Surveying the LGBTQ population(s) through social media

Author:
Marco D. TERRIBILI

Supervisor:
Prof. Assunta VITERITTI

SAPIENZA
UNIVERSITÀ DI ROMA

Surveying the LGBTQ population(s) through social media

PhD in Applied Social Sciences

Marco D. TERRIBILI

Matr. 1799725

Supervisor

Prof.ssa Assunta VITERITTI

A.A. 2019-2020

Abstract

The majority of quantitative studies regarding Lesbian, Gay, Bisexual, Transgender and Queer people, carried out worldwide, point out several recurring themes. In Italy, the main characteristics of these studies have been the intensive use of mixed-methods research and the support of the LGBTQ associations towards these studies. In fact, for a long time associations represented the main informative source on the LGBTQ community, promoting and partnering with the most important quantitative surveys about LGBTQ. Of course, the information provided has been extremely useful and informative, but social knowledge should go beyond "associationism" and associations activities.

Today we can exploit innovative data sources, mainly those from social media, which can allow us to reach, investigate and study this population(s) bypassing associations, hitherto unavoidable.

In this context, a challenging research project, called "Over the rainbow", has been carried out through a survey, which has involved all those Instagram users, listed by a web-scraping software, who tag their pictures with some of the most common LGBTQ community hashtags.

Studying the application of big data methods on issues related to gender identity and sexual orientation ensures that this project grounds on two main theoretical frameworks, which are getting more and more intertwined: gender studies and digital sociology.

The survey investigated some of the thornier subjects in LGBTQ people's daily life, such as self-perception of gender identity and sexual orientation, coming-out experience, participation in LGBTQ Pride events, homotransphobia and discrimination experiences suffered.

The survey's results give both methodological suggestions, about the use of social media big data for studying the LGBTQ population(s), but also important informative contents, for instance, the massive use of dating apps, LGBTQ's relationship with civil rights associations, and the "sneaky" nature of gender and sexual discrimination.

In this research project, the methodology becomes as important as the studied field, and the attention paid to sequentially conduct, interviews before, and a survey after, led to an acquaintance with the characteristics of the LGBTQ population(s) which could be useful for planning and promoting efficient social policies for LGBTQ inclusion.

Keywords: LGBTQ, social media data, survey design, sexual orientation, gender identity, discrimination

Contents

List of tables, images and figures

INTRODUCTION

LGBTQ stands for Lesbian, Gay, Bisexual, Transgender and Queer people. Definitions are always difficult in this field and, accordingly, the acronym has kept changing over the past few years. The LGB acronym replaced the generic and misleading term "gay" in the mid-to-late 1980s, but after a few years it changed to LGBT, to include transgender people. It finally became LGBTQ, to also include queer people, who refuse traditional identities (sex, gender, sexual orientation, or even ethnicity), defined by the predominant culture of a society (Carey-Mahoney, 2016). Other acronyms such as LGBTQI, LGBTQIA, LGBTQIA+ are used too to be as inclusive as possible, for example with intersexual and asexual people.

Referring to LGBTQ as a population, thus overlooking the heterogeneity of the groups that the acronym identifies, can be strongly inappropriate. Thereafter we will refer to LGBTQ adding an "s", within brackets, following the word population: this *population(s)* will represent, with all its inner many-sidedness, our target population.

Worldwide, through the decades we have seen much progress in unravelling the social characteristics of this population(s). However, the quantitative investigation of LGBTQ social features is still in its infancy in Italy, as in several other countries. The reason for this lack of knowledge is inevitably linked to the intrinsic complexity of this population(s), the rarity of the involved phenomena, and the elusiveness of the people making up this community (Corbisiero et al., 2013; De Rosa and Inglese, 2018). These problematic aspects entail research difficulties, that result in complexities both at a methodological and epistemological level. This set of research problems are indissolubly linked both to the investigation of a hidden population (Monaco, 2018) and to the dichotomies male/female and heterosexual/homosexual, on which people have been categorized for decades, but that today do not fit anymore with our complex society (Ruspini, 2014; Mieli, 1977). Hyper-gendered categories such as "males" and "females," "brothers" and "sisters," and "husbands" and "wives" are not enough: social research needs to rethink the measurement of sex and gender as to not reproduce statistical representations. This rethinking process should aim both to reflect the diversity of gendered lives, and to better align survey

measurement practice with contemporary gender theory(Westbrook and Saperstein, 2015). If this process did not happen the social research would neglect much of the information, and likely limit the understanding of the human processes that perpetuate social inequality dynamics, such as non-inclusivity, pay gap, social disparities, ghettoisation and gender hegemony.

The lack of information about LGBTQ entails the persistence of false stereotypes about people, their community, and stigmatizing behaviour towards LGBTQ people. Moreover, the absence of reliable quantitative data impedes to quantify negative behaviours, such as homophobia (but also biphobia and transphobia), mobbing against LGBTQ workers or bullying against LGBTQ students, and putting in place social policies to incentivize inclusion and equality. The collection, analysis and availability of accurate data could be crucial for the implementation of successful social policies and people's services (Hanft, 1981). Social Science Data can support practitioners and policy-makers in their daily work, and become a driver for change, addressing a country, or a community, towards a progressive and liberal path of civil recognition of the LGBTQ instances.

For a long time associations represented the main informative source on the LGBTQ community (Barbagli and Colombo, 2007; Inghilleri and Ruspini, 2011), promoting and partnering the most important quantitative surveys about LGBTQ. Of course, the information provided has been extremely useful and informative, but social knowledge should go beyond *"associationism"* and associations activities: studying LGBTQ people from their associations' point of view can lead to deep but biased knowledge because the target population(s) can be quite wider than the people who belong to these associations.

Today we can exploit innovative data sources, mainly those from social media, which can allow us to reach, investigate and study this population(s) bypassing associations level, hitherto unavoidable (Matthews and Cramer, 2008).

In this context, complex and constantly evolving, the focus of this research project has been on the methodology, but without neglecting the issues about gender and sexuality. After a detailed and thorough literature review, some innovative methodological concerns have been deepened. Indeed, if using big data in social research is still a methodological topic under consideration (González-Bailón, 2013; Snee et al., 2016) using them to study the LGBTQ population(s) could be even more controversial.

In this regard, this dissertation tries to answer a critical, but specific, methodological research question: *"could big data be, coming from social media, a right tool for studying LGBTQ population(s)?"*

In answering this question, a challenging research project, called "*Over the rainbow*" (indicated as OTR, from now on) has been conducted. The research project aims to observe, from an impartial and discrete point of view, the LGBTQ world, represented by the colours of the rainbow, which reflects the characteristic diversity of this community.

The present research project has been carried out through a survey, which has involved all those Instagram users, listed by a web-scraping software (Schröder, 2018), who tag their pictures with some of the most common LGBTQ community hashtags. These keywords, used in the tagline, can describe the content of the picture, but also the person who posted the photo and sometimes reveal her/his/its sexual orientation.

Web-scraping represents a cutting-edge set of informatics techniques whose purpose is to extrapolate information publicly available on a web page. These techniques, not implemented for social research purposes, but for commercial ones, allowed us to get into a social medium, often considered as a "black box", and in which we tried to shed some light. Working on a popular social media, like Instagram, led to a remarkable number of people to contact, and to submit a questionnaire to, in a direct, cheap and careful way. On the other hand, having a social network as fieldwork, necessarily includes in our theoretical framework topics such as digital sociology (Lupton, 2015; 2013), digitalization (Mäkitalo, 2020), relationships between humans and web technology (Snee et al., 2016), and self-representation on social media (Rizzo, 2018). That is the reason why we emphasise the work carried out *on* big data, and not *with* big data. A subtle but substantial difference, because we take all the advantages of big data use (massive volume, high access velocity, real-time updating, etc.) without "collaborating" in such a way with them; for instance, we did not use indexing algorithms or sponsored contents to reach a higher number of users.

Studying the application of big data methods on issues related to gender identity and sexual orientation ensures that this project grounds on two main theoretical frameworks, which are getting more and more intertwined: gender studies and digital sociology. Gender studies, and Queer studies, helped us to answer a basic but crucial question that we asked ourselves when we started to design this research project: *what we talk about when we talk about LGBTQ?*

As will be seen in the following pages, describing gender identity and sexual orientation behaviours could be really challenging: the operationalisation of these concepts need, in a quantitative context, operative and purposeful definitions and classifications, which could be seen as oppressive superstructure to shun. Adopting Gender and Queer studies as the base theoretical framework, facilitate finding the right compromise between intricacies of the studied concepts and the feasibility of their investigation.

On the other hand, working in an online environment, we had to consider that the development of the Internet of things (IoT) favours the digitalization of social, economic, political and private aspects. IoT facilitates new ways of daily self-representation, in which bodies are not just objects with borders and propriety, but they structure themselves as material and discursive phenomena (Barad, 2003). The intra-actions[1] between technologies and individuals constitute a society in which the dichotomies male/female, public/private, heterosexual/homosexual don't work anymore (Ruspini, 2014). In this context, using a gender-sensitive approach means to analyze the identity, relationships and gender role transformations.

At the same time working on a social network necessarily includes topics such as Digitalization, and of all those relationships between humans and web technology, and particularly between social media users, social researcher and social media websites and apps. Including digitalization in our theoretical framework means taking into account what Noortje Marres, in 2016, wrote: "the rise of social media, the proliferation of mobile devices, and the uptake of digital analytics across professional practice, have given rise to a new apparatus for researching social life. And, by doing so, social methods are becoming even more prominent or mainstream in our societies and cultures". (Snee et al., 2016).

Gender studies and digital sociology, the two important and broad theoretical frameworks on which this project grounds, are deeply interconnected because technology changes our daily life. As a consequence of this, our role in contemporary society is changing too, both as an object of research and as researchers. Social media have become more than websites and app, but a full-fledged new social space, on which to express themselves, their gender, their sexuality. But, on the other hand, social media have become an innovative

[1] Karen Barad introduced the intra-action concept, in opposition to interaction, to signify the mutual constitution of subjects and objects (Barad, 2007).

empirical research tool, which seems to get more important for social researchers, statisticians, scientists and scholars.

This change entails several important considerations on the research method, which were not neglectable and so have been deeply explored, to not leave unfulfilled any theoretical, methodological and epistemological aspects.

The survey's goal is to detect and identify, through the questionnaire, the thornier subjects in LGBTQ people's daily life, such as self-perception of gender identity and sexual orientation, coming-out experience, participation in LGBTQ Pride events, homotransphobia and discrimination experiences suffered by them. These subjects have been identified to produce reliable data, which can be used by all those institutions, policymakers and organizations which work (or ought to work) to plan efficient social policies for LGBTQ equal rights, such as recognition of same-sex relationships, LGBTQ parenting or adoption, anti-discrimination and hate crime laws, legal recognition and accommodation of reassigned gender, and laws concerning access to sex reassignment surgery and hormone replacement therapy.

Anyway asking for this kind of sensitive information, properly, should be a scrupulous job. Several international experiences suggest drafting the questionnaire with a consultation process based on qualitative methodologies, such as focus group among researchers (Green, 2011; Pew Research Center, 2013). This process, although methodologically convincing, would not directly represent the personal experience of the LGBTQ people; it would actually propose survey schemes already used and inevitably linked to associations world. That is the reason why qualitative interviews on LGBTQ people are to be carried out before the questionnaire drafting phase. Semi-structured interviews allowed a part of the survey target population to actively contribute to writing questions, on which the survey is based. People belonging to different LGBTQ population(s) groups have been selected, trying to find different persons in terms of age, profession, social background, educational level and involvement level in any LGBTQ associations, and then they have been interviewed. Such interviews give the analysis an added value of the direct experience. In addition, they give the research project the methodological strength of the sequential use of qualitative and quantitative methods, in a pragmatic perspective of combining both methods. Interviews have addressed the questionnaire at every step of the drafting process: from the consultation of the beta-test, finally to the

interpretation of results; that is the reason why, in the following pages, the reader will also find some interview excerpts.

The present thesis is divided into different parts, subdivided into chapters and in turn into paragraphs, to facilitate the reading process.

In the first part, as well as framing the "definition matter" and the recurring epistemological and methodological considerations (chapter 1), the most important surveys about LGBTQ, will be presented and reviewed (chapters 2-3-4). This extensive review aims to understand and frame the the-state-of-the-art of social research on this field, and how it has kept changing over the past years. Quantitative studies have undertaken very different research paths, worldwide, in Europe and our country; that is the reason why these studies have been reviewed in three different chapters. Around the world, in the last few decades, many attempts have been made to quantify the LGBTs, and/or their proportion on the whole population. These attempts have been carried out in different ways in the various countries, taking into consideration the traditions, the contexts and the degree of social recognition of the LGBTQ in each of them. In Europe, rather, the estimation of the LGBT population size, or its proportion on the total, has increasingly lost its importance, except for private opinion polls and market research which are still strongly interested in quantifying LGBTQs, economically seen as an unexplored market niche (IFOP, 2017; Lam, 2016). An Italian typical characteristic of LGBTQ studies has been instead, over the past twenty years, the use of mixed-methods (Johnson et al., 2016; Creswell and Plano Clark, 2011). This approach has guaranteed results, not only by numeric data but also by experiences and observational qualitative data, collected through interviews, focus groups, story-telling, etc. which has given a strong added value to the national studies about LGBTQ.

In the second part, the methodology applied in the OTR research project will be described, with special attention to the use of Instagram data for social research and the questionnaire drafting phase. In the fifth chapter all the methodological and theoretical implications of using social media data, coming from Instagram, for investigating the LGBTQ population(s), have been faced and addressed. Special attention was given to the specific choice of Instagram as the social media used, and to the set of hashtags, to focus on. Moreover, the web-scraping program for "breaking into the social media black box" has been deeply illustrated. In the sixth chapter, we tried to answer the question "What to ask,

and how to ask it?", presenting the sequential pragmatic approach (Rossmann and Wilson, 1985) used for drafting the questionnaire: a set of interviews propaedeutic to the survey. Qualitative and quantitative methodologies combined, in a complementary rather than competitive way, to ensure that the questionnaire contains all the right questions, asked in the right way.

Finally, in the third part, the survey's results will be presented and theoretically analyzed, regarding the five main dimensions investigated by the questionnaire: self-perception, coming-out, apps and digital technology, socialization and discrimination. Firstly, we will describe respondents, according to their socio-demographic characteristics pattern, such as age, occupational status, highest educational degree, geographical area of residence and origin, civil and/or relationship status. We will try to point out respondents homogeneous profiles, also regarding how they frame their sexual orientation and gender identity. We will reason on how the participants feel (or do not) part of a community, of associations, and the Pride movement. Finally, we will investigate discriminatory dynamics correlated to sexual orientation and gender, paying attention to who, where (at work, at school, online, etc.), and how, these homotransphobic acts are committed.

We will try to analyze investigated phenomena from an intersectional point of view, which cross different social characteristics, such as gender, sexual orientation, age classes and geographical region, to understand how these characteristics shape an individual's, or group's, life experience (Lykke, 2011).

The conclusions made by this work, want to be more than a "recap". Besides summing up what data showed us, trying to investigate the thornier issues in LGBTQ people's daily life, guidelines will be marked. These guidelines contain suggestions and advice, that this research experience bequeaths us and that we would share with who in the future will deal with surveying the LGBTQ population(s), mainly through social media.

Examining ideas about gender performed online, need and freedom of self-defining themselves, relational practices within and beyond physical spaces, imagined communities and belongings, old and new forms of discrimination, this thesis foregrounds the daily life of LGBTQ people in a society that, both in perception and reality, is constantly speeding up (Rosa, 2010). An acceleration which involves, with distinctive features, the LGBTQ people. An acceleration with which social research must deal if it wants to thoroughly study, interpret

and explain the phenomena and the behaviours of people belonging to a certain society

Through an interdisciplinary approach, grounded midway between survey methodology and social sciences, this thesis would offer suggestions, suitable for policy-maker and activist audiences; but also, it would appeal to students and scholars with interests in sexual and/or gender identities in the fields of statistics, sociology and big data management.

PART 1

LGBTQ SURVEYS AND QUANTITATIVE STUDIES REVIEWS

Sexual orientation and gender identity are aspects of who we are.
No one should feel a need to conceal their identity to avoid discrimination, hate or even violence.
But, in the European Union today, many LGBTI individuals still feel the need to do so.

EUROPEAN AGENCY FOR FUNDAMENTAL RIGHTS - A long way to go for LGBTI equality (2020)

1. Surveying LGBTQ people: a definition matter

1.1 Sex-Gender-Sexuality system: the contribution of Gender and Queer studies

LGBTQ can be considered as an umbrella term under where we can find all its components. These components can highly differ from one another, yet at the same time, they can be very close on a civil rights claim occasion. There are three distinct categories of "demographics" that, crossed, may identify LGBTQ people: sex, gender and sexual orientation. These concepts were introduced and differentiate in gender studies, a multidisciplinary theoretical perspective (concerning several academic fields, such as sociology, philosophy and psychology) which started to denaturalize reproductive roles (Butler, 1990) among individuals belonging to a society, to make a distinction amongst biological sex, social gender identity and sexual orientation.

Gender studies were born in North America during the 70s and arrived in Europe some years later. The feminist theory being a strong influence on these studies, pursued by important intellectuals, such as Judith Butler, who in 1990 introduced the concept of gender performativity. Gender becomes something "done", and not attributed. Donna Haraway, who studied the relationship between gender and science, introducing the concept of post-human artefact, in which gender is defined not only by the relationship among humans, but also among humans and nonhumans actors (Haraway, 1991; 1997). Unlike Haraway, Karen Barad tried to recompose the terms dichotomy subject-object, which presumes agencies of observations and objects of observation, as irremediably separated from each other, in scientific knowledge production.

The main definitions of the three demographics (Zevallos, 2014) introduced in gender studies, which identify our target population(s) are sex, gender identity and sexual orientation. Let us now examine these main definitions, in detail.

- *Sex* includes biological characters, according to which a person is defined as male or female. These traits could be overvalued and our role in a society can be interpreted according to our biological characters. This

process could be really misleading.

- *Gender*, on the contrary, is a social identity: it is not an innate human characteristic, but it can vary across space and time. It's assigned by society, which defines how individuals understand their identities, being a man, a woman, a transgender, a queer, and all the other gender definitions. If sex is decided by nature, gender is decided by the society that deems if a job, an activity, or a role is more appropriate for a man or a woman. Gender and sex sometimes are not aligned. This is the case of transgender and intersexuality, two gender categories that in no way regard sexual orientation.
- *Sexual orientation* deals with the sexuality of an individual, attraction, practices, and identity. Just as sex and gender, sexual orientation does not always align and it generates the rainbow of sexual shade that we are studying, called LGBTQ. As gender, sexuality is fluid, so it can change over time, being made of experiences, desires, and behaviours.

These three concepts identify a *system* called the "sex-gender-sexuality system", a term coined by Gayle Rubin, in 1984 to describe "the set of arrangements by which a society transforms biological sexuality into products of human activity". This system defines the social borders, categories and rules, to obey and to transgress, resulting in plural scenarios (Poggio and Selmi, 2012).

Even amongst these three categories, it is important to note the strong distinction between how people identify themselves and how others may label them (Irlam, 2012). Moreover, gender identity construction and the relationship between sex, gender, and bodies are "fluid" and always in progress for everyone. Individuals call into question, all life long, the acknowledgement or the refusal of social rules linked with having a male or a female body, called *gender binary*, or again a body attributable in the continuum between male and female biological antipodes (Antonelli and Ruspini, 2016).

Sex and gender are becoming increasingly recognized by people, in more and more countries, as two separate entities with more than just two possibilities. The phrase "non-binary gender" is more commonly used to refer to people who do not identify as just male or female, nor do they see themselves as a man or woman. (UNECE, 2019).

Gender may be thought of as something unlinked by the sexed body and directly linked to the person. Gender can assume a huge number of variants, called *gender spectrum*, based on the gender identity (self-perception), gender expression (performing gender, dressing, etc.), sexual orientation (choosing sexual and/or affective partner), birth sex and actual sex, in case of transsexual people. Switching from gender identity to sexual orientation definitions, we need to introduce two main currents of thought: the first one classifies people according to their partner choice, the second one considers sexuality indefinable. Actually, the classification according to the partner choice could be extremely difficult and sometimes misleading, if it does not take into account the several dimension of the interpersonal attraction between two people, such as affective, sexual, behavioural, and so on.

In this complex context of definition and classification, all the surveys shown in the following chapters try to categorize people in heterosexual, homosexual, bisexual and one or more adding modalities (sometimes pre-filled options, sometimes an open modality, to leave people the option of expressing themselves as they prefer). These modalities include all those sexual orientations "in-between", important to investigate, mainly because ignored also by sociological research in the past. The most frequent definitions given are *asexual*, one who is not sexually attracted by any other gender, *pansexual*, which have a sexual preference that is not limited by biological sex or gender, *polysexual*, who is attracted by more than one sex but refuses the term bisexual which entails the existence of just two genders. But there are many other definitions that a person can choose to define itself. Actually, definitions of sexuality can never end, and that is the reason why a second current of thought, called queer theory, considers sexual orientation as indefinable, because of the outcome of a social construction (Richardson and Seidman, 2002).

Queer theory refuses both the existence of just two genders and the tripartite division of sexual orientation (in heterosexual, homosexual, bisexual). Both these categorizations are considered "social inventions of an obsolete medicine of XIX century" (Marcus, 2005). All people who feel they have a gender or sexual identity out-of-the-box, according to the rules laid down by society have found in the term "Queer" a valid alternative to any definition (Hughes, 2006; Ahmed, 2006).

Sexual orientation operative definitions are moreover linked to three main dimensions: *identity*, cognitive dimension, *performance*, behavioural dimension,

and *attraction*, emotional dimension (Micheal et al., 1994). Kinsey (1948; 1953) underlined how the homosexual population estimate changed according to the operative definition of homosexuality adopted, and according to the reference period of the question. Sexuality is not a static attribute of a person: it can change during a time, frame and reference period have to be clarified to investigate sexual orientation status properly.

The crucial contribution of women studies first, then gender studies, and finally queer studies, has been to put in the centre of the social debate the removal of inequalities, but without distorting the differences. Identities, defined as natural superstructures, have been dismantled piece by piece, and considered instead as influenced by several different factors such as economy, religion, ethnicities, or historical period. However, in certain fields and working areas, these studies did not find a fertile ground: for instance, in some European academic field, these studies received little attention and resonance, due to the hegemony of patriarchal and heteronormative values still in force in the universities.

A gender-sensitive culture, inclusive and careful to sexuality issues could only raise in the research context, but, unfortunately, this did not happen in many countries, including Italy: academic interest in LGBTQ issues has remained limited, non-institutionalized, and circumscribed to a niche.

Relationships between gender studies and the development of the research about LGBTQ issues arises from the fact that these studies represent a "conceptual paraphernalia", fundamental for investigating the peculiarities of the LGBTQ population(s). Gender and queer studies' concepts, besides giving a solid theoretical framework, represent a toolbox to analyse and conceptualize the literature and the phenomena studied.

