to affected people, safeguarding and ethics.

Some of the existing guidance in the literature points to the importance of intersectionality in LGBTQIA+ lives and consequently data collection practices. However, much of the guidance appeared generic and broadly for LGBTQIA+ as a single uniform population. Of course, people with diverse SOGIESC are not homogeneous, and therefore things like what constitutes 'inclusive language', for example, are not the same for all people with diverse SOGIESC, even within same country/contexts. For example, "queer" may be generally understand and accepted by LGBTQIA+ people in Australia, but subsections of populations may still find the term offensive and therefore inappropriate for use in data collection instruments such as surveys, for instance older LGBTQIA+ people who grew up with "queer" used as a vicious slur.

A key recommendation was to develop partnerships and work with existing LGBTQIA+ organisations for data collection activities. While the justifications for this are logical and important, one should also consider situations and contexts where there aren't organisations to partner with – how should one approach the research then? Where they are present, LGBTQIA+ organisations don't represent everyone in LGBTQIA+ populations, and who is being represented and who is being excluded or overlooked must be considered and perhaps alternative access/recruitment strategies devised for particular people/groups. Or, what if a partner organisation has a particular focus or mandate, for example due to funding/funders (e.g., focus on key populations, or people living with HIV)? What if there are suitable

organisations, but one cannot access them? Some guidance on navigating these potential challenges is needed.

Staff training, data management and ethics are presented in the literature as key areas that need attention for developing organisational capability to conduct safe, inclusive and effective LGBTQIA+ data collection. But what are the expectations and what should happen when using contractors to conduct pieces of work, as it often typical in the field of humanitarianism? Who is accountable for their actions, and should actions like LGBTQIA+ training for organisational staff be mandatorily extended to all contractors used? Or should contractors only be used if they can demonstrate a particular level of competence or awareness? What would any of this look like in countries or environments that remain hostile towards LGBTQIA+ people, but it is still necessary to utilise local contractors?

Finally, some overarching questions should be asked and critically reflected upon for all LGBTQIA+ data collection approaches:

1. With data collection approaches, particularly surveys, are we imposing ideas or categories that suit 'us' but don't' align with diverse lived experiences of people with diverse SOGIESC? What are the implications of this?
2. What other biases are informing our data collection practices and what we consider and value as 'best practices'? This literature review, for instance, is limited in that it uses and privileges only resources available in English language, reflecting the researcher/author's own positionality and limitations. So, what guidance exists and what practices are positioned as 'best' in literature published in other languages or from other positionalities?
3. Which methods best create space for LGBTQIA+ expertise and understanding to come through?
4. How should we consider, and then reduce, power imbalances? For example, power imbalances between government and the public, between dominant groups and marginalised people, or between researchers and research subjects. Are there approaches to data collection and research that better-place marginalised people at the centre of activities, facilitating sharing and exploring of ideas together?

References

Badgett, M.V.L., & Crehan, P. (2017). Developing actionable research priorities for LGBTI inclusion. *Journal of Research in Gender Studies*, 7(1), 218-247.

Barclay, A., & Russell, M. (2017). *A guide to LGBTIQ-inclusive data collection*. Canberra LGBTIQ Community Consortium. Available at https://genderrights.org.au/wp-content/uploads/2020/03/LGBTIQ-Inclusive-Data-Collection-a-Guide.pdf (last accessed 27 September 2023).

Bauer, G.R., Braimoh, J., Scheim, A.I., & Dharma, C. (2017). Transgender-inclusive measures of sex/gender for population surveys: Mixed-methods evaluation and recommendations. *PloS one*, 12(5), e0178043.

Bell, M. (2017). Data collection in relation to LGBTI People. *Analysis and comparative review of equality data collection practices in the Eruopean Union. Luxembourg, Publication Office of the European Union, 22*. Available at https://ec.europa.eu/newsroom/just/redirection/document/45605 (last accessed 30 October 2023).

Bettinger, T.V. (2010). Ethical and methodological complexities in research involving sexual minorities. *New Horizons in Adult Education and Human Resource Development,* 24(1), 43-58.