In the following chapters and paragraphs, it will be shown as the quantitative analysis could be indelicate, insensitive and careless about the lots of shades which everyone gives to its own sexuality: furthermore, will be deeply described as in the research project Over The Rainbow has been paid a huge attention to the questions about self-definition, looking for a compromise between the phenomenon multidimensionality and the possibility of describing it properly. Moreover, both gender and queer studies, questions and challenges heteronormativity, namely the habit to consider every sexual orientation different by the heterosexual one as deviating from normality, numerically negligible, or even morally questionable. The heteronormativity could lead,

indirectly, to a large set of homotransphobic behaviours, from violent to institutionalized ones. Heteronormativity entailed that every gender identity non-cis, every sexual orientation non-hetero, and every transition was relegated just to stereotypes, clichés, and specks, far away from the complex reality, and slowing down the path of fundamental civil rights recognition for the LGBTQ people.

Concluding, both operative and conceptual definitions are complex in this field. Every psychologist, anthropologist, social researcher or statistician adopts a personal categorization of gender identity, gender behaviour, sexual orientation, etc. All these categorizations can be considered as inevitable structuralism, applied to a set of concepts "done" and not attributable (Butler, 1990), but useful to study all the LGBTQ world dimensions, and not only those dimensions directly linked to people's sexuality.

The quantitative studies reviewed in the following chapters and paragraphs adopt operative and conceptual definitions diversified, according to the aim of the survey and the socio-cultural characteristics of the country. Analyzing these definitions has been an unavoidable step to design the Over The Rainbow project research, conducting interviews, drafting a questionnaire, working on Instagram in the web-scraping phase and, finally, analyzing collected data.

1.2 Historical, political and regulatory framework of LGBTQ rights claims

To fully understand the LGBTQ population(s) nowadays, their characteristics, needs, fulfilments and claims, it is essential to frame the history of the LGBTQ movement, from a socio-political perspective, and to retrace the steps took by this community since the end of the 60s. Generally, the 1969 Stonewall riots are emblematically identified as the starting point of the contemporary homosexual liberation movement all over the world. The riots took the name from the Stonewall Inn, in the Greenwich Village of Manhattan, habitually frequented by the New York LGBTQ community. On the night of 28th June 1969, the police got into the bar to arrest those without identity documents and all people dressed in clothes of the opposite sex, which was prohibited at that time. The riot exploded with about 2000 people shouting "Gay Power!" and other slogans became iconic. The riots continued for days, bringing to light the exasperation of the LGBTQ community, accumulated over years of segregation

and secrecy. The following month, the Gay Liberation Front (GLF) was founded, and a march against the persecution of homosexuals and trans people was organized. It was the first gay pride and that is the reason why, nowadays, these marches all over the world are taken in June, the month of the 1969's Stonewall riots. If New York was the scene of the first riot, another important American metropolis was crucial in the history of the LGBTQ movement: San Francisco was the first city to accept an openly gay man, called Harvey Milk, as councillor. His battles for LGBTQ rights are a milestone, especially his opposition to Proposition 6, which would have implied the dismissal of openly gay teachers in the schools. Its battles were stopped by his premature homicide: Harvey Milk was murdered in 1978 with the major by a former city councillor, inside the town hall. That night more than 30000 people participated in a candlelit procession in memory of the first gay councillor.

In the 1980s the LGBT movement had to deal with the AIDS pandemic: HIV infected many people, mainly among gay men, causing a lot of victims. Furthermore, public opinion started to highlight the strong correlation between homosexuality and the disease, introducing a long-lasting stigma against gay LGBT people, unfortunately still existing. In this context, the LGBTQ community started to actively promote important campaigns about safe sex and the use of condoms for preventing HIV and AIDS. The strong LGBTQs social commitment helped to contrast the stigma and to point out the social engagement of the community against an invisible enemy, which can kill everyone, not only gay men.

During the early years of the new millennium, LGBT movements all over the world started to see their rights recognized: 21 countries have recognized same-sex marriage; several other countries, such as Chile, Croatia, Switzerland and Italy established civil unions between persons of the same sex and some of them allow step-child adoption too. Anyway, the path to full equality among people having different gender identities and sexual orientations is an ongoing process, which has a long road ahead, mainly because the condition all over the world are still various and sometimes really difficult: in many African and Arab countries homosexuality is still illegal, and often punished with imprisonment, sometimes a life sentence or the capital punishment too.

Unfortunately, it does not mean that the situation in Europe is idyllic: the imbalance between the countries of the Oriental bloc countries and the Northern and Mediterranean ones has worsened in recent years, with the rise of several

conservative political parties. In countries such as Russia, Poland, Belarus and Turkey, the LGBT community has to daily protect itself from constant attacks, both institutional and personal. In this context, a strong link, between the feminist struggles and that of the LGBT community, emerged over the years: the suspension of the abortion right in Poland and Turkey's exit from the Istanbul protocol brought back to the streets, side by side, feminist women, gay men, transsexual and transgender people and thousands of activists who wanted to express their solidarity. The endorsement of the European community towards LGBTQ people and their fundamental rights is clear, but too often national governments obstruct European policies, accusing them of being too liberal, and libertine.

Retracing these steps, taken with pride and struggle, and hypothesizing the next steps that the LGBTQ community will want to take (for instance, a law that establishes the crime of homophobia and transphobia) could be a preparatory task to fully understand a fundamental concept expressed in the following chapters, that is, the continuous and changing exploration of new forms of communication. The LGBTQ community, and the people who belong to it, have used social media massively and intensively to claim their fundamental rights. Digitization has created a huge sounding board for people who, for decades, have lived their identity, both gender and sexual, in secrecy and invisibility: if initially, in fact, the chats served as a tool to engineer the task to know someone else, now social media are used by LGBTQ people to spread their thoughts to all the others, first of all to heterosexual people who, immersed in the same cyber-space, can no longer look the other way.

1.3 Studying LGBTQ: epistemological and methodological considerations about a hidden population(s)

Studying LGBTQ poses important questions and challenges: which aspects to study to fully understand the people living under the above umbrella term? And how to do it? These questions introduce both epistemological and methodological considerations. The peculiar characteristics of the LGBTQ population(s) instill doubts on the correct methodology to adopt in studying the phenomena (Coffman et al., 2017): is a lesbian or a gay man ready to truthfully answer an official statistical questionnaire about her/his sexuality? Are the

standard definitions, proposed by researchers, able to fully catch a person's sexual orientation and/or gender identity? Is the standard research approach, used on most social surveys, suitable for studying a phenomena such as the LGBT population?

Another complicating factor of investigating and quantifying gender identity is that for some transgender people, privacy concerning their transgender status is of paramount importance. They can pass in public as their adopted gender expression of man or woman via their appearance and often voice, particularly if they accessed hormone treatments or have had surgery (which can include cosmetic surgery) that helped them to transition physically (UNECE, 2019). This does not necessarily prohibit researchers to ask questions about gender identity, but it does ensure that it is only done where the benefits of collecting information outweigh the intrusion of privacy. In other words, the data collection burden must be compensated with a strong utility of the research, for example when it is used in terms of creating policies and for civil rights movements.

Qualitative methods, employed to study the LGBTQ world, revealed many important aspects of this population(s), gathering many features that the quantitative approach is still struggling to fully catch. Think, for instance, to transsexual people, the scantest LGBTQ subpopulation in terms of numerousness, so strongly characterized by the personal path of every individual: in qualitative inquiry, researchers seek to understand the daily life experiences of participants, without the representativeness sample constraints. Qualitative inquiry is just concerned with "gaining an in-depth, rich understanding of LGBTQ participants' experiences of a phenomenon" (Singh and Shelton, 2011). Moreover, qualitative approaches are useful research frameworks for understanding LGBTQ individuals and communities, because of the many subgroups and the diversity which exist within this population(s).

If qualitative methods have given robust and reliable results, quantitative ones, both nationwide and abroad, have been few and mostly of an experimental nature. Best quantitative studies produced a good output if assisted by quantitative methods, adopting a *mixed methods research* approach (indicated as MMR, from now on). MMR is a strategy that uses both quantitative and qualitative approaches (Johnson et al., 2016; Creswell and Plano Clark, 2011), to enrich social research, combining positive aspects and contrasting negative ones of both approaches. MMR seems to be more appropriate and flexible than other approaches in dealing with the complexity characterizing the LGBTQ

population(s), defined as rare and elusive (De Rosa and Inglese, 2018). Rare, because of the objective limited number of the members of this population, especially of the transgender sub-population. Elusive, because people are not always willing to provide information about their sexual identity. These characteristics have to be taken into account when designing a sample survey on an LGBTQ target population. Having an adequate sample survey to obtain reliable estimates of the LGBTQ population(s) as a whole, and particularly of its sub-populations, can represent a problem: small samples, selected on the list of the whole population, may not be efficient. Moreover, the risk of having a biased estimate is high: selected people can choose to not cooperate with the survey (total or partial non-response), or the responses cannot be sincere (misreporting).

Advances in web technology provide social researchers with increased opportunities to conduct researches with populations that have historically been inaccessible, hard-to-reach and *hidden* (Matthews and Cramer, 2008). In this context, the idea of working on a social network to study LGBTQs has started to ripen.

New social environments, such as Twitter, Instagram, Grindr and many others, can easily lead to a massive number of potential candidates to contact and interview, directly and cheaply; but it can also reassure the LGBTQs' interviewed to feel safe when answering sensitive questions.

1.4 What is the aim of LGBTQ surveys?

In 2011 a report of the European Commission pointed out the importance of having "more and better data, particularly at the institutional level to draw any firm conclusion about the distributive outcomes of social protection reforms" (Nelson, 2011). International social experience research agrees on the use and combination of data for a more targeted, impactful and efficient social policy, which can lead individuals to a fulfilling life (Gluckman, 2017; OECD, 2020). Data relevance for efficient social policies is even more important if applied to sexual orientation and gender equality rights. LGBTQ people have begun to be included within government and non-government policy frameworks. This includes broad health strategies and plans, as well as specific LGBTQ policies, initiatives or programs. However, there has been little work completed to actively secure LGBTQ data to better inform relevant decision-making. Without the inclusion of

LGBTQ data in reporting mechanisms and research, it is challenging to determine the effectiveness of such initiatives. Further, without greater LGBTQ data allocation of future resources/initiatives may be hindered (Irlam, 2012). Let's think about health care: the decision to include (or not include) LGBTQ people in particular policies are often made based on the available data; in areas such as mental health, sexual health, and drug and alcohol abuse, there is significant national evidence of health disparities faced by people attracted to the same-sex. However in areas such as general health research, socio-economic data, mortality data-sets, morbidity data-sets, etc. same-sex attracted people continue to be excluded from national statistics. Moreover, in all these areas transsexual, transgender, and intersex people are still not included (Irlam, 2012).

Intersexual people's situation is the example of the lack of awareness of institutions regarding this issue: some countries are considering what to do about those who are intersex at birth registration, including similar designations such as "unknown" or "undetermined". Medical literature and individual testimony show us that there are many types and causes of ambiguous or intersex situations. An intersex designation may be applied to an individual only later in their life if their puberty is greatly delayed and this leads to new information about their sex traits. In any case, it is not well known among the general population, which could affect data quality in a survey question if it were included (UNECE, 2019).

Socio-demographic data about LGBTQ people could help not only public institutions but private stakeholders too: LGBTQ consumers could be niche markets for many businesses. Access to data on geographical locations, income, family and other general data would be of enormous benefit to companies seeking to pitch their advertising spend towards this niche market.

Given the increasing demand by governments and institutions for data on LGBTQ people, several quantitative studies have been conducted, all over the world, on this population(s). In the following pages the most important of them will be presented. Every survey review will be focused on the methodologies used, on the subjects investigated and on the geographical target area. Every quantitative study has its strengths and weaknesses, and they will be impartially presented, but it is very important to say that every study has been fundamental for this research field because it has increased the knowledge about LGBTQ population(s), both at a national and international level. Analyzing in-depth

these quantitative studies and conducting the following review has been crucial for the present research project because it has addressed the questionnaire draft, the choice of using a snowballing sampling method, and the theoretical framework in which to collocate the Over The Rainbow research project.

2. International studies

Significant progress at an international level has been made towards including LGBTQ people within national surveys across a wide range of topics. Quantifying LGBTQ population size is a challenge that has lost its research importance in favour of studying LGBTQ related issues. Measuring the LGBTQ population(s) phenomena is a methodological problem faced, in the last twenty years, in many different ways by National Statistical Institutes (abbreviated as NSIs, from now on), ministries, private research institutes, and other data releasers. What they all have in common is the inclusion of sexual orientation and gender identity questions in massive population surveys, to have in the sample a notable number of people belonging to this population(s). Several countries, for example, recognize same-sex couples in their national Census. However, this kind of phenomenon recognition is partial because it excludes all those LGBTQ people who are not in a couple (De Rosa and Inglese, 2018).

The international sociological and statistical literature acknowledges the emerging importance of collecting sexual orientation data, along with the difficulties of respondents answering questions, where sexual orientation and gender identity concepts have been poorly defined or understood. The definition matter, previously mentioned, has often been "underrated" during the planning phase of many quantitative studying experiences carried out all over the world: aside from Nepal's recent inclusion of a "third gender" in part of their national Census, there has been no international discussion identified about the inclusion of trans/transgender or intersex people within Census. The US Department of Minority Health has committed to the inclusion of gender identity within population health studies and is currently consulting and testing on question designs. A considerable number of health and population surveys include sexual orientation and gender identity indicators within them. (Irlam, 2012)

Every international quantitative experience about LGBTQ has its own strength and weakness, which address the work of all the other countries. In the following pages important survey experiences regarding the LGBTQ population(s) carried out in Australia, Canada, India, Pakistan, Nepal, New Zealand, UK, USA will be presented and analyzed.

2.1 Australian pioneering studies for estimating non-heterosexual population

Australian organizations have been international precursors in the field of LGBTI research with a long set of surveys, research and quantitative studies, started in the 90s. The Universities of Sidney and of South Wales partnered in 1996 creating the Sidney Women and Sexuality Health Survey, conducted by Acon, one of the most Australian community organizations against HIV. In the same year, the National Centre in HIV Social Research with the Kirby Institute, the state AIDS council and the state Health Department started to conduct the Gay Community Periodic Survey (in Adelaide, Canberra, Melbourne Perth, Queenisland and Sidney).

In the international official statistics field, Australia has been a pioneer too. Australian Bureau of Statistics (ABS) has carried out a set of surveys for a long time intending to estimate the Australian non-heterosexual population and to quantify all the phenomena linked with it. This set of surveys include the Household Income and Labour Dynamics in Australia survey (HILDA), carried out in 2012, the Australian Study of Health and Relationships (ASHR), carried out in 2002, and the ABS General Social Survey (2014). Moreover, the Census of Population and Housing (since 1996) and Mental Health and Wellbeing (SMHWB, since 1997) collect data about sexual orientations and, since 2011, the civil status of same-sex couples has been asked. In 2016 Australia collected data on those who identify as other than male or female in the Census. In the same year, third sex and gender options were introduced to the national standards for sex and gender variables.

In 2018, in Wilson and Shelley's metadata study, prevalence rates of the non-heterosexual population over 18 were averaged through three ABS surveys and multiplied by Estimated Resident Population (ERP) to obtain an estimate of the national population. Then, census data on same-sex couples were used to distribute the national estimates by state and territory (Wilson and Shelley, 2018). The results showed that Australia's non-heterosexual adult population, in 2016, was estimated at about 592 thousand people, representing about 3.2% of the whole adult population. A relatively small population, but with a varying prevalence by age and sex, but also between states and territories. For example, New South Wales was home to the largest non-heterosexual population (about 204,000) and the Northern Territory the smallest (4,700), while the highest

prevalence was in the Australian Capital Territory: 5.1% (Wilson and Shelley, 2018).

A higher percentage was estimated in 2015 by Roy Morgan Research Institute (RMRI), which provocatively titled its study "Australia is getting gayer". RMRI conducted a longitudinal 3-waves study, between 2006 and 2014, asking almost 180,000 Australians aged 14+ to define their sexual orientation. In 2006-08 2.4% of the population defined themselves as homosexual. By 2009-11, this percentage had risen to 3.1%. During the latest triennium 2012-2014, the figure had risen again: 3.4%. Higher than the Wilson and Shelley's estimate of 3.2%.

In 2017 ABS was commissioned by the government also to design and conduct the Australian Marriage Law Postal Survey. The postal survey aimed to gauge support for legalizing same-sex marriage in Australia. ABS sent to every person on the federal electoral roll, a reply-paid mail asking the question "Should the law be changed to allow same-sex couples to marry?". The survey returned 61.6% "Yes" and 38.4% "No", which pointed out a definitive approval of the Australian population to same-sex marriage, which has been legalized since December 2017, by the Marriage Amendment.

2.2 Statistics Canada: a benchmark for many others NSIs

Statistics Canada (StatCan) introduced in 2003 questions about sexual orientation in its Community Health Survey. During the computer-assisted telephonic interview (CATI) and the personal interview (CAPI) they asked 160,000 randomly selected individuals to define their sexual identity (heterosexual, homosexual, bisexual, unsure, other). Results pointed out an estimated percentage of homosexuality at around 1% of the whole population aged between 18 and 59. Bisexuality was instead estimated at around 0.7%.

In 2006 StatCan considered the issue of sexual orientation as part of the Census too. During the census design phase, StatCan conducted several focus groups. The use of this qualitative research methodology showed that the survey context was important to provide an explanation as to why the question was being asked, and how the data could be used. For example, participants were most willing to answer questions within the context of health surveys or in a discrimination and human rights survey. Most participants did not approve of including a sexual orientation question on the Census (StatCan, 2006). Despite

focus-groups suggestions, since 2011, the civil status of same-sex couples has been asked in census, analogously to Australia, and test questions have been refined. For example, to be more sensitive regarding proxy response, StatCan changed the question wording from "what is your gender identity?" to "what is your gender?". The 2016 population Census collected data on those who identify as other than male or female. StatCan became a benchmark on this field for the other NSIs and, in April 2018, published standards for investigating sex assigned at birth, and gender in the social surveys. Canadian experience gives much importance to the qualitative perspective, on which the 2021 Census consultation process grounds.

Qualitative methods introduced in the design phase (such as focus groups, interviews, cognitive testing of questions, etc.) many advantages: discussing with people belonging to the LGBTQ target population about arguments, wording, and questions proposed in the questionnaire, guarantying a better final survey result, for example in terms of a lower non-response rate. Qualitative methods will be introduced in the design phase also by many other NSIs, universities, and research centres. One of them is the McCreary Centre Society, which surveyed young people in British Columbia (BC, the western Canadian province) in three BC Adolescent Health Survey waves (1992, 1998 and 2003). This longitudinal survey tested various methods before measuring sexual orientation, trying to take into account various aspects: who someone is sexually and emotionally attracted to, who they are sexually active with (if they are sexually active), and how they actually identify or self-label. The importance of these aspects may not be reliable, especially during the adolescent phase: in this life period, a consistent part of teens are not yet sexually active, while some teens may have been coerced to have sex with someone they are not attracted to. Some teens may not be experiencing attractions yet at all, and others may be unsure what their orientation is. Moreover, since being gay, lesbian, or bisexual can be stigmatized, some students may choose to hide an LGB orientation by dating someone of the opposite gender; analogously, a bisexual student with an opposite-gender partner might be assumed to be heterosexual. To be safe, students may choose not to identify their orientation publicly ("come out"), they could also be from a culture that does not recognize the labels or identities that are commonly used in Canada (Saewyc et al., 2007).

Longitudinal survey data analysis revealed both hopeful and worrying trends, fewer than expected differences between rural and urban LGB youth, and

ongoing health disparities for LGB teens compared to their heterosexual peers. Females in 2003 were more likely to identify as "mostly heterosexual" or bisexual, than in previous years. Among them, the rate of sexual and physical abuse increased, while it declined instead among gay males, during the observation period. Between 1992 and 2003 the percentage of LGB students reporting sexual orientation discrimination increased for gay males and bisexual teens; correlated to the latter evidence is the rate of suicide attempts, which increase for lesbian and bisexual females over the three surveys, but declined for gay and bisexual males. According to Rural and Urban Differences, the longitudinal survey showed a strong difference between contexts: rural gay and bisexual males were more likely to report sexual abuse, and more likely to have attempted suicide in the past year. At the same time, both male and female rural LGB teens were more likely to report that they had been in contact with a stranger on the Internet who made them feel unsafe. LGB boys and girls resulted to be current smokers, to have tried alcohol, or to have used other drugs. Moreover, they were more likely to have reported emotional stress, suicidal thoughts, and suicide attempts (Saewyc et al., 2007).

Canadian survey experiences about the LGBTQ population(s) has been, not only exhaustive and a benchmark for many other countries, but also specific regarding important aspects such as the LGB adolescences field, the introduction of qualitative methods for drafting the questionnaire and employ of mixed-mode data collection (CATI-CAPI) in statistical surveys.

2.3 Quantifying the Hijras, ancient third gender of the Indian subcontinent

LGBTQ living conditions in India, Pakistan and Nepal are really different from those in Canada and Australia. In India, for example, homosexuality has been decriminalized just since 2018, after decades of illegality, started during British colonial control of India. Gender reassignment surgery (GRS) was allowed in 2014.

This legislative rigidity collides with the cultural acknowledgement from antiquity all over the sub-continent of an indigenous third gender, called *Hijra*. This third gender has been included since 2014 as an option on passports, administrative documents and civil registers. In 2011 the Indian census schedule allowed to elect a gender indicator other than male or female (Irlam, 2012). The

resulting estimate of this third gender in India was of 490,000 individuals in 2011, 0.04 per cent of the whole population. Transgender activists estimated a number six or seven times higher but they were surprised that such a large number of people identified themselves as the third gender, even though census counting occurred three years before the Supreme Court order gave legal recognition to Hijra, in 2014 (Nagarajan, 2014). Furthermore, activists specified that not all hijras would want to identify as "third gender", suggesting that many would prefer to identify as male or female.

In Pakistan, the Lahore High Court issued an order to include a third gender option in its 6th Population and Housing Census, in 2017: 10,418 people defined their gender as neither man nor woman. The forms were interviewer-administered, as in India and Nepal. With these data collection modes (CAPI or PAPI) responses may have been affected as this is a sensitive question and people may not have been comfortable disclosing such information face to face. However, an interviewer may have helped clarify any confusion leading to better data quality (UNECE, 2019).

Nepal, in 2011, planned to introduce the third gender option in the census schedule, being the first Central Bureau of Statistics to do it, but the estimate was not given. Contrary to what the Nepalese Bureau had announced before the census enumeration, the third genders were not recognized in the census. In fact, the third gender category was not included in the detailed questionnaire, but in the household listing form (Chhetri, 2017). This "change of mind" makes the number of hijra in Nepal unfortunately still unknown. Several key issues, such as the unclear definition of the third gender, inadequate enumerators training, and a lack of software able to distinguish three gender categories caused the census process interruption (Knight, 2011).

Approximately one year after the conclusion of the 2011 Nepalese census, The Williams Institute of UCLA, in partnership with the Blue Diamond Society (an important Nepalese sexual health and human rights organization) designed and implemented a survey, focused on demographics, self-identification, discrimination, and HIV. The aim of the William Institute's survey was not to estimate the population but "to develop best practices for gathering data and to provide future survey designs with a model of survey questions that accurately reflect and include sexual and gender minorities" (Knight, 2014). This survey proved that an LGBTQ survey can be informative, useful, and accurate also if it has no inference ambition. Estimation of the whole target population is one of the

main goals for a quantitative study, but it is not definitively the only one: trying to understand and quantify discrimination and self-identification modalities of the LGBTQ population(s).

2.4 A long design phase to introduce SOGI questions in the survey: the New Zealand experience

For several years, Stats NZ has been working, alongside interest groups and agencies, to better understand gender and sexual orientation topics and how to collect more robust data. This ongoing work ultimately aims at including these topics in all social surveys, and finally in the 2023 Census (after two postponements heavily disapproved by the Kiwi LGBTQ community). The project has been conducted across several government agencies, also employing qualitative methodologies during a deep consultation phase. After this consultation, Stats NZ released a new statistical standard for gender identity. This standard gives guidance on the collection, classification, and dissemination of information on gender identity, which was the first of its kind in the world. The Stats NZ standard is similar to the Canadian one, which includes cisgender and transgender people in the definitions of male and female. Till now questions about sexual identity have been introduced in the General Social Survey (GSS), which provides information on the wellbeing of New Zealanders, and (for the first time) in the 2019/20 Household Economic Survey (HES). A combination of these two massive surveys will show the well-being and the economic disparities of Kiwis declaring different sexual identities. The objective of this innovative project is to enable New Zealand's LGBTQI+ community to see themselves reflected in the collection of household data and enable these groups to be better reflected in a range of social and economic outcomes covered by Stats NZ's household surveys (UNECE, 2019).

In the 2018 GSS, 96.5% of adults in New Zealand identified themselves as heterosexual, 1.9% as bisexual, 1.1% as gay or lesbian and 0.5% as other identities, such as asexual, pansexual and takatāpui (a Maori term which defines the partner of the same sex). The results of the 2018 GSS are consistent with those reported by the New Zealand Health Survey conducted by the Ministry. On the contrary, GSS results are quite lower than those reported by the 2016 Attitude and Values

Longitudinal Study, conducted by the University of Auckland on a smaller sample, which makes its estimates less reliable than those published by Stats NZ.

New Zealand conducted a long design phase to introduce gender identity and sexual orientation questions in the survey schedules. This lengthy process is giving great results, in terms of output estimates quality and data analysis capability.

2.5 The importance of data for the LGBT Action Plan in the U.K.

In 2006 the British Office of National Statistics (ONS) started to think about the inclusion of sexual orientation questions in the 2011 Census (Irlam, 2012).

In the same period, the Parliament of the United Kingdom promulgated the Equality Act, which aimed to address a set of anti-discrimination laws and practices in Great Britain. This act made data about gender identity and sexual orientation of British citizens more and more needed for "equality monitoring, policy planning, and public service provision" (UNECE, 2019).

In this context ONS was responsible for collecting reliable data on LGBTQ people; some necessitude has caused a stall in the new survey's design, and ONS does not currently measure gender identity on any of its surveys, but it recently published plans to collect data on binary sex and gender identity in the next 2021 census. In these years, ONS researchers have conducted a testing phase to understand how to accurately combine a binary sex question with gender identity, minimizing the non-response item effect.

In 2017, the Government Equalities Office (GEO) launched a nationwide LGBT survey, asking the LGBT population in the UK questions regarding public services and daily experiences. The survey has become the largest national survey to date of LGBTs anywhere in the world, receiving over 108,000 responses, during the 3 months data collection period. The survey was open to anyone over 16 living in the UK and who identified as LGBT, including both any minority sexual orientation (such as asexual, pansexual, etc.) and gender identity (such as non-binary or genderqueer), but also intersexual individuals. The online survey is an example of massive mixed-methods research (MMR) collecting and mixing both quantitative and qualitative data. Voluntary respondents were invited to participate in the survey during the 2017 LGBT pride celebrations, via national media coverage, and on social media. The self-selected sample does not

guarantee any representativeness of the entire LGBT population in the UK. Most importantly, respondents had to be willing to self-identify as LGBT; these people may have a different experience to those who are unwilling to identify in this way, even in an anonymous survey.

The survey aimed to develop a better understanding of the life experiences of LGBTIs, particularly in the areas where LGBTI people face the largest inequalities, such as health, education, personal safety, and employment. Understanding the proportion of the different components of the population(s) is crucial to interpret the phenomena. This was the composition of the sample, release in the summary report (2018): 61% of respondents identified as gay or lesbian and a quarter (26%) identified as bisexual. A small number identified as pansexual (4%), asexual (2%) and queer (1%). These labels distribution varied by age. For example, younger respondents were more likely to identify as bisexual, asexual, pansexual, queer, or 'other' (39% of cisgender respondents under 35 compared to 14% of cisgender respondents over 35). This reflects work undertaken by the ONS that shows younger people were more likely to be bisexual than older people. Besides, 13% of the respondents were transgender (or trans). Of the total sample, 6.9% of respondents were non-binary (i.e. they identified as having a gender that was neither exclusively that of a man nor a woman), 3.5% were MtoF trans women and 2.9% were FtoM trans men.

Another goal of the survey was to investigate the life satisfaction of LGBTQ people, to understand if their gender identity and/or sexual orientation could negatively influence their daily satisfaction. On average, respondents were less satisfied with their life nowadays than the general population, scoring it 6.5 out of 10, compared with 7.7 for the general UK population. Among cisgender respondents, gay/lesbian people had the highest scores (6.9) and pansexual or asexual people had the lowest scores (both 5.9). Transgenders also had low scores: trans men scored 5.1, trans women scored 5.5 and non-binary people scored 5.5. Overall being LGBTQ in the UK is socially accepted and over half of the respondents (56%) felt comfortable, rating their comfort as a 4 or 5 out of a 5 Likert scale, to identify themselves as part of the LGBTQ population.

On the other hand, some findings pointed out by the survey are extremely concerning: data confirms that LGBT people are at greater risk than the general population, of being victims of crime. In-depth, 40% of respondents had experienced physical, emotional or verbal abuse in the 12 months preceding the survey committed by someone they did not live with and because they were

LGBT. Around a quarter (26%) had experienced verbal harassment, insults or other hurtful comments, 14% had experienced disclosure of their LGBT status without permission, 6% had been threatened with physical or sexual harassment or violence, 2% had experienced physical violence and 2% had experienced sexual violence (GEO, 2018).

The National LGBT Survey collected a huge amount of data and an impressive set of quantitative and qualitative analysis. Both in the summary and the research report (published on the GEO's website) tables, graphs and interview shreds are presented to exhaustively illustrate the LGBTQ living condition.

Results are relatively good, in terms of integration and inclusivity, but they are defined as "sobering" too because promoting civil rights policies is a long way. That is the reason why a comprehensive LGBT Action Plan has been published, as the final output of this remarkable national survey.

2.6 Underestimation of the LGBT American population size and anti-gay sentiment magnitude

The U.S. Census Bureau, the official American statistics producer, does not currently collect information on gender identity on their censuses or social surveys. The question of sex is still asked through a binary response option but since 2016 the U.S. Department of Labour has been analyzing the "feasibility of adding SOGI (sexual orientation and gender identity) questions" to the Current Population Survey (CPS) (UNECE, 2019). The research aims to ask about SOGI in the context of a labour force survey. Focus groups have been conducted to deal with the accuracy and sensitivity of the questions and to understand if interviewed people could be able to answer a new formulation of these questions. Respondents do not find SOGI questions difficult or sensitive to report for themselves or others in their households. However, lesbian, gay, bisexual, and trans (LGBT) respondents found it hard to "align their self-identity with the response options provided" (Ellis et al., 2017), especially transgender respondents.

If American official statistics surveys about LGBTQ are still in progress, universities and private research institutes have been working since the 90s in this field. *Williams Institute on Sexual Orientation and Gender Identity Law and Public*

Policy estimated in 2012 that 3.4% of the whole American population defines itself as LGBT (Gates and Newport, 2012). In the second edition of the research, based on Gallup's survey, it increased to 4.1% (Gates, 2017), 0.6% of which transgender.

The list of the quantitative studies carried out for studying the LGBTQ American population(s) is boundless (Green, 2012; Gates, 2011; Randall, 2020; Pew Research, 2013; ILGA, 2018) and all of them, over time, give less importance to the proportion of LGBT subpopulations on the total (which is always between 1.5 and 6.5%). Changing data collection mode, over the last few years, has worsened the data reliability issue. The spread of web-interview led to higher estimates but also a higher non-response rate, problem indissolubly linked with social desirability, privacy, and anonymity issues. Instead, more and more significance is given to daily life, to homophobic events, and all those problems faced by the LGBTQ population(s). Understanding these issues over the last two decades has acquired a scientific value, practical importance but mainly for policy purposes.

In this context, the use of qualitative methods for improving the surveys, such as interviews, ethnography, but mainly focus groups, has been encouraged and employed (Fryrear, 2016; Singh and Shelton, 2011; Wronski, 2019; Ellis, 2017; Badgett, 2009), also instilling doubts about the use of quantitative methods in this field. In 2013, Coffman wrote: "*the size of the LGBT population and the magnitude of anti-gay sentiment are substantially underestimated*", because people were not ready to answer truthfully to an official statistical questionnaire about their sexuality.

American extensive experience of qualitative sociology pointed out that the standard research approach, used on most social surveys was not suitable for studying the phenomena regarding the LGBTQ population(s). Moreover, people involved in a survey about sensitive arguments want to be reassured about privacy conditions, anonymity, the aim of the research, and the credibility of the institution, which is conducting the research. In this context, the questionnaire-drafting phase assumes a crucial role, that will be deeply discussed in the following paragraphs.

3. European studies

3.1 The "long way to go for LGBTI equality" and the awareness of data need in the European Union

The 21st article of the EU Charter of Fundamental Rights articles guarantees that "*any discrimination based on any ground such as sex, […], age or sexual orientation shall be prohibited*". Despite the cultural advancement of Europe (although with national characteristics of the single countries) both national and European institutions recognized the lack of robust and comparable data about the LGBTQ community as a whole, which is necessary to guarantee respect, protection and fulfilment of their fundamental civil rights (FRA, 2012).

European LGBTQ people have enjoyed several important rights in the last decades but with strong differences between countries: 13, out of the 27 countries belonging to the European Community, legalized same-sex marriage (Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Luxembourg, Malta, the Netherlands, Portugal, Spain, and Sweden). Further, 8 European countries legalized civil unions for same-sex couples (Croatia, Cyprus, Czech Republic, Estonia, Greece, Hungary, Italy, and Slovenia). Joint adoption by same-sex couples is allowed in 13 countries, the same which allow same-sex marriage, while step adoption is allowed also in Croatia, Slovenia, and Estonia. Homosexual people are allowed to serve in the armed forces in every European country, except for Cyprus. Most of the countries have generic anti-discrimination laws, and some of them promulgate specific laws against homosexual hate crimes.

The European Union (EU) and the Council of Europe, supported by the United Nations (UN), have tried to standardize the recognition process of non-discrimination and equality for LGBTQ people: in 1999, articles 10 and 19 of the Treaty of Amsterdam officialised the European purpose of combating every kind of discrimination on the grounds of sexual orientation. After ten years article 21 of the Charter of Fundamental Rights declared the prohibition of any discrimination, explicitly including those based on sexual orientation.

In this heterogeneous and variable context, the situation of the European LGBTQ population(s) is no longer a marginalized issue, but a recognized human

rights concern (FRA, 2012). Awareness of this issue has led to an increase of attention towards data regarding LGBTQ: Europe has put in place an important research project to investigate its LGBTQ population(s), intending to plan efficient and adequate EU social policies.

The European Union has tried for years to collaborate with national governments and satellite institutions to solve the lack of knowledge about LGBTQ data by providing data to national and European stakeholders. In the following paragraphs the most important surveys carried out in Europe by the main institutions, will be presented and analysed, highlighting differences and similarities with international experiences, previously presented.

3.2 How many people are not heterosexual in European countries? A question whose importance is waning.

If the European Union, and several of its satellite institutions, have made notable efforts that led to an increasing volume and quality of LGBTQ data, on the other hand, single European countries are still working on this field. National surveys find it hard to investigate exhaustively gender identity and sexual orientation: privacy problems, ethical qualms and political obstructionism hinder the work of NSIs and quantitative research centres. In this context, the most important and useful surveys about LGTBQ, were not directly conducted on the target population but the whole one, to provide a current overview on peoples' attitude towards LGBTQ (FADA, 2017; Istat, 2011). In this way, national ministers and governments studied discriminatory behaviour without directly asking about the sexual orientation of individuals. In Germany, for instance, the Federal Anti-Discrimination Agency (FADA) carried out in 2017 a telephonic survey on 2,000 people age 16 or older to understand the attitude of Germans towards lesbians, gays and bisexuals. The survey pointed out that 81% of respondents perceive that LGB people still experience discrimination, with the persistence of prejudice but a decreasing trend of traditional homophobic behaviours (12% of respondents showed bad attitudes such as disparaging homosexuality as being immoral or unnatural as well as denying equal). A similar survey was conducted in Italy to investigate the persistence of stereotypes and discrimination (it will be specifically presented later).

Single European countries are leaving behind the idea of estimating the proportion of LGBTQ people overall population because the main civil rights national policies on this field can be successfully planned without knowing that proportion. The magnitude of peoples' attitude towards LGBTQ, such as inclusivity and acceptance, as well as, homophobia and prejudice, have become of higher interest for national policymakers, NSIs and public social research centres.

3.3 Counting the European LGBTQ population: a "private matter"

During the last few years, in European countries, the proportion of LGBTQ in the whole population has seemed to become more interesting for private opinion polls and market research than for public policy planning.

In France, for example, the French Institute of Public Opinion (IFOP) conducted two surveys, the first in 2011 and the second three years later, to quantify the sexuality aspects of the French population. In the first one 6.6% of respondents identified themselves as homosexual (3.6%) or bisexual (3%), and 90.8% as heterosexual; in the second one the heterosexual percentage decreased to 90%, with strong differences depending on age, gender and region. IFOP also conducted an important job about feminine sexuality, studying sexual attraction only among women, through a web survey conducted both in France and in other seven European countries (IFOP, 2017).

Another important experience of LGBT demographics is Dalia's one (Lam, 2016). This private market and opinion research centre conducted, in August 2016, a "census-representative" survey of 11754 people across the EU. Dalia's survey disclosed a 5.9% proportion of LGBT people on the whole European population. The survey directly asked "*Do you identify as lesbian, gay, bisexual, or transgender?*" and the possible answers were "*Yes/No/Prefer not to say*". The sexual orientation was further investigated through a question based on the heterosexual/homosexual rating scale[2] proposed by Kinsey in the late 40s. According to this second question, the share of people who identify as not only

[2] Kinsey's scale (1948; 1953) has been widely used in sociology to detect the perception of one's sexual preference, based on four dimensions: sexual identification, type of sexual attraction, type of sexual fantasy, and sexual behavior.

heterosexual was nearly twice as high (10%) as the percentage of those who identify themselves as LGBT from the yes/no question (5.9%). Results showed also strong gender and age differences: more women than men identify themselves as homosexual, and young people are more likely to describe their sexual orientation as something other than only heterosexual (16% between the ages of 14 and 29, compared to 7.5% between 30 and 65).

This research, conducted by private enterprises, shines a light on the sexual emancipation of people and on the right to perform its own gender, but their results have been often criticized. In fact, Dalia and IFOP percentages are quite high and two poll institutes don't hide the substantial size difference between their survey results and others. Private opinion polls and market research enterprises know that SOGI questions are subject to misreporting, uncertainty and non-response problems, which can lead to biased estimates. But, at the same time, Dalia's researchers specify, in the final report, that private survey "offers a degree of anonymity to encourage honest answers, and get a better count", and though imperfect, their LGBT population count could be an important step forward for the LGBT minority often unrepresented in national census data.

3.4 Life on Margins'. A study to fill the data gap in South-eastern Europe

One of the most significant surveys in Europe about LGBTI was carried out by the World Bank Group (WBG) in Albania, Bosnia and Herzegovina, Croatia, Kosovo, Macedonia, Montenegro, and Slovenia. In 2018, 2300 LGBTI interviewed people from all over South-eastern (S-E) Europe, shared their experiences in the largest-ever CAWI survey for sexual and gender minorities in the region (Van Gelder et al., 2018).

The research, named "Life on the Margins", is the perfect example of the survey which does not have any ambitious aims of estimating the proportion of the LGBTQ overall population. Interviewing just LGBTI people is impossible to calculate the proportion of the whole population. WBG preferred to provide, to both European and national policymakers, the magnitude of homotransphobia behaviours in S-E Europe, telling a story of discrimination, exclusion, and violence, through a detailed quantitative report.

The main characteristics of Life on Margins is the focus on two main issues: the economic problem led by a discriminating environment, and the importance

of giving reliable data to institutions that plan economic policies. Observing from an economical point of view the discrimination against the LGBTQ population(s) is such an innovation for the quantitative study proposed in this field. WBG proposed a vision, by which "exclusion based on sexual orientation and gender identity (SOGI) is costly to both the individuals concerned as well as a country's economy and society as a whole". WBG argued with socio-economic data that social inclusion of LGBTI people is therefore important in and of itself, but also because it is the smart thing to do. South-eastern Europe is currently the most developing scenario in Europe; countries are rising up from the ashes of the old Yugoslavia, investing assets and efforts in tourism, infrastructures and services. From a macro-economic point of view, WBG suggests that "more inclusive societies are, more likely to make the most of their entire stock of human capital. More open and inclusive cities are better placed to attract international capital and talent. More open and inclusive countries make attractive international tourist destinations" (World Bank, 2018). In this context of economic and social growth, the data contained in the WBG report presents the challenges and the obstacles experienced by the LGBTI population(s) in S-E Europe. Solving the inclusion problem of this particular population(s) would bring many benefits, both social and economic, to the whole region.

The second issue on which the report focuses is the lack of knowledge about LGBTI, all over the world, but more accentuated in Ex-Yugoslavia. WBG staff claims that "the LGBTI data gap remains large, and further research and data collection are necessary to better understand the lived experience of LGBTI people and the challenges they face. National statistical agencies should begin to collect LGBTI-disaggregated data to provide the up-to-date evidence needed to build more inclusive policies and programs at the country level, thereby aligning themselves with statistical agencies in advanced countries" (World Bank, 2018).

Life on Margins' survey, for the first time, sheds light on the experience of LGBTI people in the region and gives a macroeconomic interpretation of the social policies that should be undertaken. Better LGBTQ inclusion conditions will entail not only a quality-life gain for individuals but an improvement of national economies and societies (Van Gelder et al., 2018).

3.5 Europe's main actor against discrimination: the Fundamental Rights Agency

The European institution responsible to counter discrimination is the Fundamental Rights Agency (FRA), an "independent centre of reference and excellence for promoting and protecting human rights in the EU". The FRA collects and analyzes comparable data with the aim of helping better law-making and implementation.

To fulfil the need for comparable data about LGBTI the FRA launched in 2012 the European Union (EU) online survey of LGBT persons' experiences of discrimination, violence and harassment. The Agency carried out a CAWI survey to address the lack of robust, statistical data on the life experiences of LGBT people in the EU.FRA's survey, partnered by the opinion polls company Gallup and the International Lesbian and Gay Association (ILGA), was the largest ever EU-wide LGBT survey, with over 93,000 respondents. The target population was made aware of the survey through several websites, magazines and associations' portals. This "invite-all" strategy involved a set of activities to publicize the survey, motivate people to drag others into the survey and finally cover different respondent groups, both at socio-demographic and at membership levels.

The main thematic area of the questionnaire referred to many important issues, such as background and self-perception of LGBT individuals, public perception of them, hoped-for policies against homotransphobia behaviours, civil rights consciousness, discrimination (perceived or experienced), visibility in social environments and institutions, aside from the usual socio-demographic questions.

Discrimination in the workplace was one of the most important questionnaire areas, distinguishing job hunting and job conduction phases. Surveys pointed out that, in the last 12 months, 20% of LGB workers (or job-searchers) felt discriminated at work for their sexual orientation; this percentage rises to 33% among transgender people, who had more difficulties finding a job, mainly because of their appearance and problems with ID card changes.

Another peculiarity of FRA's survey was the consultation introduction during the survey design phase. Stakeholders, academic experts and institutions that work for, or promote, LGBTQ equality and human rights in the EU were involved in defining the main methodological survey aspects, such as sampling, communication, questions target, and privacy issues. The questionnaire has been

tested in 5 countries through a cognitive test, to improve and fine-tune research tools.

The sample survey was auto-selected because there is no population frame from which to select a random sample. It means that associations and institutions advertised the survey through their own formal or informal social networks. The direct consequence was the higher probability of reaching people involved in "awareness-raising activities" (FRA, 2012). This is an unavoidable characteristic of every LGBTQ survey, and maybe its most problematic aspect. Trying to adjust this bias, FRA introduced a sampling weight system to correct the over/under representativeness of LGBTQ associations, over the countries. FRA, in 2012, did not estimate the global amount of LGBTQ people in Europe, or the proportion on the whole European population, but focused on the quantification of bad behaviours against LGBT people, such as bullying, mobbing, or legislative discrimination.

In 2019 a second survey was carried out to give the research the added value of a longitudinal study, which aims to compare results from the previous surveys to assess the effectiveness of policies and measures to combat discrimination and victimisation, and to promote equal participation in society (FRA, 2020). With almost 140,000 participants, it was the largest survey of its kind in the European Union (including North-Macedonia and Serbia, which do not belong to the EU).

Results confrontation between two 2012 and 2019 surveys is pitiless, if anyone expected a great leap forward in LGBTQ fundamental rights and daily life conditions, also if there are relevant differences between the Member States. For instance, discrimination remains a reality, mainly at work: in 2019, the proportion of who felt discriminated against in job-search (11%) is only slightly smaller than it was in 2012 (13%); the proportion of respondents who feel discriminated against at work in 2019 increased (21%), compared to 2012 (19%). For trans people, this proportion was higher (22% in 2012) but it increases further in 2019 (36%). Reporting discrimination incidents remains difficult but the share of LGBT respondents who reported slightly increased, from 13% in 2012 to 17% in 2019.