Brown, T.N.T., & Herman, J. (2020). Exploring International Priorities and Best Practices for the Collection of Data about Gender Minorities. A focus on South America. *The Williams Institute*. Pp. 1-30. Available at https://williamsinstitute.law.ucla.edu/wp-content/uploads/WPATH-English-Mar-2020.pdf (last accessed 18 October 2023).

Clarivate (2023). *Web of Science Core Collection Overview*. Clarivate Web of Science Help website. https://webofscience.help.clarivate.com/en-us/Content/wos-core-collection/wos-core-collection.htm#:~:text=Web%20of%20Science%20Core%20Collection%20is%20the%20world's%20leading%20citation,on%20your%20institutions%20subscription%20depth (last accessed 30 November 2023).

Craig, S. L., Eaton, A. D., Pascoe, R., Egag, E., McInroy, L. B., Fang, L., ... & Dentato, M. P. (2020). QueerVIEW: Protocol for a technology-mediated qualitative photo elicitation study with sexual and gender minority youth in Ontario, Canada. *JMIR research protocols*, 9(11), e20547.

Cruz, T. M., & Paine, E. A. (2021). Capturing patients, missing inequities: Data standardization on sexual orientation and gender identity across unequal clinical contexts. *Social Science & Medicine*, 285, 114295.

Daigle, M. (2021, January 1). *How Humanitarians should consider LGBT+ issues in their work.* Overseas Development Institute website. Available at https://odi.org/en/insights/how-humanitarians-should-consider-lgbt-issues-in-their-work/ (last accessed 30 October 2023).

del Río-González, A. M., Lameiras-Fernández, M., Modrakovic, D., Aguayo-Romero, R., Glickman, C., Bowleg, L., & Zea, M. C. (2021). Global scoping review of HIV prevention research with transgender people: transcending from trans-subsumed to trans-centred research. *Journal of the International AIDS Society*, 24(9), e25786.

Dwyer, E. (2021). *The Only Way Is Up: Monitoring and Encouraging Diverse SOGIESC Inclusion in the Humanitarian and DRR Sectors.* UN Women Regional Office for Asia and the Pacific, Bangkok. Pp.104. Available at https://www.edgeeffect.org/wp-content/uploads/2021/03/TheOnlyWayIsUp_Web.pdf (last accessed 30 October 2023).

Edenborg, E. (2019). Visibility in Global Queer Politics, in Bosia (Ed.) *The Oxford Handbook of Global LGBT and Sexual Diversity Politics.* Oxford University Press, 348-363 https://doi.org/10.1093/oxfordhb/9780190673741.001.0001

Family Safety Victoria (2019). *Victorian Family Violence Data Collection Framework: Data collection standards - Lesbian, gay, bisexual, transgender and intersex communities*. Victorian Government. Available at https://www.vic.gov.au/victorian-family-violence-data-collection-framework/data-collection-standards-lesbian-gay-bisexual (last accessed 27 September 2023).

Guyan, K. (2022). *Queer data: using gender, sex and sexuality data for action*. Bloomsbury Academic.

Harkness, A., Rogers, B. G., Mayo, D., Smith-Alvarez, R., Pachankis, J. E., & Safren, S. A. (2022). A Relational Framework for Engaging Latino Sexual Minority Men in Sexual and Behavioral Health Research. *Archives of sexual behavior*, 51(5), 2399-2412.

Haworth, B. T., Cassal, L. C. B., & de Paula Muniz, T. (2023). "No-one knows how to care for LGBT community like LGBT do": LGBTQIA+ experiences of COVID-19 in the UK and Brazil. *Disasters*, 47(3), 584-607.

Haworth, B. T., McKinnon, S., & Eriksen, C. (2022). Advancing disaster geographies: From marginalisation to inclusion of gender and sexual minorities. *Geography Compass*, 16(11), e12664.

Koehler, D., & Menzies, N. (2023). Surveys, big data, and experiments: how can we best learn about LGBTI development outcomes?. In *Handbook on Economics of Discrimination and Affirmative Action* (pp. 1-23). Singapore: Springer Nature Singapore.