Out of the workplace, the situation gets better: the share of LGBT respondents, older than 18, opening up about their own sexual orientation increased from 36% in 2012 to 52% in 2019. Being open at school or university (for respondents between 18 and 24) is still difficult but the trend is getting better, passing from 47% to 41% in 2019. Anyway, the main concern for the European

institutions remains physical and sexual attacks: 10% of respondent experienced violence in the five years before the survey; more than 50% of LGBTQ respondents still avoid holding their partner's hand in public and prefer to be "discreet" (FRA, 2020).

The EU Commission has called for implementing a "List of Actions to advance LGBTI Equality" to tackle discrimination against LGBTI people (European Union, 2016). These actions cover all policy areas that are relevant for LGBTI people, such as non-discrimination, education, employment, health, free movement, asylum, hate speech crime, enlargement and foreign policy. These actions aim to improve the social acceptance of LGBTI people and enforce EU legislation. For implementing these actions the Commission counts on the collaboration of several important institutions, agencies, international organizations (such as the OECD and the UN) and, last but not least, the European civil society.

In conclusion, we could say that European institutions are working (and collaborating) hard to improve LGBTQs life conditions, through important and challenging actions. But beyond those actions, aimed to change states and institutions policies, all of us should look at what we can do in our streets, our cities, and in our neighbourhoods. Fighting prejudice and intolerance have to become a daily behaviour that everyone pursues. Beyond the slogans, Europe should become a place where everybody is free to be who he or she is.

The path is long, but the EU seems to be motivated to guarantee the fundamental rights of LGBTQ people, supporting member states in promulgating efficient social policies.

4. Italian studies

4.1 The national context and the use of mixed-methods in the research about LGBTQ in Italy

In Italy, until 1969, homosexuality was quite a taboo and some, bipartisan, legislative proposals tried to proscribe every kind of non-heterosexual behaviours. Social condemnation against homosexuality was so strong as to force many gay men to move from rural areas to bigger cities like Rome, Milan or other metropolitan cities; at the same time, lesbian women were about invisible in the eye of society (Benadusi, 2007).

During the 70s LGBT Italian activism started to take shape and to claim basic civil rights, also involving politics (Zanola, 2014). In the following years, the most important LGBTQ associations were founded and some political parties, such as Radicals and the Italian Communist Party, started to open and support LGBT claims (Corbisiero, 2013).

In this context, social research started to orientate its interest through the LGBTQ issue, to unravel the sociological, demographic and economic characteristics of this population(s). Studies on gays and lesbians (Richardson and Seidman, 2002) began to spread in our country with the interdisciplinary approach typical of sexuality issues, to deeply analyze identities, representations and behaviours (Benadusi, 2007). At the same time, Italian institutions started to notice the lack of reliable data about LGBTQ and also to slowly invest in data retrieval.

Social researchers started in the 2000s to tackle the LGBTQ issue through both qualitative and quantitative methods. Qualitative research explored in-depth the intimate relationships within families: Rinaldi and Cappotto (2014) analysed the coming-out process in the "normal and traditional" Sicilian families, from a deviance point of view; Paterlini (2006) described, from a narrative point of view, the Italian way of "gay marriage"; Allegro (2006) focused on coming-out, describing the strategies of lesbian mothers who tell their children about their sexual orientation; Franchi and Selmi analysed the persisting heteronormativity of the Italian context against gay and lesbian parents (2018; 2020); Carnelli et al.

(2009), finally, addressed the issue of having and rearing children (De Rose and Marquette, 2011).

On the other hand, quantitative studies started to focus on the estimation of the LGB (and LGBT) population in Italy and of its proportion on the whole population (Istat, 2011). An integrated approach of qualitative and quantitative methods has appeared to be more appropriate and more flexible in dealing with the complexity characterizing the LGBTQ population(s) (Corbisiero, 2013; Monaco, 2019). That is the reason why, in Italy, an important contribution to this field, was made in mixed methods research (MMR) application, which uses and combines quantitative and qualitative approaches (Johnson 2007; Creswell, Plano Clarc 2007). In line with the MMR approach, some important works were carried out, such as by Barbagli and Colombo (2007), by Chiara Saraceno (2003) and "I am, I work" (2010), Lelleri (2006), Porrovecchio (2011) and Monaco (2019).

Although MMR increased the understanding of the LGBTQ population(s) - despite some limits, usually given by their local and circumscribed nature- quantitative studies, nationwide, have just been a few and mostly of pioneering and experimental nature: Fabris and Davis, in 1978, published "*Il mito del sesso*", the first report on sexual behaviour amongst Italian people, which for the very first time touched upon, from a quantitative point of view, the argument of homosexuality; "*Il sorriso di afrodite*" was edited in 1991 by Fiore and partnered by Ispes, and represents the first report on homosexual conditions in Italy.

Other suggested quantitative studies were conducted in the Neapolitan area in 2010 and 2015, respectively called "*Certe cose si fanno*" and "*Napoli DiverCity*". They were carried out and edited by Fabio Corbisiero, partnered by local Arcigay "Antinoo" and local institutions. The aim of these studies was mainly to observe and gain knowledge on how LGBTQ people live in the city and urban spaces, for socializing and for performing their gender and their sexuality. The last Neapolitan quantitative experience, not yet published, was conducted in 2018/2019 by the LGBT observatory of the Federico II University of Naples, which carried out a CAWI survey on 1600 individuals to investigate lifestyles, interpersonal relationships and daily life aspects of LGBT+ people.

In regards to massive statistical studies, in Italy, have been carried out by the Arcigay collaborating with the National Institute of Health (ISS)[3], in 2005, and by

[3] ISS's survey, called "Modidi", focused on health, safe sex behavior and perception of HIV among homosexuals in Italy, but giving also an insight about family formation and childbearing. It was

Istat, in 2011; this latter survey will be specifically presented in the following paragraphs.

Italian politics seems to be increasingly aware of LGBTQ civil rights claims, which led to the recognition of same-sex unions in Italy, in 2016. On the other hand, awareness of discrimination, self-representation and coming-out modalities are still in their dawning in Italy yet and data, about this phenomena, have to be provided to stakeholder and policymakers to encourage an adequate recognition rights process.

In the following paragraphs, the most relevant quantitative and mixed-method researches regarding the LGBTQ issue will be profoundly analyzed, mainly discussing the changes of Italian researchers' point of view on the methodologies used or that should be used for investigating lesbian, gay, bisexual, transgender and queer people.

4.2 The "modern homosexual" concept

Marzio Barbagli and Asher Colombo, both Sociology professor at the University of Bologna, in 2001 introduced the concept of modernity in the LGB debate. Their research aimed to obtain a 360-degree knowledge of men and women who love same-sex people, beyond every stereotype. In the early 2000s, Italian public opinion towards homosexuals was changing, along with the behaviours of LGB people: coming-out experience, the spread of meeting place and gay-tourism (Corbisiero, 2013), the elaboration of a political and cultural identity changed both the daily life of LGB people and their acceptance by the Italian society (Trappolin, 2008). To analyze all these fundamental aspects authors carried out a survey which involved a convenience sample of 3502 people, who were interviewed between 1995 and 1996. Besides, authors and their staff recorded 144 biographies among LGB people, to enrich quantitative data, come from the survey, with the direct experience of the people interviewed; the research employed also metadata coming from other surveys, ethnography and other qualitative methods, intending to give sociological robustness to the research.

---

carried out using a snowball sample, through a self-filled questionnaire, and collaborating with the main LGBT national associations. (De Rose and Marquette, 2011)

Barbagli and Colombo's results pointed out that homosexuality in Italy no longer reflects gender behaviour stereotypes in terms of a certain "dress code", physical attributes or attitudes: relationships in same-sex couples do not follow asymmetrical stereotypical sexual roles (first of all passive/active) and are more "egalitarian" and less eccentric than what stereotypes impose. Results showed that LGB no longer tends to meet up with other or new people anymore in clandestine situations, such as *buttuage* or gay cruising spots; instead, an interconnected network of gays and lesbians associations promote cultural, entertainment and socialization activities, which led to a higher community spirit and to a need of recognition of civil rights.

"Modern homosexuals" usually live as a couple in a cohabitant union, having an active sexual lifeless libertine than traditional negative stereotypes of lesbian and gays still suggested in Italian society: more than 50% of lesbians have been in a relationship. Among gays, 39.0% of those in the age group 18-24 were in a couple, at the time of the survey, and the percentage rises to 47% in the age range 35-39. Moreover, between the 35 and 40-year-olds, 1 out 5 gays and 1 out 3 lesbians live together with a same-sex partner, without being married. When it came to the issue of parenthood, research pointed out that 10% of homosexual men and 19% of homosexual women (over the age of 35) have children, mainly from a previous heterosexual relationship. Gender and age differences are quite pronounced on this argument because younger gays showed the highest percentage concerning their desire to have children. Barbagli and Colombo's data counters the misleading stereotype that LGB are less likely to have stable and durable relationships than the heterosexual population, and argue that, also from a sociological and demographic point of view, LGB must be considered as an integrated part of our society, avoiding every kind of discrimination.

The research provided by Barbagli and Colombo was really appreciated both by researchers and the LGBTQ community, because it filled the LGB void in the sociological debate. The limitation of this study, given by the purposeful sample on which the data are based, is, unfortunately, a constant of the quantitative studies on this field: interviewing people at meeting places for gays and lesbians can lead to a biased sample in terms of different age groups and social strata of the homosexual population as a whole, because these places are mainly attended by younger people and are more spread in bigger cities as opposed to smaller towns. It should be clarified that, in this context, analyzing an elusive population(s), the aim is not to make statistical inference nor is it to have a

perfectly representative sample, but to increase the knowledge about the targeted population. In this context, the implementation of qualitative methods fixes this limitation of quantitative ones and gives precious information on issues barely investigable by a survey, such as the first homosexual intercourse, death wish and the process of coming-out in households.

"*Omosessuali moderni*" rightfully became a milestone in the field of sociological studies on homosexuality and bisexuality in Italy, and an unavoidable bibliography reference for all the following researches (this thesis is an example of those), which share two main goals: including other sexual orientations and gender, such as queer and transgender people, and it observes the phenomena only through the "associationism" lens, without acknowledging the huge contribution of the LGBTQ associations to the debate.

4.3 Different from whom? The Saraceno's metropolitan research experience

In 2003 Chiara Saraceno, Alessandro Casiccia, Chiara Bertone and Paola Torrioni conduct a pioneering mixed methodology research to study the LGBT population(s), in Turin's metropolitan area.

Data analyzed came from both a survey and qualitative methods, such as in-depth interviews, focus groups and interviews of members of several associations. The author underlines the plurality of people included under the umbrella term LGBT, which try to cross the profoundly distinct concepts of sexual orientation and gender identity. This research proposed a procedure to recognize the sexual identity of a person, which was then employed also in other studies (Istat, 2011), based on three main dimensions: erotic attraction, sexual behaviours, and sexual experiences. Women resulted more fluid than men, and less inclined to define their sexuality statically: they experienced more attraction for other women, without the need of defining themselves necessarily as lesbian, in what was called the "uncertainty society". Saraceno's book also focused on the "first time", not intended as the first sexual encounter, but as the moment in which and how they first discovered an attraction to the same sex: at what age it happened and for whom. What emerged is a sense of inappropriateness, compared with the heterosexist traditional model, which lasts for most of adolescence and youth, leading to a temporary isolation period, mainly in the household. Chiara Saraceno is one of the most important European experts of

family Sociology and she brought attention to coming-out at home, to understand the family member (or members) to whom the non-heterosexual orientation is most often confided, which strategies are developed within the home, the family's reactions, and finally possible changes in the relationships after the revelation of one's sexual orientation, which represent the most common of bad behaviours towards LGBTQ people and which can lead to tragic suicide cases as well.

Discrimination is obviously the main problem for all people interviewed, but the problem is heightened for transgender people, who are not included in the Barbagli and Colombo research. In 2003, even in a big city like Turin, transgender people experienced strong difficulties in daily life, such as retrieving information about sex reassignment surgery, accessing the labour market, and establishing an affective or love relationship. Transgender people resulted as the most vulnerable victims of a cultural inadequacy of our society, which continues to propose old social schemes and marginalize those who are out of these (Romano, 2011; Papuli, 2019).

"*Diversi da chi?*", the original title of the research edited by Chiara Saraceno, presents how a metropolitan context like Turin's can be ghettoizing and marginalizing for LGBTQ people. The conclusion of this pioneering research could be generalized, even today, to the majority of Italian cities, because not enough has been done in terms of social policies, during the last seventeen years. On the other hand, the studies main weakness is to not include an analysis of the little towns, which represent almost all of the Italian municipalities. In these small cities, bad attitudes towards LGBTQ people are sadly widespread: coming-out is more difficult, violence, sadly episodes of mobbing and bullying are more frequent and there are no associations or listening centres, capable of giving social and psychological support to homotransphobia victims.

4.4 A set of surveys and not a single one. Istat choice for a SOGI-sensitive official statistics

In previous paragraphs, we saw that several national statistical institutes (NSIs) have introduced questions about sexual orientation and gender identity in the census and social surveys. In 2011, the Italian National Institute of Statistics (Istat) also surveyed "gender discrimination, sexual orientation and ethnic

belonging", partnered with the Italian Minister of Equal Opportunities, in which were introduced –for the first time in Italian official statistic- questions covering the respondents' sexual orientation and gender identity (SOGI).

The survey involved 7,725 Italians between 18 and 74 years old, interviewed with the CAPI technique, about every type of discrimination (gender, sexual, ethnic, etc.) and, for homosexual and bisexual people, about their coming-out experience and possible discriminations against them. 77% of interviewed people define themselves as heterosexual, 0.1% as transsexual, 4% chose the "other" option, 15.6% did not answer and 2.4% of the population declared to be homo-or bisexual, for a total estimation of about one million people; considering who, during their lives, fell (or are) in love with a same-sex person, or that have felt attraction or had sexual intercourse with a same-sex individual, the estimate rises to about 3 million people (6.7% of the population between 18 and 74 years old) with relevant socio-demographic differences between genders, geographical areas and age classes: who identifies herself/himself as homosexual is more men than women, more northerners than southerners, more younger than older people (Istat, 2011). Sexual orientation proved to be strongly correlated with the individual propensity of being victims of discrimination, in every social life phase: at school or university (24% of homosexuals were victims of discrimination, against 14.2% of heterosexuals), at work (22.1% against 12.7%) but not in the job-hunting phase (29.5% against 31.3%). On top of this coming-out data show that it was still difficult, for homosexuals and bisexuals people, to reveal their own orientation to their parents (20%), to their siblings (45.9%) and colleagues (55.7%). The survey, further pointed out a discomforting perception of the whole population towards LGBT people: 27% of the population do not feel the need to condemn discrimination against LGBT people, 25% thinks that homosexuality is an illness and 26% thinks that it is immoral.

Survey results stirred up several criticisms and objections, both about the results and methodology applied (Cafasso, 2017). As already said in paragraph 1.1.1, designing a sample survey representing LGBTQs could be a really complex task. Rarity and elusiveness of the population(s) could complicate the survey design, mainly from a sampling point of view: inadequate sample size, non-responses and misreporting in SOGI questions can lead to very biased results. On the other hand, Istat has recently switched to a sample-based permanent

census[4], from a traditional one, which makes it no longer possible to ask SOGI questions to the whole population. Anyway, the introduction of SOGI questions in the permanent census questionnaire is a hypothesis under consideration. Right now the only information collected on this field by the census are same-sex cohabitant couples, since 2011, and civil unions, since 2016.

As already seen at a European level, in paragraph 1.3.1, several European NSIs are abandoning the idea of estimating the LGBTQ population in proportion to the overall population, focusing solely on estimating good attitudes, like inclusivity measures, or homophobic behaviours towards LGBTQ.

Istat is also planning on putting into place a set of surveys, collaborating with the National Anti-Racial Discrimination Office (UNAR[5], from now on) to give a cognitive framework about LGBT people in the workforce and a magnitude measure of the discrimination phenomena against them. The set of surveys are oriented both towards individuals, enterprises or other stakeholders.

Three individual surveys will be carried out between 2019 and 2021:

1. census of people who have got in a civil union, since 2016, based on the administrative register;

[4] Since October 2018, Istat has been yearly conducting a sample survey by collecting the main characteristics of the Italian resident population and its social and economic conditions at national, regional and local levels. The permanent census of the Population and Housing does not involve all Italian households anymore, but only a sample of them every year: about 1,400,000 resident households located in 2,800 Italian municipalities. Through integrating information from statistical sample surveys and data from administrative sources, the permanent census yearly provides data representing the entire population, while reducing costs and response burden. (Istat, 2018)

[5] UNAR is the office destined by the Italian State to guarantee the population the right of an equal treatment, regardless their ethnicity or race, age, religious belief, sexual orientation, gender identity or disability. The office manages discrimination cases, studies possible solutions and promotes the culture of respecting human rights and equal opportunities.

2. a survey on LGBT individuals selected through an advanced snowball sampling method, called *Respondent Driven Sampling*[6] (RDS), starting with association members;
3. focus survey on transgender people, starting with association members and counselling centres users;

Enterprises will instead be involved in the set of surveys asking them if any diversity management policies have been applied to promote inclusivity and to discourage any discrimination based on sexual orientation and/or gender identity.

Choosing to conduct a set of surveys and not a single one, Istat rough out a "road map" for SOGI-sensitive official statistics, which is not interested in just the number of LGBT people anymore, but which wants to know and to understand the phenomena linked to being LGBTQ in our society, supporting policymakers to promote efficient diversity inclusion policies (De Rosa and Inglese, 2018; UNAR-Istat, 2020).

Future prospects of this research field in Italy have been clearly described by Elisabetta Ruspini, in 2013, who wishes for more attentive, sensitive, and accurate research on LGBTQ subjects. Sociology, statistics and all other related fields have to read with professionalism and critical thinking social phenomena and go further into a simplistic point of view and stereotyped scenarios. Moreover, LGBTQ quantitative research has to become continuous and longitudinal to show how phenomena change across demographic cohorts and generations. For example, it will be interesting to understand how Millennials will perform their gender and their sexual orientation in a completely different way from what previous generations did. Finally, on the methodological side, LGBTQ research has to continue to focus on sampling and recruitment methods, to represent the LGBTQ population(s) as correctly as possible, exploiting MMR as the right tool to mediate between the need of quantifying and that of deep understanding.

[6] *Respondent-driven sampling* is a chain-referral sampling method where participants recommend other target people they know to interview in turn, as it happens in snowballing methods. The main difference between RDS and simple snowballing is that the first one is mathematically oriented to add an element of randomness to the sample, launching many different snowballs, and to employ a mixture of non-probabilistic and probabilistic methods (Heckathorn, 1997)

LGBTQ surveys and studies reviews: concluding remarks

In previous paragraphs, we have gone through many surveys and quantitative studies, conducted all around the world, about LGBTQ individual populations and as a whole. The studies reviewed show us several recurring themes, which can give important suggestions -summarized in the bulleted list below- that should allow social research to undertake much more significant steps in this field.

- What we talk about when we talk about LGBTQ? Definitions are really tricky in this field, but defining a proper operative and functional definition is essential for fruitful research. Evidently, categorizing social gender identity, gender behaviour and sexual orientation introduce significant considerations, both from an epistemological and methodological point of view. Every categorization can be in fact considered as heavy structuralism to avoid. Both sex and gender are becoming increasingly recognized by people as having both separate dimensions and more than two possibilities. Social researchers, therefore, have to design sexual orientation and gender identity (SOGI) questions referring to gender and queer studies, to give them as much freedom as possible of defining themselves to people and to mediate between the complexity of reality and the feasibility of its investigation.
- The growing interest of governments, ministries and institutions about the LGBTQ population(s) leads to a higher need for quantitative data on this issue. Sociology, statistics and many other disciplines are jointly called to fill this lack of knowledge, dealing with all the difficulties of investigating a hidden population. Moreover, the research sometimes has to face a sort of "political filibuster", which tries to get in the way of LGBTQ civil rights recognition, hindering research and surveys on this issue, as it happened, for example, in Nepal (paragraph 2.3).
- In the last few years the number of LGBTQ people, or its proportion on the whole population, has lost its importance in the eye of governments,

ministries and policymakers, mainly in Europe. Institutions principally need data about bad behaviour towards LGBTQ for addressing their social policies (FADA, 2017). That is how LGBTQ research has changed, abandoning the aim of finding a definitive number of people and starting to quantify the phenomena related to those people. On the other hand, private opinion polls and market research are still strongly interested in quantifying the size of the LGBTQ population(s), considered an unexplored market niche (IFOP, 2017; Lam, 2016).

- The main outline area for future study should be the periodicity, because only a panel approach can give a longitudinal perspective on all those investigated phenomena involving the LGBTQ population(s). Ruspini (in Corbisiero et al., 2013) focused, for example, on the comparison between LGBTQ Millennials and their previous generations. This kind of evaluation could outline how performing gender and sexual orientation has changed through generations, also and especially about the use of digital technologies. Gender Studies and digitalization are two theoretical frameworks more and more interconnected, which have frequently underlined that gender is neither something that "one has" nor something that "one is", but something that "one does" and "says" (Cozza, 2008).
- To solve the problem of lack of knowledge about LGBTQ, researchers first have to fix all the significant methodological flaws that we have seen thus far, such as sampling bias, survey modes and independence. It is not difficult to observe that till now, as Barbagli and Colombo highlighted in Omosessuali moderni (2007), gay associations (mainly Arcigay in Italy) represented the main informative resource on the LGBTQ community, promoting and partnering the main studies. Of course, the information provided has been extremely useful and informative, but social knowledge should go beyond the "*associationism*" and associative activities. Studying LGBTQ people from their associations' point of view can lead to deep but biased knowledge, because the target population(s) can be wider than the people who belong to the associations. For example, associations are more localized in big cities than in small towns and their members are clearly more involved in the process of recognition of civic, cultural and political

rights than most of the other LGBTQ people who, for several reasons, cannot belong to any association.

- Italian experience on this issue, although with its gaps, has shown that at the forefront of social research in this sector. Both from a statistical and from a sociological point of view, data regarding the LGBTQ population(s) have been produced, discussed and improved. Moreover, in Italian academic research, the use of Mixed-Methods Research has demonstrated that qualitative methods could be functional and necessary for quantitative ones. An example, strongly recommended also by international institutions, is the employment of interviews, focus groups or story-telling, for drafting the questionnaire. Investigating which aspects of daily life to ask and how, taking into account the questions sensitivity and the complexity of gender identity and sexual orientation issues (Ruspini, 2014).

PART 2

METHODOLOGIES USED AND THEORETICAL IMPLICATION

A change in Quantity also entails a change in Quality

FRIEDRICH ENGELS - Anti-Dühring (1877)

5 Why use big data for studying LGBTQ population(s)

Quantitative studies review shows that most of the survey carried out on LGBTQ population(s) has been partnered, promoted or incentivized by associations. Their contribution has been crucial for the knowledge of the LGBTQ community. Consider, for example, all those studies on transgender people, hidden and invisible to institutions, defenceless victims of many stereotypes, and hard to reach also for the most zealous social researcher.

LGBTQ associations have represented, for decades, the main informative source on the community, as already said since the introduction. Associations have provided useful information about LGBTQ, but social knowledge should go beyond "associationism". Studying LGBTQ people from associations' point of view can lead to a deep but biased knowledge because the ideal target population(s) can be quite wider than the people who belong to the associations. For example, associations are more localized in big cities than in small towns and their members are clearly more involved in the process of recognition of civic, cultural and political rights than most of the other LGBTQ people who, for several reasons, cannot belong to any association.

The concept of social research about LGBTQ beyond the associationism encouraged to use of big data, coming from social media, in this research project, introducing an innovative approach in the field.

5.1 Instagram users data for the "Over The Rainbow" survey

The above innovative approach has been applied on Instagram, one of the most popular and used social media worldwide, which allows millions of users to share their pictures and to describe themselves by using keywords, called *hashtags*, in the tagline. Social media are almost "closed" and they work as "black boxes", but using web-scraping software it is possible to quickly download all posts containing a specific hashtag, to design, plan and finally conduct a survey on a massive number of people who likely come out on Instagram. Following the web-scraping users' data collection, a survey through Instagram has been

conducted, sending them a questionnaire about the most unknown and interesting aspects of the LGBTQ population(s).

In this way, Instagram has become our *multi-purpose research tool* which has allowed us to carry out every data collection phase: through web-scraping Instagram, we found LGBTQ people, through Instagram direct messages (DM) they were contacted, and, finally, through an Instagram page they were reassured about privacy, accuracy and aim of the project.

Instagram showed very high potential in social research, for many reasons:

- it is widely used, which is more than we can say for many other social media, first and foremost Twitter, which has changed during the last few years, becoming a specific tool for politicians, journalists and communication experts, cutting normal users off.
- it is globally transverse, because it crosses nationalities, ethnicities, social statuses, genders and sexual orientations. This is not completely true for age classes, because young people are more likely to use this social network than others (i.e. Facebook, which has been left to adults), but this disproportion is mitigated by the huge number of users as well as it is unavoidable, working on this field.
- it is not completely closed because it allows the use of Application Programming Interfaces, commonly called APIs[7]. By the way, the Instagram API policy has changed in the last years, converging towards Facebook's one, increasingly closed and self-referential.
- Instagram's direct messaging system allows to directly contact users also without following them (or being followed by), which is more than we can say for many other social media, first and foremost LinkedIn.

Instagram, and data "scraped" from it (that will be specifically discussed in paragraph 5.6) allow to trace all those users who use, or have used it in the recent past, the most common LGBTQ hashtags to tag and index their picture, by argument and content. The use of big data in this research project remains circumscribed to the data collection phase, without implications in the most futuristic methods of digital social science, such as algorithms, machine learning,

[7] An API is a part of the website server that receives requests and sends responses, allowing users to do certain acts.

visual analysis, etc. However, Instagram has been used as a research tool, also and foremost in the phases following the data collection, such as to contact users, to solicit them, to reassure and inform them, mainly about privacy conditions of the survey. In this project, thus, Instagram represents both an empirical research tool and a new social space: subjectifying, self-expressive and free, though subject to its netiquettes (Snee et al., 2016).

Working on social media necessarily includes theoretical topics such as *digitalization* and all those relationships between humans and web technology (Mäkitalo, 2020). Think about smartphones, tablets, smartwatches, etc. All these devices have had a deep influence on the everyday life of many people spread all over the planet: the communication and the data flow between people have never been so quick and so reliable. Sociology wasn't indifferent to this epochal change: social researchers tried to explain it, focusing on the impact, development and use of digital technologies and their incorporation into social worlds. This sociological field started to be called Digital Sociology (DS) in 2009, but only in 2015 did this line of investigation of sociology became official with the publication of Lupton's manual, indeed named "Digital Sociology".

Moreover, using big data, and relying on new tools, makes social researchers aware of new epistemological problems: scientific knowledge limits are changing, and with them both the reality in which we live and the methodology to investigate it (Amaturo and Aragona, 2019).

5.2 Big data and social research: methodological and theoretical implications of new tools

Social sciences often use statistical tools to argue their thesis from a quantitative point of view. However, in the last few years, statistics have increasingly exploited all those *footprints* that all of us daily leave on the web, commonly called big data. In this way, the gap between big data and social sciences has been collaterally shortened. However, getting close to the digital world and social research involves a discussion about the nature of both of them. This discussion inevitably brings up the "coming crisis of empirical sociology" (Savage e Burrows, 2007) and the unfounded fear that, in the foreseeable future, data-scientists could replace sociologist, anthropologist, economists and even psychologists.

But, let's start from the beginning to better understand the situation: big data consists of an informative dataset, remarkable in terms of volume, velocity and variety[8] (Laney, 2001) whose size require ad-hoc technologies, both of storage and elaboration, as well as specific analytic methods for extrapolating the data intrinsic informative value. Big data include all those traces (the footprints, mentioned above) that we leave on the internet, willingly or unwillingly, creating a huge amount of data, different and continuously in update: tweets, pictures, videos, self-tracking data (Lupton, 2015; 2013), GPS signals, emoticons, scanned QR codes, cookies, etc.

Moreover, for someone, big data not only represent a powerful tool, but a real *paradigm shift* (Kitchin, 2014). Big data implications are considered a new form of empiricism, which could mark "the end of theory", as Chris Anderson said in 2008, on Wired, arousing amazement. From this point of view, would be obsolete and useless, for a scientist, starting from hypothesis, verify them through experiments, collecting data and, finally, achieve a theory. Today everyone could directly start its study from data. This approach, known as *data-driven*, lets algorithms run on available data, until they find the most likely theory to describe a phenomenon (Coletta and Kitchin, 2017). This kind of process would represent a new Copernican revolution on the scientific method, in which the theory is not the starting point anymore, but, on the contrary, the final achievement of the research process.

Epistemological implications of this process lead Prensky (2009) to coin the term "Homo Sapiens Digital", a new kind of person whose thinking and wisdom[9] are given by a symbiosis of the human brain and its digital enhancements. This new race lives in a high-speed society (Rosa, 2010), in which the speed of communication has quickly changed everything, also the rules of knowledge. This epistemological change downgrades the *causation* importance to a strict and

[8] In a famous and very cited article, Douglas Laney, in 2001, defines for the very first time big data according the 3V definitions: Volume, which indicates the massive amount of data, Velocity, which means data are in real-time updating and Variety, which means they could be text, images, sounds and they have to be handled to be structured and used. Some authors add other V characteristic to the initial definitions, such as Variability, Virality, Veridicity and Value, but they could be thought as declinations of the starting 3V characteristics.

[9] Prensky introduced in 2001 the concept of *digital wisdom*, a human quality which develops as a result of the empowerment that the natural human skills can receive through a creative and clever use of digital technologies.

unverifiable theoretical condition. A significant *correlation*, rather, would be more than enough to formulate a theory in a big data-driven approach.

Obviously, this is just a scenario and it's not necessarily said that Kitchin's paradigm shift is going to happen. But, in social sciences, some considerations have to be clarified. This new empiricism represents for many scientists the positivist (and post-positivist) full accomplishment: the realization of a project of social control and social prediction made possible by the incalculable amount of available data. On the other side, it is not said that the advent of big data will delete all of the social sciences' problems about quantifying humans behaviour, in one clean slate.

But rather, working *on* big data could help solve some methodological problems that characterize quantitative studies, such as social desirability, interviewer effect, the economic sustainability of surveys or (as it happened in the present research project) the absence of a frame from which to select people, belonging to a hidden population, to interview.

Working *on* big data is not equal to working *with* big data. In the first case, the social researcher takes advantage of big data use without negotiating his/her key role in the process. Working with big data, instead, means collaborating in such a way with them. This collaboration involves acquiring a new consciousness about the social context in which we interact: a context of hyper-surveillance (Lyon, 2014) and oversharing (Agger, 2015). In other words, in both cases, the discussion about big data cannot be exhausted just talking about data size or of the informatics procedures used to handle them. For instance, privacy online and "rights to be forgotten" have to be clearly faced and solved (Lombi, 2015).

Concluding, we can say that sociology has recognized the potentiality of digital media and of big data; sometimes adopting perhaps a too pessimistic point of view, like in critical digital sociology (Lyon, 2014), sometimes adopting a too "idyllic" point of view, thinking about new digital media as the answer to all the questions that traditional sociology posed. Both these two points of view are way too drastic and could be misleading (Marres, 2012). Innovative data sources, mainly social media, have brought an informative richness (Daas, 2014) that social researchers have to use for their studies. Although they have to be careful on how social media users represent themselves, in a kind of digital "self-disclosure", and how they construct an online profile and for what aim those profiles are used for (Rizzo, 2018).

The progressive use of big data in social research needs to be sustained by plural and inclusive epistemological framework, that allows to include new data and tools within the different paradigmatic traditions that coexist in social sciences. In order to affirm this "digital epistemology" researchers have also to adopt a new methodological point of view, trying to exploit the advantages that digital techniques entail, alongside digital methods with traditional ones, both qualitative and quantitative (Amaturo and Aragona, 2019).

5.3 Affirming and performing gender identity and sexual orientation on social media

When Anthony Giddens wrote, in 1992, "The Transformation of Intimacy: Sexuality, Love, and Eroticism in Modern Societies", he had understood that profound changes occurred to human relationships, far beyond the private sphere. What he didn't expect was that many other revolutionary changes would have happened in the following decades, with the World Wide Web advent.

The Internet has offered new important opportunities for action and aggregation, allowing for the "showcasing" of diverse sexuality. LGBTQs are actively engaging in social media, which has offered provisions to a minority, involving user interactivity and self-production-oriented modes. LGBTQ people have taken several online advantages to test their identities within virtual environments, and to make connections previously unimagined. This may be evident in the provision of coming-out, dating, blogging, affirming themselves, performing their own gender identity and sexual orientations, associating and mobilizing political ideologies. This new electronic age makes the historical isolation and rejection of sexual diversity relatively distant for all the LGBTQ community (Pullen and Cooper, 2010).

In the beginning forums and chats existed, which opened the door of the "self-games, body play and to cybersex" (Waskul, 2003). They were pioneering because allowed people to meet up, to discuss taboos without any inhibitions or restrictions, providing a safe place (also if virtual) to the community, mainly to the part which didn't live its condition proudly and openly. LGBTQ chats, nevertheless, had the same old problems of gay bars: they were just for gay men. Not even lesbians or bisexual were included. And it can generate a sort of "ghettoisation", given by the lack of admixture with other people, firstly

heterosexual friends. Furthermore, chatrooms were usually used to find sexual partners and so it became frustrating for all those people looking for friendship, support or an affective relationship (Barbagli and Colombo, 2007).

The legacy of chats and forums was continued by social media, which tried to integrate the virtual LGBTQ community with targeted actions for making their websites as inclusive and safe as possible. The intent of LGBTQ people on social media was no longer only to date. The goal had become to also find out information, meet new people and keep in touch with old friends, creating a web community capable of self-identifying, mainly in time of need.

This new main goal can be deconstructed in the following four aspects:

1. SOCIALIZATION: LGBTQ people join social media to seek out people of the same sexual orientation, both for a sense of belonging and to socialize. Websites become networked publics which are simultaneously "*the space constructed through networked technologies and the imagined community that emerges as a result of the intersection of people, technology, and practice*". In these publics, people meet up, flirt, have relationships or affairs, fall in love and break up (Boyd, 2008).

   Heterosexual people do the same, of course, but with the big difference that they can also have a non-sexual-discriminating off-line life. On the contrary, for non-heterosexual young people and/or small-town citizens (an intersectional approach, to cross different discrimination conditions, seems to be more appropriate) online tools are not a choice for socializing with other LGBTQ people. For instance, LGBTQ youth, compared to youth in general, have limited use of public spaces or are limited in their self-identity-expression in public such as at school, in their spare time, or sports.

   Social media have become a socialization catalyst. An important role that has been played until now by associations (Ross et al., 2014). Social networking sites are part of lived experiences, not separate from them, able to introduce people and to redefine the (virtual or real) locations of engagement and signification for LGBTQ people (MacIntosh and Bryson, 2007).

2. INFORMATION: Social media provide every kind of information we need. Sometimes we feel over-exposed too, and it is a reality that some information could be really misleading (conspiracy theories, fake news, deceptive advertising, etc.). The LGBTQ community exploited this informative richness for solving the problem of having information about their rights (Zhang et al.,2009). Moreover, by understanding the background of the use of social networking in the LGBT community, we can understand why the social network becomes a very important role in the dissemination of information about this community. The clearest example in this field is the transgender case: trans people took advantage of social media not only for meeting but also to search and share information, difficult to find before, about transitioning, both from a surgical and a psychological point of view. In many Facebook groups, it is possible to ask for tips and suggestions about transition and on YouTube, there are several channels about the same subject.

3. COMMUNITY: Generically, social media is potentially a very useful tool for all those organizations that concern themselves with political action for social justice and support, such as LGBTQ community groups (Ross et al., 2014). Social networking sites, in particular, offer the possibility of communicating with multiple constituencies and can be used to publicize services, campaigns, engage potential sponsors, create peer networks, as well as communicate directly with existing and new service users (Jenzen and Karl, 2014).

4. SELF-REPRESENTATION Social media incorporate plenty of interactive media, published by professionals or average users, that everybody can comment or share. Users of social media become producers of a social artefact which is the online persona, an online expression of the self (Rizzo, 2018). Users self-revelation is not just an act of sharing personal details, but also an active construction of one's perception of who one is (Cooper and Dzara, 2010). It often involves a re-mediation process of online and offline media. Self-expression on social media has also changed over the years, passing from the personalization of a profile (like MySpace or MSNspace pages) to a form of self-expression that prioritises social relations and being

networked (like what happens on Facebook, Twitter and Instagram). In this way, new forms of social media are incorporated into everyday life, complicating some practices and reinforcing others (Lombi, 2015). On the other hand, new technologies reshape public life, but users' engagement (especially younger ones') also reconfigures the technology itself, generating an interchange between humans and technology (Boyd, 2008).

In summary, social media has been playing, since early 2010, an important role in aggregating LGBTQs and allowing them to organize community advocacy all over the world, to further their impact and to share information. In this way, social media have become the perfect tools to help LGBTQ people to stand together and demand justice when an unethical act occurs to a specific individual. Facebook, Twitter, Instagram, and many others, help normalize this community, making it finally visible to outsiders who may not have much exposure to LGBT people in their day-to-day lives.

Social media involve negative aspects too, and bad behaviours towards the LGBTQ community can happen also online, giving rise to cyber-bullying. However, the socio-political climate of our societies, during this informational capitalism age (Castells, 2014), is becoming more progressive on LGBTQ issues, which suggests that the positive effects of social media are edging out the negative ones (Wallace, 2019).

5.4 Rainbow social media: different networked publics for different users

As previously shown, social media have become an essential part of people's social, romantic, and sexual lives. Social network websites are important for meeting, dating, breaking up, but they also represent important information resources about sexual health and identities (Pascoe, 2011).

The way the LGBTQ community exploits social media potentiality is not equally developed among different providers. The three main social network websites (Facebook, Twitter, and Instagram) have different goals, managers, corporate policy, and target users. Each of them provides different advantages

and disadvantages to LGBTQs, and also to social researchers who want to study them.

Facebook has proved to be more than sensitive to LGBTQ community claims. The star example was the rainbow filter to personalize the profile picture when in the USA, in 2015, the Supreme Court approved same-sex marriage. This kind of gestures drew strong criticism on Zuckerberg's enterprise, both from the most conservative users and from a part of the LGBTQ community too, which labelled these gestures as merely profit-driven. This diatribe is motivated by the advertisements system used by Facebook to reach a specific audience, targeted according to its characteristics, such as geographical location, gender, age, profession, relationship status, interests, music and possibly sexual orientation as well. Considering this viewpoint, applying the rainbow filter on a profile picture can be considered one of those small data palatable for marketers and investors.

Furthermore, LGBTQs deplored Facebook for its "real name policy", the principle of using ones real name and surname, instead of nicknames to avoid fake accounts. Sometimes this policy may result in discriminatory and dangerous, mainly in those countries where homosexuality is still illegal, and a nickname can allow people to express themselves through useful social artefacts (Pinch and Bijker,1984).

Facebook, nevertheless, has impressive numbers which represent the strength of this social network: on Facebook, there are more than 76'000 groups concerning LGBTQs, where more than 7'500 LGBTQ events have been organized, involving over 1.5 million users, 0.87‰ of Facebook's total active users daily[10] (Dara, 2017). Zuckerberg's social media, although a huge user number, is not a networked public for conducting social research: its API does not allow any non-commercial query, most profiles are private and groups closed, needing an invitation or approval to join. Facebook became very sensitive to the issue of privacy after the Cambridge Analytica data breach, in 2018, when Facebook leaked millions of users' data, then used it for political advertising.

Twitter has been the favoured communication channel of the most important LGBTQ awareness campaigns all over the world, such as #ItGetsBetter, following

[10] Facebook daily active users (DAUs) were 1.73 billion on average in March 2020, an increase of 11% year-over-year. Facebook monthly active users (MAUs) were 2.60 billion as of March 31, 2020, an increase of 10% year-over-year

the suicides of two young gay boys victims of bullying, #lovewins, tweeted by Barack Obama after the approval by the American supreme court of same-sex marriage, and #dumpStoli, which aimed to boycott the Russian vodka Stolichnaya, indirectly owned by the Kremlin, which has always been accused of inciting homophobia.

On the other hand, looking for LGBTQ-related issues, on Twitter, was not easy. This social media, until very recently (January 2020), allowed for the free publication of porn material. Because of this when searching, #gay #lesbian or #transgender many misleading results were given. Besides, these results confirmed the unresolved stigma regarding words such as lesbian or trans, not used for pointing out a sexual orientation but just for tagging pornographic material.

Twitter is also not a public network in which to conduct social research because its users' number is down: in 2019 it lost 4 millions users compared to 2018. Twitter has been an excellent social media for years, but mainly for politicians, journalists, insiders, experts and professionals of the most diverse topics. But now it seems to be consulted mainly by external users, who tap into tweets without logging in; but this is not actually the case with LGBTQs, for which information on Twitter is few and often misleading.

Still talking about social media numbers, **Instagram** is certainly one of the fastest-growing Internet social media on the Web. In 2019 it counted 1 billion monthly active users, half of them interact through stories and posts once a day, 35% of them several times a day. Its catchment area is public (not many accounts are closed) and balanced, according to socio-demographic user characteristics: on Instagram, the proportion of men to women is 44:56, the proportion of users between the ages 18 and 34 is higher than on any other social media network, but six in ten online adults have Instagram accounts. As a matter of fact, Instagram is also very popular in all the countries in which Facebook has been banned or hindered, such as Russia, Eastern Europe and Central Asia (Omnicore, 2020). The LGBTQ community has enjoyed friendly Instagram's milieu since its inception. Instagram has been used uniquely and differently compared to all other social media: while Facebook promotes socialization, through events, the web community, through groups and fan pages, and information (Twitter too, focuses on this last goal), Instagram is completely focused on the self-representation of its users. Already in early 2010, many LGBTQs understood, the potentiality of a tool that allowed them to represent themselves, their own gender identity, their

own sexual orientation, through social artefacts and accounts not necessarily personal. Consider, for example, all the Instagram accounts of drag queens, in which there appear pictures of the male-cisgender version of them (often doing make-up), and sometimes the link to their male-cisgender account. And, vice versa, the account of men in which they compare pictures of themselves in their drag queen version, tagging the drag queen account. One person, two gender identities. One person, two accounts. Nothing is more gender theory-oriented than the possibility of being whoever you want, whenever you want.

Due to Facebook's real name policy, this kind of account management is not possible, while on Twitter it would be off-target, being an incubator of news, mainly specific ones. Instagram allows greater freedom in terms of self-representation and, inevitably, of their own gender identity and sexual orientation and their performance in daily life.

Before being acquired by Zuckerberg's empire (in 2012), Instagram emulated all the best and popular aspects of Twitter: it indirectly pushed its users to open their accounts, making them accessible also by non-followers and even by those who did not have an Instagram account at all (using the on-line version). Moreover, Instagram introduced post indexing, through hashtags. Indexing posts makes them easily findable by other users interested in a certain topic. This can be achieved through the simple and effective use of a hashtag, a word (preceded by the characteristic symbol #) used in the tagline, that allows users to find all the posts tagged using that word. Developed on the IRC[11] system during the 80s, the use of hashtags becomes viral with Twitter and is then adopted by Instagram as well, to make posted pictures tagged according to their content or other kinds of information related to the photo. A hashtag can be a feeling (#happiness), a place (#igersroma), a movement (#metoo), a person (#cr7) or information about the user who posted the photo, such as job (#nurselife), appearance and looks (#skinny) or sexual orientation.

From this suggestion, came the idea of finding LGBTQs on Instagram, to reach a wide catchment area, finally independent by associations, and just composed by people who freely self-represent and perform their own gender

[11] Internet Relay Chat: a client software used to "communicate, share, play or work" with others on IRC networks, broadly used in 90s and 2000s all over the world

identity and sexual orientation or, at least, people who support LGBTQ community in rights claims.

5.5 Using Instagram hashtags in social research

The previous paragraph showed that Instagram's characteristics facilitate its use in social research, for many reasons, such as its popularity, its ample usage, but it is not as close nit as Facebook, as its indexing system, it uses hashtags but they are not misleading like on Twitter, its non-compliance with any real name policy, which precludes the gender identity on the social media, and finally its well-balanced user base, in terms of gender, age, nationality, social background, etc.

Instagram has slowly started to be used in different social research fields (Munk et al., 2016; Chung et al., 2016), but usually by paying more attention to pictures, rather than to their captions. On the contrary, taglines could be really informative. For instance, finding the most common hashtag used to describe and tag pictures of users belonging, or supporting, the LGBTQ community can lead to obtaining a wide frame of users to interview for surveys. Different hashtags identify different people, but we were interested in surveying the whole LGBTQ people so we have to be as inclusive as possible: in June 2018, Instagram celebrated Pride Month, colouring hashtags as a rainbow[12], identified by the GLAAD (an important international organization at the forefront LGBT+ rights), as the hashtags that are often used in the LGBT+ community. Some of these hashtags can identify many components of the population(s) and, used in the tagline of a picture, they could likely describe the sexual orientation of the user.