Lanham, M., Ridgeway, K., Dayton, R., Castillo, B. M., Brennan, C., Davis, D. A., ... & Evens, E. (2019). "We're going to leave you for last, because of how you are": Transgender women's experiences of gender-based violence in healthcare, education, and police encounters in Latin America and the Caribbean. *Violence and Gender*, 6(1), 37-46.

Mazurana, D., Marshak, A., & Spears, K. (2023). *Sex, age (and more) still matter. Data collection, analysis, and use in humanitarian practice*. Care International. Pp. 70. Available at https://fic.tufts.edu/assets/Sex-age-and-more-still-matter_Final-report.pdf?cmdf=ocha+inclusion+sogi (last accessed 6 October 2023).

Medina, C., & Mahowald, L. (2022). *Collecting data about LGBTQI+ and other sexual and gender-diverse communities. Best practices and key considerations.* The Center for American Progress. Available at https://www.americanprogress.org/article/collecting-data-about-lgbtqi-and-other-sexual-and-gender-diverse-communities/ (last accessed 29 September 2023).

Moseson, H., Lunn, M. R., Katz, A., Fix, L., Durden, M., Stoeffler, A., ... & Obedin-Maliver, J. (2020). Development of an affirming and customizable electronic survey of sexual and reproductive health experiences for transgender and gender nonbinary people. *PLoS One*, 15(5), e0232154.

Park, A. (2016). Reachable: Data collection methods for sexual orientation and gender identity. *The Williams Institue*. Pp. 1-8. Available at https://williamsinstitute.law.ucla.edu/wp-content/uploads/SOGI-Data-Collection-Mar-2016.pdf (last accessed 17 October 2023).

Picq, M.L. (2019). Decolonising Indigenous Sexualities: Between Erasure and Resurgence, in Bosia (Ed.) *The Oxford Handbook of Global LGBT and Sexual Diversity Politics.* Oxford University Press, pp.168-184. https://doi.org/10.1093/oxfordhb/9780190673741.001.0001.

Roth, D., Blackwell, A., Canavera, M., & Falb, K. (2021). *When "we know nothing": Recommendations for ethical research and learning with and for LGBTQI people in humanitarian settings*. New York: International Rescue Committee. Available at https://www.rescue.org/report/when-we-know-nothing-recommendations-ethical-research-and-learning-and-lgbtqi-people (last accessed 30 October 2023).

Ruberg, B., & Ruelos, S. (2020). Data for queer lives: How LGBTQ gender and sexuality identities challenge norms of demographics. *Big Data & Society*, 7(1). https://doi.org/10.1177/2053951720933286.

Smith, J. A., Mccullough, R., Critchlow, C., & Luke, M. (2017). Proposing an initiative research methodology for LGBTQ+ youth: Photo-elicitation and thematic analysis. *Journal of LGBT Issues in Counseling*, 11(4), 271-284.

Suen, L.W., Lunn, M.R., Katie, K., Finn, S., Duncan, L., Jae, S., . . . Obedin-Maliver Juno. (2020). What sexual and gender minority people want researchers to know about sexual orientation and gender identity questions: A qualitative study. *Archives of Sexual Behavior*, 49(7), 2301-2318. https://doi.org/10.1007/s10508-020-01810-y

Tordoff, D. M., Minalga, B., Gross, B. B., Martin, A., Caracciolo, B., Barbee, L. A., ... & Khosropour, C. M. (2022). Erasure and health equity implications of using binary male/female categories in sexual health research and human immunodeficiency virus/sexually transmitted infection surveillance: recommendations for transgender-inclusive data collection and reporting. *Sexually Transmitted Diseases*, 49(2), e45-e49.

van der Merwe, L., Cloete, A., Savva, H., Skinner, D., November, G., & Fisher, Z. Z. (2023). Engaging transgender women in HIV research in South Africa. *BMC Public Health*, 23(1), 1-9.

Wolfe, H. L., Boyer, T. L., Rodriguez, K. L., Klima, G. J., Shipherd, J. C., Kauth, M. R., & Blosnich, J. R. (2023). Exploring research engagement and priorities of transgender and gender diverse veterans. *Military Medicine*, 188(5-6), e1224-e1231.