In the following table the order of magnitude of the posts containing a certain hashtag, all over the world and of those referred to Italy, is reported:

[12] #pride #gay #pride2018 #lesbian #bisexual #trans #queer #asexual #loveislove #instagay #lgbtpride #genderqueer #intersex #gaypride #transgender

Tab. 1. - Number of Instagram posts containing hashtag (up to 20th September 2018)

Hashtag	*n. posts*	*Hashtag*	*n. posts*
#gay	59'700'000	#gayitalia	228'000
#lesbian	15'300'000	#lesbianitaly	1'687
#bisexual	7'100'000	#bisexitalia	331
#transgender	7'000'000	#transitalia	849
#queer	6'500'000	#queeritaly	113
#loveislove	13'000'000	#loveisloveitaly	6
#instagay	30'600'000	#instagayitalia	2339
#lgbt	22'400'000	#lgbtitalia	26'100
#lgbtq	7'600'000	#lgbtqitaly	210

The volume of the number of posts, coming from all over the world, is massive and also the number of posts tagged with an Italy-referred hashtag is considerable. The different declinations of the LGBTQ population are not represented in the same way, and there is a strong over-representation of gay people, rather than of lesbians or bisexuals. Transgender people (numerically the smallest component of the LGBTQ) on the contrary are well represented, as far as Instagram is concerned.

Wordcloud is a visual representation of text data commonly used to represent the link between words (or hashtags), and their importance in a text, a web page or a file. The size of words that form the cloud is directly proportional to their use, in our case to the number of posts.

The following wordcloud represents hashtags that appear when the main one (#gay) is mentioned, and their size proportionate to the main one:

Img. 1 – Hashtags Wordcloud

The wordcloud shows that in posts containing the main hashtag *#gay*, several other hashtags are often used. Most of them are other sexual orientations (such as #lesbian, #bisexual, #asexual, or #pansexual), gender identities (#nonbinary, #transgender, #trans, #ftm, #genderfluid), words to endorse membership in the community (#lgbtq, #lgbtpride, #pride, #loveislove), or in the web-community (#instagay). Sometimes instead, hashtags used together with #gay, are just frequent off-topics, such as #anime, #cute, #art, #instagood. The tag #love is also usually present when other ones are used, as if to say "love has no other labels".

In table 1, the intention of reporting the number of posts containing Italy-referred hashtags is to understand whether conducting a country based survey, with a questionnaire in Italian, is a reliable measure. In fact, from a data-collection point of view, it is definitively easier, mainly to avoid the language barrier and problems in translating a questionnaire.

Using a program able to automatically download posts, containing a certain hashtag, from Instagram can generate a sampling frame, in other words, a list of Instagram's users to contact and to interview in a direct, cheap and careful way. These types of programs are usually called *web-scraping*. They allow us to capture information stored on a website, or (as in our case) on social media.

5.6 Breaking into the social media black box: web-scraping Instagram

As we have already said, a relevant part of our social interactions and personal behaviours are conducted online, mainly on social media, and thus captured digitally. A social network website works through algorithms and computational structures that we don't know, or at least not all of them. Communicating with social media is never a direct process and an intermediary is always required: this intermediary is the API (already mentioned in paragraph 5.1), which collects supply and demand of data between the website servers and some users, including web-scrapers. That's the reason why we refer to social media as a black box. From a researcher's perspective, into the black box, there is a huge amount of textual and visual user-generated content, which can offer very interesting insights. But, in order to work with such large sets of data, and to organize them in a usable form, computer programs become necessary. These programs commonly are referred to as *web-scraping* (Schroeder, 2018; Denny, 2017).

Web-scraping is a set of digital computer techniques whose purpose is to extrapolate information stored, and usually publicly available, on a web page. This information can assume many data types (files, texts, pictures, videos, audios, etc.) so, once collected, data have to be organized into a usable format, usually a table, more friendly for social research aims. Web-scraping is generally implemented through a script, usually written in an open-source programming language (R or Python), but for the most basic use also software applications can be employed[13].

The black box which we broke into was Instagram. An R package, called InstaCrawlR[14], was used to do it. InstaCrawlR is a collection of scripts (attached in Appendix R), which downloads the most recent posts, for any specified

[13] Most common webscraping software applications are *Import.io, Octoparse, Google Spreadsheets, ScraperApi, ScrapingBee*. Specifically for web-scraping social network websites *Tagsleuth* has to be mentioned.

[14] The R package InstaCrawlR wasdeveloped by Jonas Schroeder, at the University of Mannheim, which designed this package for marketing research purposes. After an Instagram's API change we collaborate for updating the package and to fix some script errors, as describe in his Medium's article (see Appendix R)
(https://medium.com/@jonas.schroeder1991/update-instacrawlr-still-crawling-6500cd376ea3)

hashtag, and exports a CSV file, containing any information about the post, such as user profile ID, post's URLs, post's text, mentions to other users, related hashtags from post text, and the number of likes and comments to the post. The structure of the CSV output is illustrated in the table below.

Tab. 2. – CSV file structure. A subset of posts containing the hashtag #transitalia

(*) User name	*post URL*	*post text*	*(*) post mentions*	*post hashtags*	*(*) post date*	*post likes*	*post comment*
@xxx	https://www...	Il coming out, …	@yyy	#lesbianitaly	d/m/y	46	2
@xxx	https://www...	Beauty! …		#lgbtitalia …	d/m/y	25	0
@xxx	https://www...	#goldenhour …		#goldenhour	d/m/y	53	2
@xxx	https://www...	... That Spirit …		#usnavyseals	d/m/y	18	0
@xxx	https://www...	#gay #lgbtq …		#lemonade …	d/m/y	25	0
@xxx	https://www...	#transitalia …		#transitalia …	d/m/y	41	0
@xxx	https://www...	Questa immagine…		#loveislove …	d/m/y	49	1
@xxx	https://www...	#spuntidiriflessioni	@yyy	#spuntidirifles sioni	d/m/y	255	6
@xxx	https://www...	i c o n i c a		#jodiefoster …	d/m/y	39	1
@xxx	https://www...	//@yyy	@yyy	#lgbtqita …	d/m/y	64	0
@xxx	https://www...	Con #omocausto		#omocausto …	d/m/y	52	0
@xxx	https://www...	HAPPY BIRTHDAY		#mybirthday	d/m/y	16	1

(*) columns marked has been anonymized, to hide usernames and date

The script was run on the most common hashtags, Italy-referred, which can represent all the components of the LGBTQ population(s). In the table below the number of posts, tagged with a certain hashtag, and the number of users who used the hashtag, is reported.

Tab. 3. - Number of downloaded Instagram posts, and the corresponding number of users, containing LGBTQ Italy-referred hashtags. (20th March 2019)

Hashtag	*n. posts*	*n.users*
#gayitalia	27'338	2'730
#gayitaliani	3'257	448
#queeritaly	412	50
#lesbicaitalia	511	51
#bisexitalia	616	39
#transitalia	1'590	225
#lesbicaitaliana	538	79
#lesbicheitalia	693	109
#lgbtitalia	32'564	5'153
#lgbtqitaly	2086	331
#lgbtqitalia	218	39
#lesbicheitaliane	1'956	180
TOT.	71'779	9'434
UNIQUE	-	**8'292**

The InstaCrawlR downloaded information about 71'779 pictures posted on Instagram by 9'434 users. Often the same user tags more than once his/her pictures with the same hashtag, or analogously a user can employ more than one LGBTQ hashtag to tag his/her pictures. 8'292 univocal users have been listed to avoid erroneous duplications.

Talking about the black box, it is important to specify some loose ends of this web-scraping case study. Downloaded posts containing a certain hashtag do not represent all those pictures ever posted on Instagram since it went online, but just the "latest". Instagram's API does not give a specific time frame for the output of our request. It just quits when it reaches an unknown limit, confirming the main weakness of social media web-scraping: the lack of transparency. Actually, this explains the difference between posts in table 1 and table 3. For instance, post tagged with #gayitalia, up to 20th September 2018 were 228'000, but InstaCrawlR script downloaded (on 20th March 2019) just 2'730 posts. On the other hand, for some other hashtags, the size of downloaded posts is higher than those reported in table 1, because of a delay between the explorative analysis and the effective download, due to troubleshooting problems with API's update: the

most glaring case is #lgbtitalia, with 32'564 downloaded posts in March 2019, and 26'100 posts up to September 2018.

To recap, we can say that web-scraping represents for social research a mature computational method to use (McDonnell, 2020). It can be carried out through many packages, both in Python and in R, and software applications, which allow engineering the data collection phase from a web page; thus, web-scraping permits the reshaping of data into ready-to-use table formats. In our case web-scraping allowed us to list the most recent Instagram posts containing the most common hashtags of the LGBTQ Italian population(s). From this list, a list of username has been extrapolated. This user list represented a sampling frame for the survey we carried out, through the questionnaire that will be deeply presented in the next chapters. For the first time, the list of LGBTQ people we asked to participate in the survey is independent of LGBTQ associations, which usually partnered with surveys on this field, as we often saw in the previous part of this dissertation.

However, web scraping has its methodological weaknesses. Its use walks on a tightrope and to base social research on web-scraping could be risky: web pages are frequently updated, therefore changes to their structure can break a script, though to be reliable; some web sites (especially social media) start to block the scraping of their contents, which they are jealous of. Besides computational problems, it should be pointed out that the ethical implications of web-scraping for social science research cannot be overviewed (Mäkitalo, 2020): frequent scraping scripts can overload a server by making too many requests, causing the website to crash. This possibility marks a fine line between collecting information automatically and stressing a website's hosting services. Web-scraping could be considered by some websites as a cyber-attack called DOS[15] (Denny, 2017).

[15] DOS stands for Denial Of Service. It is a malicious, intentional attack meant to bring down a website. Its functionis as basic as efficient: the attacker sets up a computer to make as many requests as possible for information from a website, automatically, until the server crashes

In conclusion, we can say that breaking into the black box of social media websites to get information has lots of advantages, but a downside too. Web-scraping Instagram enabled us to have a widespread list of likely-LGTBQ people, finally separate from associations, to send our questionnaire to. Moreover, the script immediately produces an output in table format (shown in table 2). On the other side, we had to entrust an important part of our research to a script, which communicates with the website through an API, in a not fully transparent way. If the website structure changed our script would not work anymore. This consideration has to warn all social researchers who want to use web-scraping, about the dependence of the social research by the used tool.

6 The sequential use of qualitative and quantitative methods for studying LBGTQs.

6.1 What to ask and how to ask it? Drafting a proper questionnaire about LGBTQ

In the previous chapter, we saw how to extrapolate a sampling frame from Instagram using a web scraping script. In this way, we have obtained a user list to send a web questionnaire link to. But planning a survey about LGBTQ entails answering important questions: what is the best way to ask about sexual orientation, gender identity and all other issues related thereto? Moreover, which are the most important aspects to investigate about the community?

Designing a questionnaire is not at all trivial. Identifying the aspects to investigate, writing questions capable of arguing or confuting the research hypothesis, choosing the order of the questions which would not influence the interviewee, are just some of the several best practices proposed in the literature to researchers who are planning a survey (Gobo, 2015). If a survey deals with LGBTQ (Lesbian, Gay, Bisexual, Transsexual and Queer people) the methodological attention needs to take into account further increases.

Methodological reports of most of the surveys, introduced in the early chapters, warn about the complexity of drafting a questionnaire on sensitive subjects, such as sexual orientation and gender identity, self-perception and homophobia: writing questions, choosing their order, and picking the words used to ask them, has to be a scrupulous job. For instance, in 2012 the European Union Agency for Fundamental Rights (FRA) carried out the first European Union LGBT survey, probably the most important statistical survey ever conducted on this field until then. Its methodological report focuses on the importance of the consultation and of the careful examination of the final version of the questionnaire to submit to the users. FRA coordinated and supervised the consultation process, to draft and sharpen the questionnaire, in collaboration with the LGBT community. The consultation process focuses on terminology, contents, layout and question order, leading to the final version of an online questionnaire submitted to LGBT people spread across 28 European countries.

Following the FRA recommendations (and those of other organizations which carried out other important sociological quantitative studies about LGBT, such as Green (2012) and Pew Research Center (2013)) several semi-structured interviews have been carried out. These interviews aim to identify some issues in addressing online questionnaire drafting. Additionally, semi-structured interviews directly allowed a part of the survey target population to actively contribute to writing questions, on which the survey is based. Interviews give the analysis the added value of the direct experience. In this way, the research project acquires the methodological strength of the sequential use of qualitative and quantitative methods, in a pragmatic perspective of combining both methods.

6.2 Qualitative vs. Quantitative: the sequential pragmatic approach

Talking about labels, through bipartitions, is always difficult in this research project. As we have already seen, defining sexual orientation, dividing people into heterosexual and homosexual is not enough either. Talking about gender identity, dividing people into male and female is no longer enough, in the same way. Analogously, in methodological terms, distinguishing categorically between qualitative and quantitative research is not enough either.

Rossmann and Wilson proposed, as early as 1985, that quantitative and qualitative methods can be combined in a single evaluation study to better understand the phenomenon in question, according to three main perspectives on combining methods:

- the purist approach where the two methods are seen as mutually exclusive,
- the situationalist approach that views them as separate but equal,
- the pragmatist approach that suggests integration is possible.

From the pragmatist position, it is argued that either method can be used at the analysis stage to corroborate, elaborate or validate findings from the other method.

In the social science research literature, the use of multiple methods is usually described with several different names: multi-method/multi-trait (Campbell and

Fiske, 1959), convergent validation, or triangulation (Webb et al, 1999). These various notions share the concept that qualitative and quantitative methods should be viewed as complementary rather than competitive. Jick, in 1983, underscores the desirability of mixing methods given the strengths and weaknesses found in a single method design. Mainly for planning a survey, whose purpose is to elucidate causal relationships or even to provide descriptive statistics, qualitative methods should be used to ensure that the questionnaire contains all the right questions, asked in the right way (Gable, 1994). Moreover, designing a survey about LGBTQ without exploiting informative richness features of qualitative methods can lead to a misleading questionnaire and then to non-robust results, a synonym of partial and patchy knowledge of the phenomena.

In this context, a pragmatic approach has been used which combines, in a sequential way, a qualitative method (semi-structured interviews) before, and a quantitative survey after. 6 people belonging to different LGBTQ population(s) groups have been interviewed: two young lesbian women, one of which defines herself as queer, three gay men and a FtoM transsexual person. The people interviewed have been selected in a non-probabilistic way, trying to find different people in terms of age, profession, social background, educational level and involvement level in any LGBTQ associations.

Each was asked which aspects of daily life to investigate and how, taking into account question sensitivity and the complexity of gender identity and sexual orientation issues (Ruspini, 2014).

Interviews have been conducted letting interviewed people express themselves as freely as possible, instead of strictly organizing questions. The following topics are proposed for discussion during the interviews:

- interactions on/with social media;
- belonging to LGBTQ community, associations and participation to their activities;
- homotransphobic behaviours to investigate, and how;
- how to ask about self-perception, according to the main definitions of sex, gender identity and sexual orientation.

Each interview lasted about an hour and they were sometimes conducted in university rooms, and other times in bookshops or bars. The participants of the interviews accepted to being recorded and quoted in the research project framework, in scientific publications or PhD thesis. In the following paragraphs, fictitious names are used in quoting the interviewed people and in reporting what they said.

The interviewees, while discussing the above topics, highlighted recurring themes, despite their differences in terms of age, job, social background, educational level but also, and above all, sex, gender and sexual orientation, which collocate them into the LGBTQ community. These recurring themes have been theoretically analysed, regarding two main frameworks on which this research project grounds: gender and digitalization.

This analysis pointed out several important issues which addressed, in a theoretical way, the questions proposed in the survey and how they were asked.

6.3 How digital technologies changed the LGBTQ community daily life

The first field which emerges through interviews, talking about the relationship between LGBTQ and new technology, is the revolutionary impact generated by dating apps: Grindr, Her, Tinder, just to mention a few which are widely used, mainly among gay men. They enormously increased the possibility of easily meeting new people, knowing in advance their sexual orientation, interests, location and all the other characteristics which are important when meeting a new person (Card et al., 2017; Grov et al., 2014).

When asking interviewees how many of their LGBTQ acquaintances use, or have used, a dating app all of them answered with percentages between 80 and 95%. Giorgio explains to us the main advantages of using this kind of app and the reasons behind their success:

I downloaded Grindr because a friend of mine suggested it to me […] in that period I was complaining about the indifference of people, nobody looks at you or winks anymore, everyone is with their smartphones, so I asked him "how can I pick up someone?" and he answered with "why don't you try with Grindr?" and so I tried and it was great […] the aim was to meet up in real life, according to the proximity between two people. In Grindr distance is really important: you localize yourself and you say "hey, I am here and I can move in this range to meet up" because if I wanted a pen friend I would accept him in Australia too, but I need a real person and I need the contact with him.

Luca confirms what Giorgio said, regarding the importance of knowing others gender, sexual orientation and especially where the person you are meeting is. He argues:

I have lived during smartphones and the dating app dawn. geolocalization was an unthinkable thing before […] when I was younger there were some chats to talk on through PCs but then, since 2007/2008 with smartphones arrival, dating apps like Grindr took dominance. It is like a closed social network with the aim of meeting, not only for having sex… for example I met Emanuele on Grindr and we have been engaged since 2012

Transgender people took advantage of social networking websites not only for meeting but also for searching and sharing information, difficult to find before: in many Facebook groups it is possible to ask for tips and suggestions about transitioning and on YouTube, there are several channels about the same subject. Emi, a FtoM transsexual person, is the admin of one of these Facebook groups. He founded it to share information with people looking for the support he wanted in the past and that was hard to find. This is what he told us:

We have a group called "FtoM e MtoF" tips, about doctors, endocrinologists, etc… it is a Facebook page in which people looking for suggestions can ask whatever; there is also a post-transition group in which people ask things like "This happening …what about you?" Some years ago everything was much more difficult, I thought to be the only one on this planet to experience

certain things. I thought "why I did it?!" Nowadays social media make it easier because you know what you are going to do. In 2014 I found out much info about transition, really difficult to get some years before: now you can find someone able to help you if you need it.

Technology, said to encourage people to always be online and thus colder and blasé in real daily life, does not seem to penalize socialization. On the contrary, although in ways characterizing our informational capitalism age (Castells, 2002). Today's society is presented with a way of meeting people different from some years ago. Geolocalization undoubtedly represents the main advantage of dating apps: the possibility of knowing who is close and inclined to meet up makes our devices "perform" users' bodies, feelings and desires (Butler, 2015) but at the same time dating apps construct gender just as a rigid category that has more to do with a matching profile function than with identity (MacLeod and McArthur, 2018)

This new socialization ease has brought decreasing participation in all those activities, mainly sponsored by LGBTQ associations, whose aim was to promote socialization between people sharing the same sexual orientation, such as cineforum, disco events or soirées in gay-bars (Ross et al., 2014). This side-effect introduced by new technology, confirmed by a member of the board of the roman LGBTQ association named after "Mario Mieli", is perceived by respondents. This is what Chiara told me about the decrease in socialization activities promoted by associations:

I don't know if associations are growing, although the important work they have carried out in the last few years, mainly in suburbs and in small towns, because social media make all of us more and more isolated: you can think that your field of action is comparable to a collective, but it is not, and maybe some people think that it's useless to be part of an association to lobby.

LGBTQ associations nevertheless remain fundamental as political pressure institutions on the front of civil rights recognition and the fight against homotransphobia (Barbagli and Colombo, 2007). The associations' prestige thus has not decreased but has stabilized, as it has happened to all other institutions involved in political activism, such as political parties or trade unions (Simon Rosser et al., 2008). Individualistic drift, deepened by new technologies in daily

life, still doesn't work in the political struggle where the self, political subject, becomes "us" in public (Butler, 2015).

Online social media are not only used to organize encounters but also to correspond and sometimes to endorse positions, by sharing them. Carlotta told us in her interview that the Internet was her first interlocutor when she realized her sexual orientation and it soothed her:

> *When I realised this "thing" [sexual attraction for same-sex people] I remembered that some lesbian girls had suggested a Facebook page to me, actually it is a magazine, called Lezpop. I remember it was reassuring so I followed the page, but without liking it, because I was afraid someone could discover it…*

Sometimes social media are used to read and then share positions. This kind of endorsement often does not coincide with traditional political activism, like that one carried out by associations. Many LGBTQ people usually attend main events in person, such as the Pride. Social media become a "megaphone" from which everyone can shout their political commitment, trying to involve as many people as possible. Interviewed people declared feeling part of the LGBTQ community mainly during grieving and pride moments, but not really in daily life: social networking websites fill this gap, integrating everyday life with the sharing of information, contacts or events, contributing in constituting a new way of political attitude and democratic participation (Zhang et al., 2010).

According to what has emerged in interviews, about the relationship between digital technologies and the LGBTQ community, several findings pointed out to have been "operationalized" in the questionnaire, through specific questions, both about social media and dating apps, such as:

- "Which social network do you use, between Facebook, Twitter, Instagram, others? And how often?",
- "With which device do you log in?",
- "Which social network do you use, among Facebook, Twitter, Instagram, and others to post videos and pictures, to be informed, to chat, and to follow people?",
- "What do you think about social media?" (5-levels Likert scale applied to a set of sentences),

- "Have you ever used a social network to support the LGBTQ community? which one between Facebook, Twitter, Instagram?",
- "Is your sexual orientation deductible by your social media profile?",
- "Have you ever used dating apps?",
- "Have you ever hung out with someone met on a dating app?",
- "Do you have a public profile on dating apps or a closed one?",
- "If you stopped using dating apps, why did you stop?",

Moreover, several other questions have been introduced in the questionnaire to understand (also through Likert scale) the relationship between LGBTQs and associations, mainly to clarify if the "belonging" concept has changed over the years, also and foremost due to the social network advent.

- "Do you frequent LGBTQ associations?"
- "Do you frequent gay bars or LGBTQ disco nights?"
- "What do you think about social media and associations?" (Likert scale applied to a set of sentences),
- "Do you feel an integral part of the LGBTQ community?"
- "Which LGBTQ claims do you agree with (same-sex marriages, adoptions, homotransphobia crime, etc.)?
- "Have you ever been to a Pride*?"*

It is important to say that semi-structured interviews, which propose topics instead of specific questions have contributed to point out characteristic aspects of the current LGBTQ community: widespread dating app, the use of social networks to socialize, to keep up with the news, to claim rights and to perform its own identity, LGBTQ associations transformation and their role towards civil society, etc.

All these aspects, during the continuation of this research project, would have been overlooked if interviews had not been carried out. The aim of the questionnaire became also to understand how social media, and generically digital technologies, have changed the community, modifying the socialization process and the way of supporting civil right claims.

6.4 Gender matters: self-perception and discrimination within LGBTQ

One of the most important aspects of the interview was the method in which they, through questionnaires, asked about sex, gender identity and sexual orientation: how many and which modalities must we consider to analyse the heterogeneity and the complexity of the LGBTQ population(s)? The people interviewed showed a "conscious disinterest" on this matter: they responded that the high variety of definitions, related to both gender identity and sexual orientation, has been useful till today to show the world, hegemonically heterosexual, the strong versatility of the LGBTQ community; but today the proliferation of definitions is not as essential to understand the key characteristics of the LGBTQ people. This is what Luca said about the umbrella term LGBTQ:

> *I don't feel it is like a label, or a bad thing. It's just an acronym, enlarging more and more, sometimes in a paroxysmal way, including really different worlds.*

In fact about the choice of modalities to insert in the self-perception questions Luca adds:

> *I don't like all this politically correct terminology, for example, I don't like to use asterisks to say "salve a tutt*"*

From the interviews, it clearly emerges that Gender is performed in daily life, fluid and so that it's not right to categorize it in a binary way (Butler, 1990); but at the same time it is not right to try to categorize it in other several ways and being afraid of using the wrong definition. For this reason, the term queer is used more and more often, by all those people who feel they have a gender or sexual identity out-of-the-box, according to the rule laid down by society (Hughes, 2005; Ahmed, 2006). Chiara endorses this queer thinking:

> *Queer is an English word that means weird, strange, it was a denigrating word to identify who didn't represent heterosexuality with a reproductive aim: every deviance was queer. I have had experiences with boys (never again!) but I don't consider myself a bisexual person. I'm a lesbian queer woman. I give a political value to this definition, according to the queer thinking of refusing the society's mandatory duty of giving birth.*

The attention given to all the gender declinations (transgender, intersex, etc.) and all the sexual orientations (pansexual, polysexual, asexual, etc.) seems to be, for those interviewed, excessive, redundant and often also "too politically correct" (Redattore Sociale, 2013). Nevertheless, terminology attention could be very important for someone. On that basis it was agreed to leave an open answer option, to leave everyone free to define herself/himself in the questions about gender identity, sexual orientation and in all the other related questions. Giorgio, ironically, offered this suggestion:

> *I am gay and I feel male. I have never denied it. Anyway, some people want to feel neutral, how to treat them? I suggest you consider a modality "other"*

Accordingly, interviewees pointed out that sex, gender identity and sexual orientation were combined effortlessly with other socio-demographic variables, proposing an open option to leave everyone free to define herself/himself.

The options proposed for selecting one's sex were male, female, or intersex. For gender identity they were male, female, transgender or "queer, gender-fluid, non-binary" and the open option "other". For sexual orientation, the options were asexual, bisexual, heterosexual, homosexual, and the open option to define it freely. Obviously, categorizing gender identity and sexual orientation through a few-option answer could be reductive, but useful, mainly in a MAWI[16]-oriented survey, which requires a simple and feasible questionnaire. Anyway, the "other" option guarantees everyone the opportunity of self-definition.

The questionnaire continues with a Likert-scale question about the self-perception of gender identity and sexual orientation, which aims to understand how the LGBTQ population(s) lives these concepts.

Finally, a set of coming-out questions ask who was the first to find out their sexual orientation (none, someone in the household, friends, colleagues, etc.) and who knows it right now. The difference between the first and the others underlines the people who can be unconditionally trusted, mainly in youth, and people who need more time to be completely trusted.

[16] MAWI stands for Mobile Assisted Web Interview, a questionnaire that could be easily filled out using a mobile device, such as smartphones or tablets.

The last part of the interview was about sexual discrimination, how it happens and the way to investigate it. Interviewed people suggested how to make a question on this sensitive field. This is what Chiara suggested:

> *I would divide discrimination in 3 types: family, job, strangers. When you come out, or you don't but someone suspects you are queer, discriminatory behaviours could occur: old friends start to procrastinate to meet you or they stop contacting you, but also worse things, like sons or daughters kicked out by their parents. Then I think about discrimination on the job: lots of people have lost their job due to their sexual orientation. And finally, there are all the discriminating behaviours of strangers who insult or attack LGBTQ people in the streets.*

During interviews, discrimination has been often described as a set of behaviours less violent than expected but not less agonizing. A certain behaviour above all, the "*ghosting*", namely the situation in which friends, relatives and colleagues disappear from the life of those who come out (Spitale, 2015), as told by Carlotta:

> *Someone drifted away over time, my ex-boyfriend actually hates me: he didn't know it from me and we have never discussed it, he simply distanced himself from me, all the contacts, removed [by all the social media]*

Giorgio said that ghosting can be also stronger in the family environment:

> *Family discrimination is really strong because it's difficult to be indifferent with people you are often in contact with: siblings, in-laws… you start to feel something bad in their conduct, they start to not leave their children with you, I mean, your nephews… you start to understand you are avoided, not invited to the family gatherings, and so on… your relatives start to be afraid that you could address the sexual orientation of their children…*

The discrimination matter assumes a characteristic dimension for transsexual people more than all the other topics touched during interviews. Using healthcare, long-lasting transition, slow bureaucracy put at risk of suffering from

discrimination many transsexual people, mainly MtoFs, for which it is harder to hide their past identity. Circumstances of discriminatory events which transsexual feel the most is undoubtedly on the labour market: transsexual people often denounce the hardness due to their identity to find a job, as Emi told us:

> *Most people think a transsexual is a streetwalker but the reality is different: we are doctors, singers, engineers, etc. and if nobody talks about transsexuality we will always be relegated to that prostitute stereotype. I want to establish that we are not streetwalkers and I want to help women on the streets find a better job. That's the reason why I left associations because they didn't want to push this kind of project: I would like to help people, talking with institutions and enterprises… for example, IKEA now accepts a transsexual labour force*

Four questions about discrimination have been asked, to make the questionnaire as light as possible on this sensitive subject. These four questions are:

- "Have you ever been a victim of discriminatory events in your household, and by whom (your mother, your father, your sister, your brother, other relatives)?"
- "Have you ever been a victim of discriminatory events outside of the household, and by whom? (my mother, my father, my sister, my brother, other relatives)?"
- "Have you ever suffered from mobbing for sexual reasons, and by whom?"
- "Have you ever been bullied for sexual reasons at school/university, and by whom?"
- "Have you ever been discriminated in sport, and by whom?"

The first two questions consider the answer option "People started to avoid me after I came out", specifically considered to operationalise in ghosting, emerged in the interviews, as an unconventional and underrated way of discrimination, both inside and outside the household.

In conclusion, it can be said that asking about violent acts of homophobia, such as beating or injuries, seems not to be enough for investigating discrimination. Less direct homonegativity behaviours, like ghosting or

mobbing, have to be taken into account. Both of them are ways of avoidance: hanging out with LGBTQ people is avoided, working with LGBTQ people is avoided, pronouncing "gay" or "lesbian" is avoided, saying "partner" is avoided, and so on. All this avoidance represents a sort of opposition, or at least impropriety, towards LGBTQ people (Graglia, 2012) and this is what could be resolved in social policies.

Methodologies used and their theoretical implication: concluding remarks

In the previous chapters, we have seen how web-scraping script allows us to obtain a sampling frame of people likely LGBTQ (or at least supporting the community), independent by LGBTQ associations. This technique allowed us to get into the social media black box, but not without reasoning over what these techniques involve in the social sciences field, and, in particular, in this research project.

The epistemological debate about the use of big data in social research is heated and constantly evolving. Two opposite poles of the debate: someone adopt a pessimistic point of view, like in critical digital sociology, and someone else instead has a too "idyllic" point of view, thinking about new digital media as the answer to all the questions that traditional sociology posed. Both these two points of view may be too drastic and they could be misleading as well (Marres, 2012). In this animated and contrasting context, this research project pursues another direction: working *on* big data and not *with* big data, leveraging the data richness guaranteed by the power tool, without leaving out that complex algorithms will steer the results and the knowledge about the LGBTQ community.

Social media has thus been considered as a new social space, self-expressive and free, though subject to its netiquettes (Snee et al., 2016). However, social media represents a research environment too, with its strengths and weaknesses. These websites (and their mobile applications) allow their users to perform their own gender identification and sexual orientation, in favour of the LGBTQ community. Among all the social media which benefits this community, and the social research on it, Instagram is actually the one that mainly allows the free self-representation of LGBTQ people. Other social media websites have been stricter in the last years, like Facebook, with its real name policy, or misleading, like Twitter, which has allowed for years pornographic contents, which increased stigmas and stereotypes.

Instagram became the framework not only during the web-scraping process, but also during the data-collection phase, directly messaging interviewed people and reassuring them about the privacy conditions of the survey, through the research project page.

The questionnaire took into account the digital orientation of the project and to do so, it asked which aspects to investigate and how, through semi-structured interviews, considering the new opportunities that digital technologies have led for the LGBTQ community.

These interviews pointed out some issues that would have been completely neglected if interviews had not been carried out. These recurring themes were "operationalized" in the questionnaire through multiple-choice questions, which often contain a free response option, to leave the interviewed person as free as possible, mainly in the questions about his/her own gender identity, sexual orientation, discrimination endured and self-representation online.

In the concluding remarks of this PART 2, it is worthwhile to remember that the web-scraping script, working through an API, makes the work as practical as "cryptic": not downloading all the posts, containing a certain set of hashtags, but just the most recent, a bias is unavoidably introduced. Another bias to consider is given by a selection of just Instagram users who use LGBTQ hashtags to index their posts. To fix this bias and to increase the catchment area of interviewed users, a snowballing technique has been introduced in the survey. Snowballing is a type of purposive sample[17], useful for the researchers who are trying to recruit people who are difficult to identify, or rare and elusive, as De Rosa and Inglese (2018) defined the LGBTQ population(s). Snowballing can be used to ease data collection: find one person who participates in the survey, ask him or her to recommend several other people who have the traits we are looking for. From there a participant list can easily grow (Abdul Quader et al., 2006).

[17] The sample of LGBTQ individuals to include in the survey, and to select for answering the questionnaire, is definitively a non-probabilistic one. Because there is not a frame for a target population to select a sample from. As it's not possible to calculate the inclusion probability of every sampling unit, and hence the sampling weights, it's not possible to make inference on it. So we cannot extend the sampling results to the whole population in a statistical way, but in methodological literature several *purposive sampling methods* are presented as the solution to solve this statistical problem.

At the end of the questionnaire (see Appendix Q) it is asked to suggest up to three Instagram users with the same sexual orientation as them, and interview them in turn. This suggestion can lead to interviewing people who don't come out on Instagram because they don't use the hashtags analyzed in the tagline. Using the snowball method entails two main benefits for the survey. First, it increases the selection frame, because we are not interviewing just people using coming-out hashtags anymore. Secondly, snowballing can approximate a random sample and so it gives more robust and reliable results (Abdul Quader et al., 2006).

PART 3

SURVEY RESULTS

No law's gonna change us
We have to change us. Whatever God you believe in
We come from the same one
Strip away the fear
Underneath it's all the same love

MACKLEMORE – Same love (2012)

7 "Who would answer to an LGBTQ survey on Instagram?": socio-demographic patterns of the respondents

7.1 Data collection phase, sample size and response rate

The Over The Rainbow survey has been carried out online, through the Google Forms tool. The link to the questionnaire has been individually shared through Instagram Direct Messages (DM, from now on) to 4216 users, randomly selected from those 8290 users listed using the web-scraping, a procedure previously described. Moreover, the questionnaire has been sent to 638 users named by other respondents, according to the snowballing procedure, introduced, as previously stated, to fix sampling bias and to increase the catchment area of interviewed users.

Answering the questionnaire has been possible from 1st April 2019 to 31st January 2020. During data-collection months the number of daily respondents was highly variable, and generally decreasing after the first six months, as illustrated in the following graph:

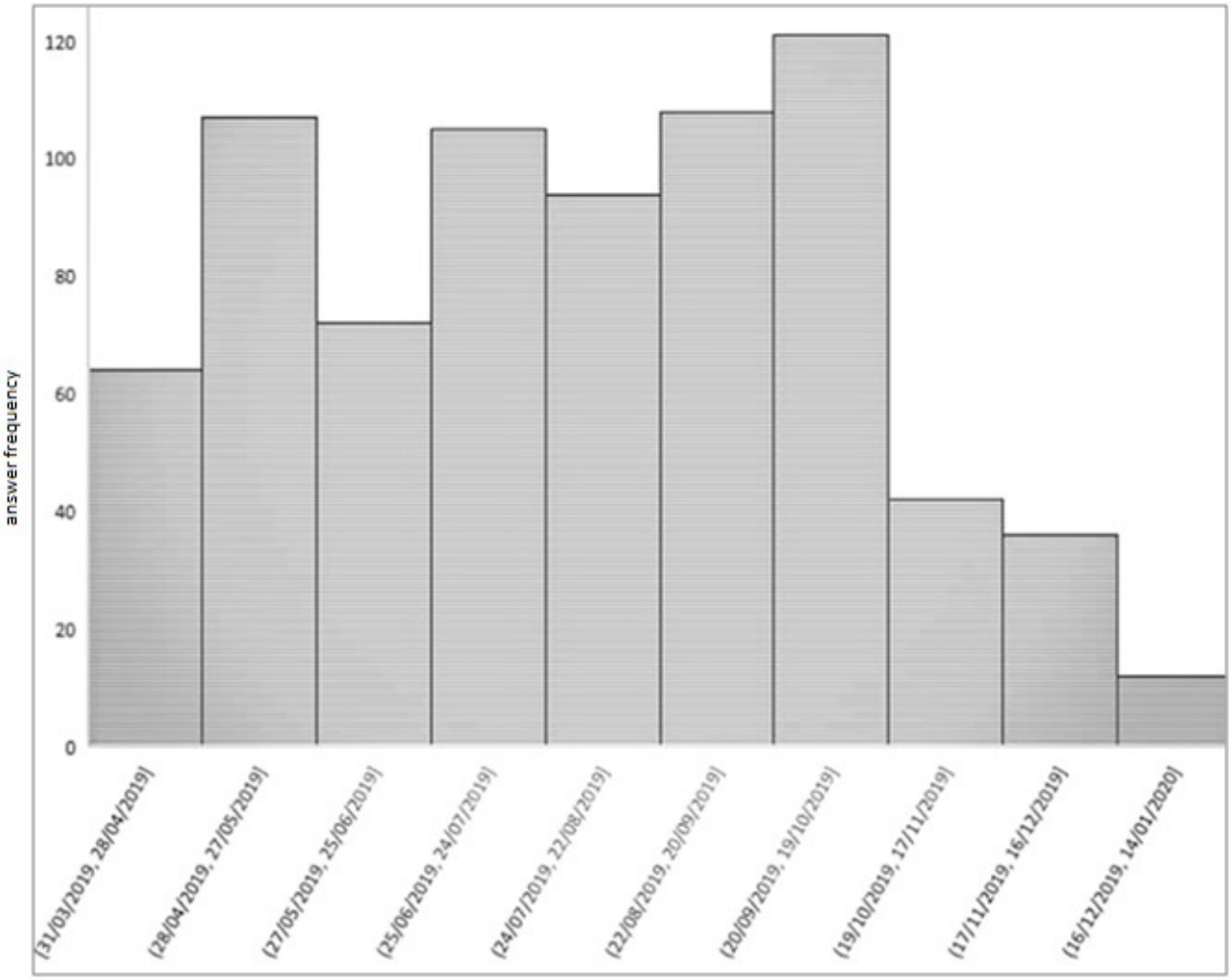


Fig. 1 – Trend of monthly survey responses

The decrease in response numbers was principally due to Instagram's anti-spam policies, which blocked temporarily, and repeatedly, the messages sent by the OTR account, suspected of being spam messages. These blocks were the main problem of the data-collection phase, and the reason for a response rate lower than expected.

The final number of survey respondents was 763.

45 respondents were removed because they did not authorize the data treatment, as requested at the end of the questionnaire, according to the Italian privacy law (Legislative Decree 96/2003) about the processing of respondents' personal information, the data-storage in an adequately safe database, the anonymity of given answers, and, finally, in regards to the statistical and sociological purpose of the survey.

The responses considered as valid, which will be deeply analysed in the following paragraphs and chapters, in the end, were **718**. The response rate, given by the ratio of valid responses and contacted users (4216+638), is **14.8%**.

Respondent people, mainly intercepted according to their usage of a set of hashtags on their Instagram posts, revealed some peculiarities from a socio-demographic point of view. These characteristics emerge from the questions, introduced in the first part of the questionnaire which aimed to understand the respondents' features and the survey users' catchment area.

In this chapter, we will try to understand who are the OTR survey respondents, why their collaboration is so important, and which interpretative potential their answers give to this survey. To do so, we will evaluate certain significant socio-demographic characteristics of respondents, such as age, employment, geographical area, both of residence and origin, and relationship status. In the next chapters, we will cross these socio-demographic characteristics with other variables, such as biological sex, gender identity, sexual orientation and many others survey variables, adopting an intersectional approach for studying the LGBTQ population(s). The Intersectionality concept is a theoretical and methodological dispositive to analyse how the intersection between several social and power dimensions contribute to perform and reproduce conditions of inequality and discrimination in the co-construction of social categorizations as gender, social background and sexuality (Lykke, 2011).

7.2 A young and free LGBTQ population(s)

The age of those who collaborated with the survey is the first piece of evidence, and probably the most important, of data analysis: respondents are young, some of them teenagers, but not because of this uncertainty of their gender identity and their sexual orientation.

In the next chapter, talking about self-perception, we will see that most of the people who participated in the survey defined themselves as non-heterosexual, each in its own way, and often they define their gender identity as non-binary. It is striking that the respondents' age distribution shows an over-representation of young age classes, with a modal age class under 21, and a relevant representation of minors (13.5% under 18 years old), till now never involved by most LGBTQ surveys, previously reviewed.

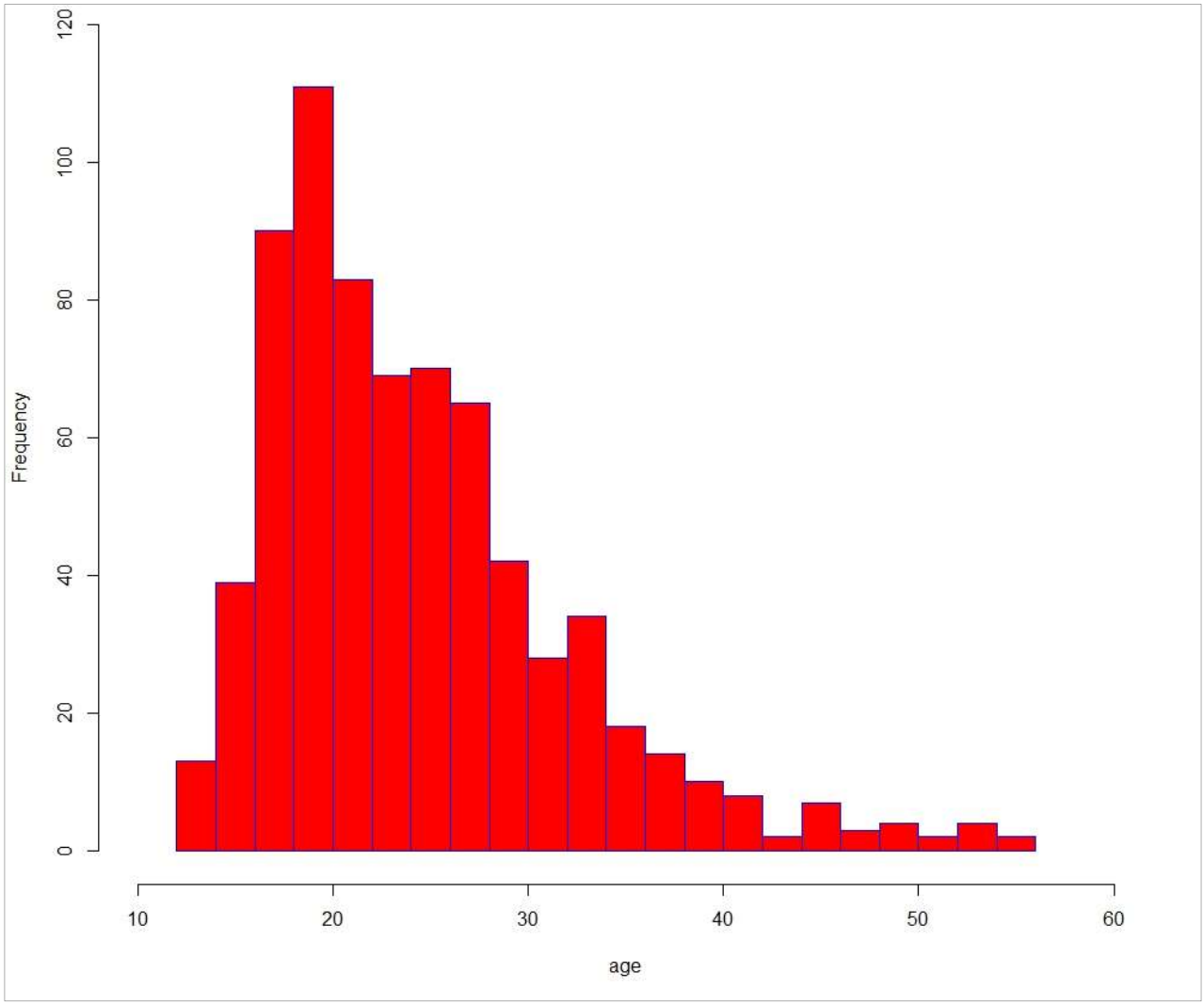


Fig. 2 – Respondents' age distribution

The respondents' average age is 24.84 years old, a mean value which decreases observing the median (23), and even more the mode (20). From a statistical point of view, this "unbalance" would introduce a bias, if we would make inference in the canonical and traditional way, respecting the strict assumptions of the sampling theory. However, working on a hidden population makes the inferential ambition misleading, since no reference population, nor known totals, are available. Working on a social network gave the possibility of intercepting a young, free and invisible, until now to social research, LGBTQ population(s). In fact, the majority of surveys already seen (PART 1) are oriented towards an adult LBGTQ population(s), both for confidentiality and unavailability issues. On the other hand, neglecting LGBTQ minors entails a negative side effect, underestimating some important bad behaviours, typical of

the young age classes, such as bullying and cyber-bullying (see Par. 11.4): schools, colleges and universities should be research environments to focus on, for understanding discrimination and affirmation dynamics of the youngest LGBTQ age classes. From this point of view, the high collaboration of young LGBTQs could be seen as an opportunity and not just as a source of statistical bias, because it allows investigating a part of the target population never deeply analyzed.

Respondents, although being young, are not just students as illustrated in the graph below, which shows the main activity (between studying, working, both, looking for a job, other activities) of respondents:

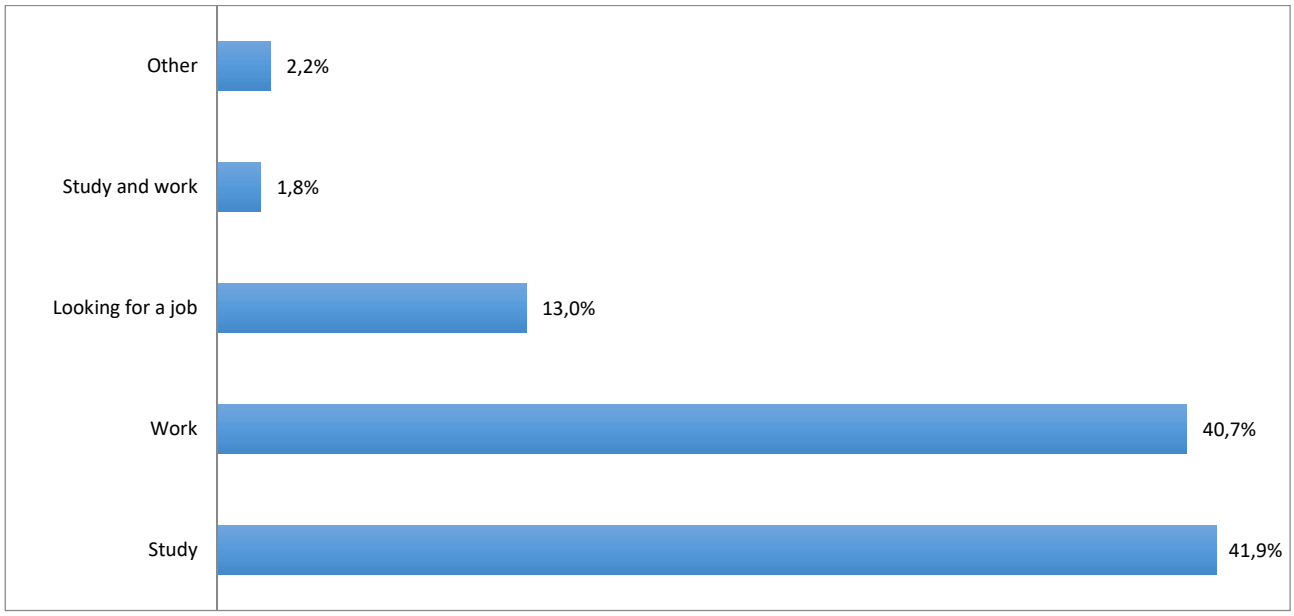


Fig. 3 – Respondents' main activity distribution

41.92% of respondents study, at university or in high school, 40.67% works, and 1.81% is simultaneously both a student and worker. 12,95% is looking for a job. The remaining 2.23% carries out other activities (specified in the open modality of the question), often attributable to "hybrid" activities, between education and work, such as stages, PhD, specialization, traineeships, internships, etc.

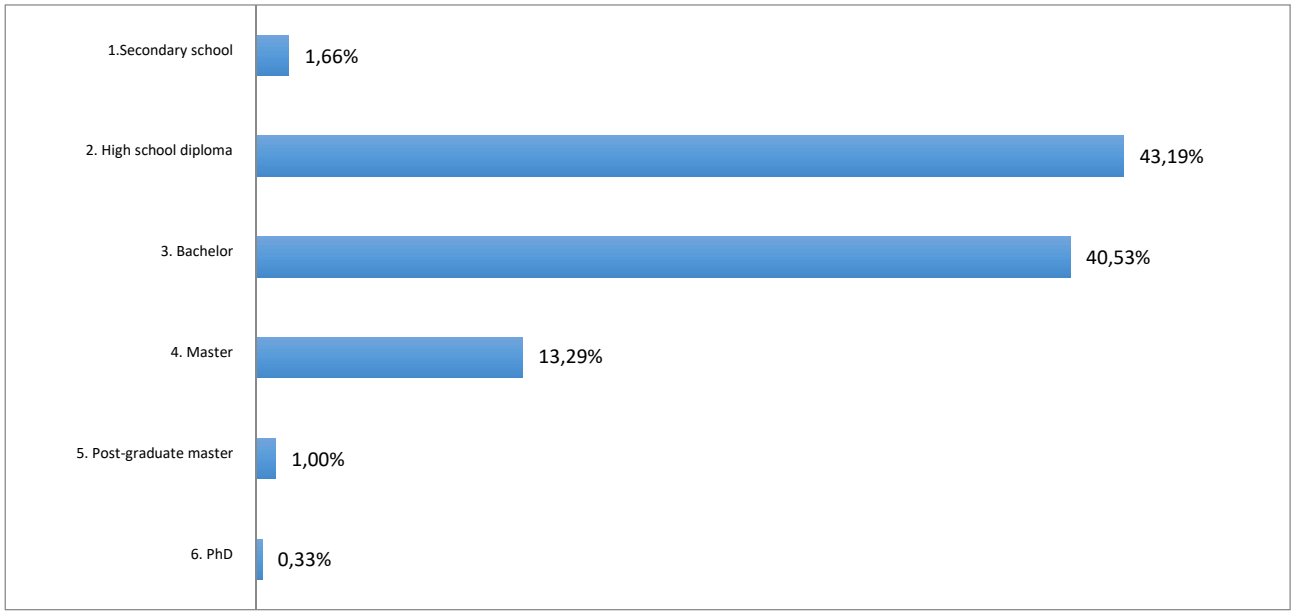


Fig. 4 – Respondent students' degree achieving

As seen in Fig. 3, 43.73%[18] of respondents are still studying, 43.19% of which is achieving a high school diploma, 40.53% a bachelor degree, 13.29% a master degree, 1.66% is a secondary school student and 1.33% of students in achieving a PhD. According to the degree that every student was achieving, type of high school or academic area was asked to students, for understanding if (and how) sexual orientation and gender identity influences education trajectories of LGBTQ students. The results of these analyses will be deeply analyzed in paragraph 11.3 looking over sexual discrimination at school and university.

Having intercepted young people allows us to compare different generations of LGBTQs. In many occidental countries, this issue has started to be studied, mainly comparing generation Z (people born between 1996 and 2007) and the previous one, Millennials (born between 1980 and 1995). For instance, an important survey, called "Beyond Binary. The lives and choices of Generation Z", conducted in 2018, by a British market research company, pointed out that 66% of Generation Z thinks of itself as exclusively heterosexual, compared with 71% of Millennials, 85% of Generation X and 88% of Baby Boomers. Moreover, 60% of 15-16-year-olds British schoolchildren think sexuality is not a binary definition, but a scale on which it is possible to be "somewhere in the middle". The same survey also focuses on the gender fluidness of Generation Z which, just to give a

[18] 41.92% students + 1.81% simultaneously both students and workers

pragmatic example, prefers gender-neutral clothing (61%), compared with 43% of Millennials (Duffy et al., 2018).

Awareness of ones own sexual and gender identity seems to be a process that has been anticipated during the last few years, and an increase in the number of people who identify themselves out-of-the-box. Statistics and social research cannot avoid this issue any longer, leaving a consistent part of the LGBTQ population(s) uninvestigated.

7.3 LGBTQ in the labour market: the workers respondents

Even though the respondents are young, the OTR survey also caught workers (40.67%), some student-workers (1.81%), and others who are job searching (12.95%), as shown in Fig. 3. In a survey about LGBTQ population(s), it is fundamental to catch workers to investigate affirmation and discrimination dynamics in the working context, which could, unfortunately, lead to sexual discrimination episodes sometimes, that we generically will refer to *mobbing* (see par. 11.2). Furthermore, it is important to also intercept a wide range of jobs, to fully understand the studied phenomena in a variegated working field.

The graph below points out how respondent that work distribute themselves according to job classification widely used in official statistics.

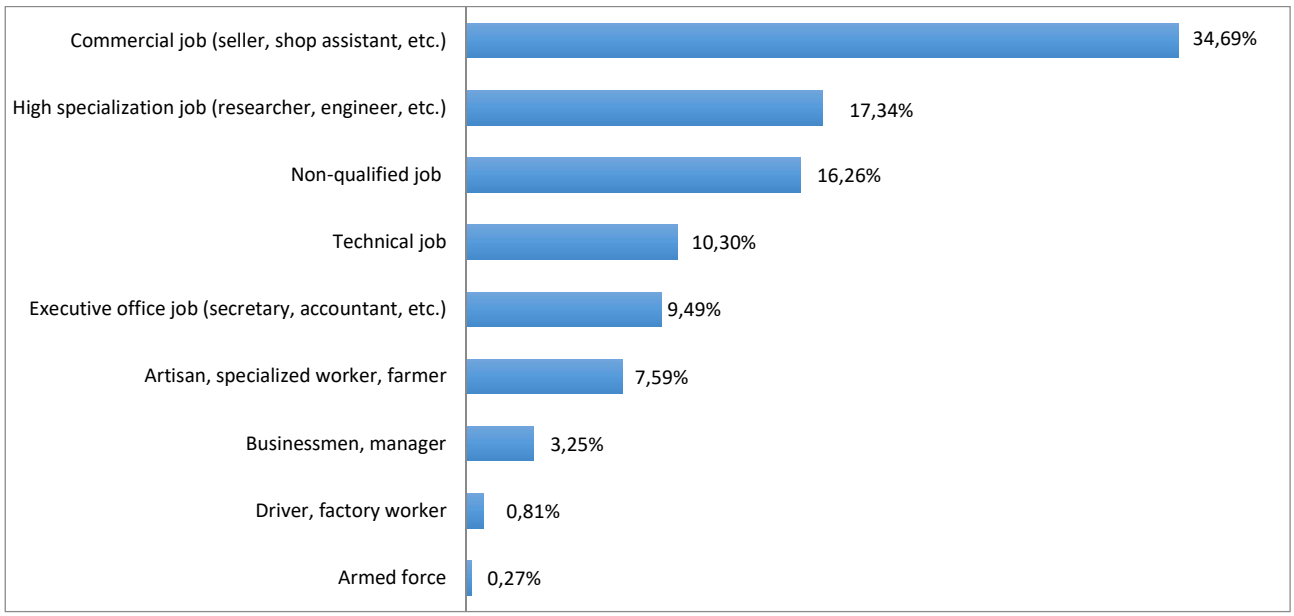


Fig. 5 – Respondent workers' type of job

From 42.48% of respondent workers, 34.69% has a qualified job in the field of commerce and services, such as a seller, shop assistant, etc.; 17.34% carries out an intellectual work activity, scientific and high-level specialization, such as researcher, engineer, consultant, etc. (this percentage is aligned with those of the respondents holding a degree (20%)). 16.26% carries out non-qualified jobs. Then, technical jobs (10.30%), secretaries and accountants (9.49%), artisans and farmers (7.59%), businessmen and managers (3.25%), drivers and factory workers (0.81%) and, last, soldiers (0.27%).

Looking at the monthly net wage distribution of respondents, in the graph below, we can consider that 1001-1500€ is the most frequent wage class and 79% of respondents have a net salary lower than 1501€ per month. In Italy, in 2019, the average net wage was 1550€ per month[19], even though there is strong differences in age, geographical area, job type and/or sex due to the gender pay gap.

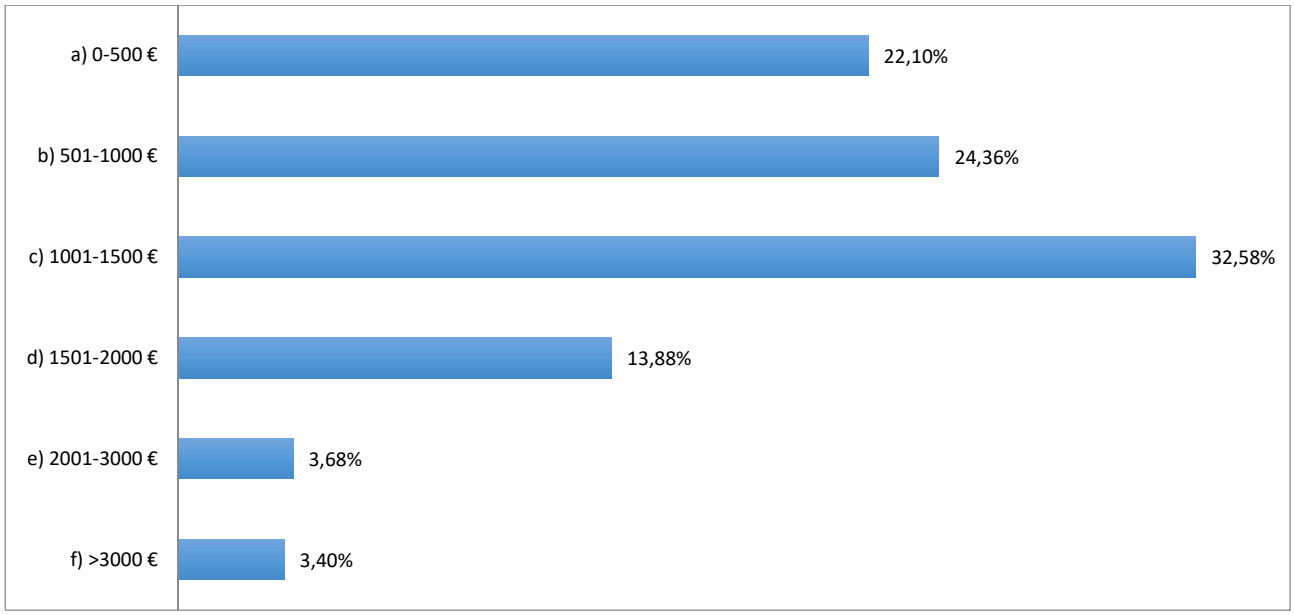


Fig. 6 – Respondent workers' monthly net wage

The phenomenon which led LGBTQ people's salaries to be, averagely, lower than those of heterosexuals is called the *gay pay gap*. Despite the name, this microeconomic discriminatory phenomenon is not exclusively male, on the contrary, following an intersectional approach, we could study if the gender

[19] average net wage calculated by JobPricing and Infojobs, on an average gross annual earnings of 29'352€

identity has a higher impact on lesbians than on gays, on MtoF transsexuals than on FtoM, on homosexuals than on bisexuals, and so on. These studies, if conducted just "in average", without taking into account correlations with other variables, nor comparing different people according to other conditions, such as education, social class, type of job, could lead to misleading results.

Affirming that the net monthly salary of OTR respondents is lower than the average Italians' salary is a hypothesis too hasty, given by the respondents' unbalance of younger age classes. Anyway, from a methodological point of view, it is extremely important to focus on the collaborative behaviour toward an online survey of so many workers, which could allow studying also economic discriminatory phenomena in the working field, such as mobbing or, indeed, the LGBTQ pay gap.

7.4 A respondents' geographic distribution which represents the whole population

Whether the sample was structurally biased by age classes, having caught more young people, on the other side it can be considered unbiased by geographical area of respondents. The graph below clearly shows the percentage distribution of respondents by region, both of birthplace and residence, and the percentage distribution of the whole Italian population.

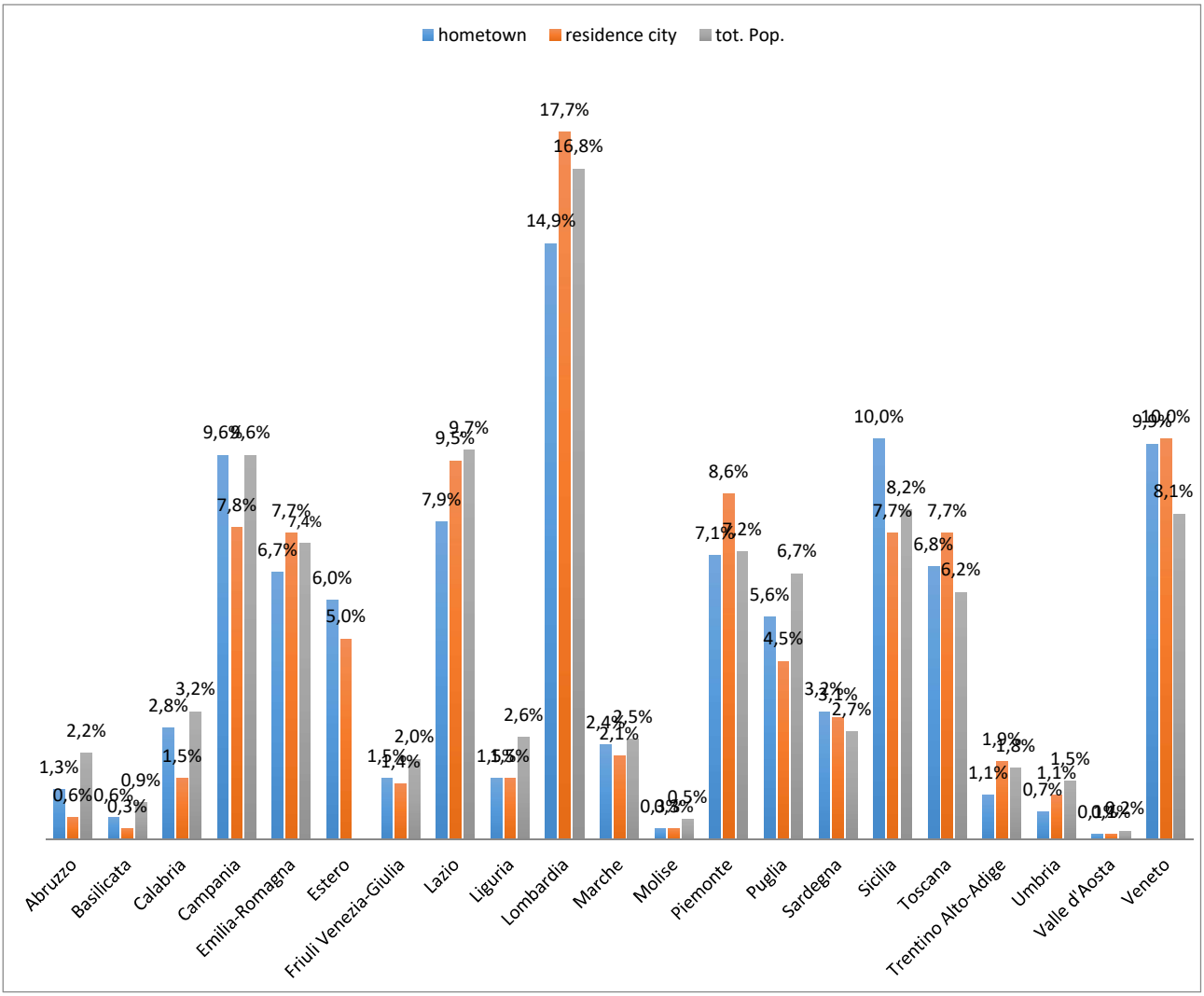


Fig. 7 – Respondents distribution by hometown and residence region, compared with the distribution of Italian population

The chart visibly shows that the percentages of the whole population among Italian regions, denoted by green bars, is correctly respected by the distribution of respondents to the survey OTR. This evidence confirms that the selection of users from social media, through the web-scraping procedure, strongly reduced the geographical bias which bothered most of the previous surveys about the LGBTQ population(s). As already noted, by not relying on LGBTQ associations, mainly located in big cities and absent in small towns, OTR involved in the survey (in a proportional size) also those regions which, for instance, do not have a metropolitan city as the regional seat, such as Basilicata, Molise, Valle d'Aosta, Marche, Umbria, etc.

The graph also points out how, in several regions (Lombardia, Lazio, Emilia-Romagna, Piemonte, Toscana) the proportion of residence respondents (red bars) is higher than the proportion of native respondents (blue bars).

This evidence could indicate that in those regions which have a metropolitan city as regional seat (respectively Milano, Roma, Bologna, Torino and Firenze) people arrive in adulthood, for working or for studying purposes. But since OTR respondents are mainly non-heterosexual, this phenomenon could be explained also as an escape from their birth town, for avoiding the stigma for a sexual orientation out-of-the-box, which could penalize people more distinctly in provinces than in a metropolitan city (Corbisiero and Monaco, 2017). A kind of national Queer diaspora (Wesling, 2008). This tendency could be clearly confirmed by the following graph which demonstrates the respondents' percentage distribution according to population size, both of residence and birth city:

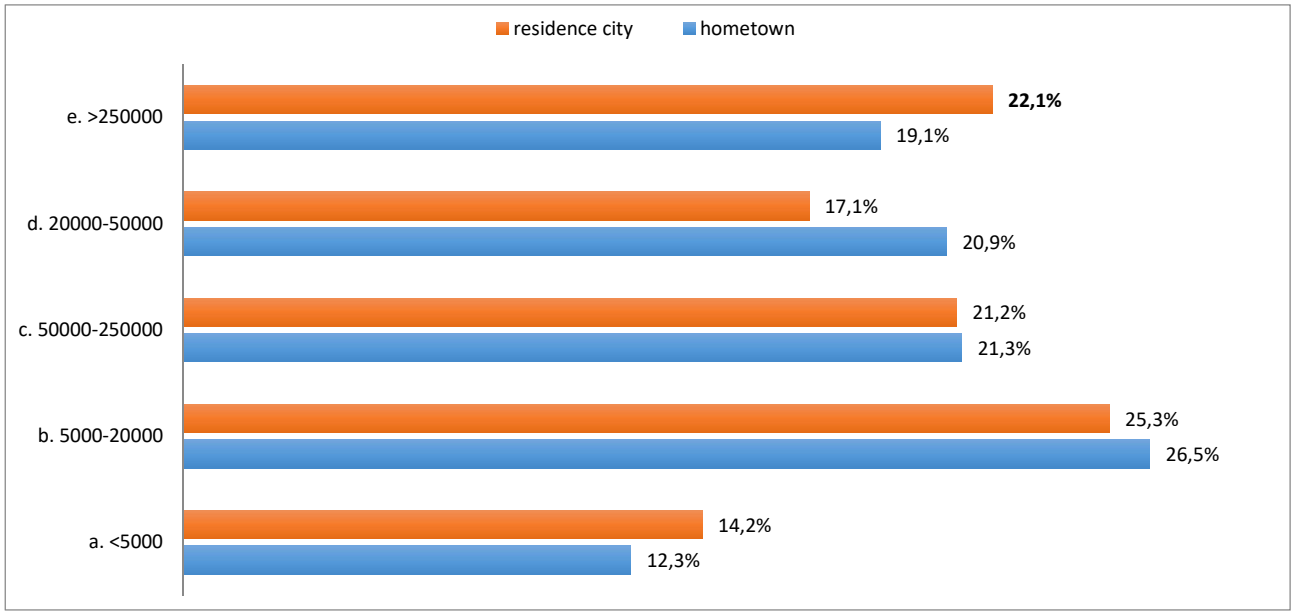


Fig. 8 – Respondents distribution by hometown and residence city population size class

The widespread diffusion of social media guarantees the geographical *equidistribution* of usernames extrapolated with the web-scraping software, overall the Italian municipalities, independently of their population size. This characteristic represents one of the most important and appreciable advantages of social research conducted on, and with, social media, as already noted in chapter 5. The geographical distribution of OTR respondents confirms this

feature and makes it possible to represent the LGBTQ population(s) by intercepting it on the web but spread all over the Italian countryside.

7.5 Not only civil union: the respondents' relationship status

93.31% of people who answered the OTR survey are unmarried, 3.76% are in a same-sex civil union, introduced in the Italian law in 2016, 2.23% declare to be married and 0.7% divorced. Such a high percentage of unmarried respondents is obviously linked with the young age of respondents, already discussed in paragraph 7.1, but not only: to the question "are you, at the moment, in a relationship?", 42.76% people answered "yes, in a same-gender relationship"; 42.20% answered, "No, I am single". The sample is thus equilibrated, between people involved in a same-sex relationship and singles, as shown in the chart below:

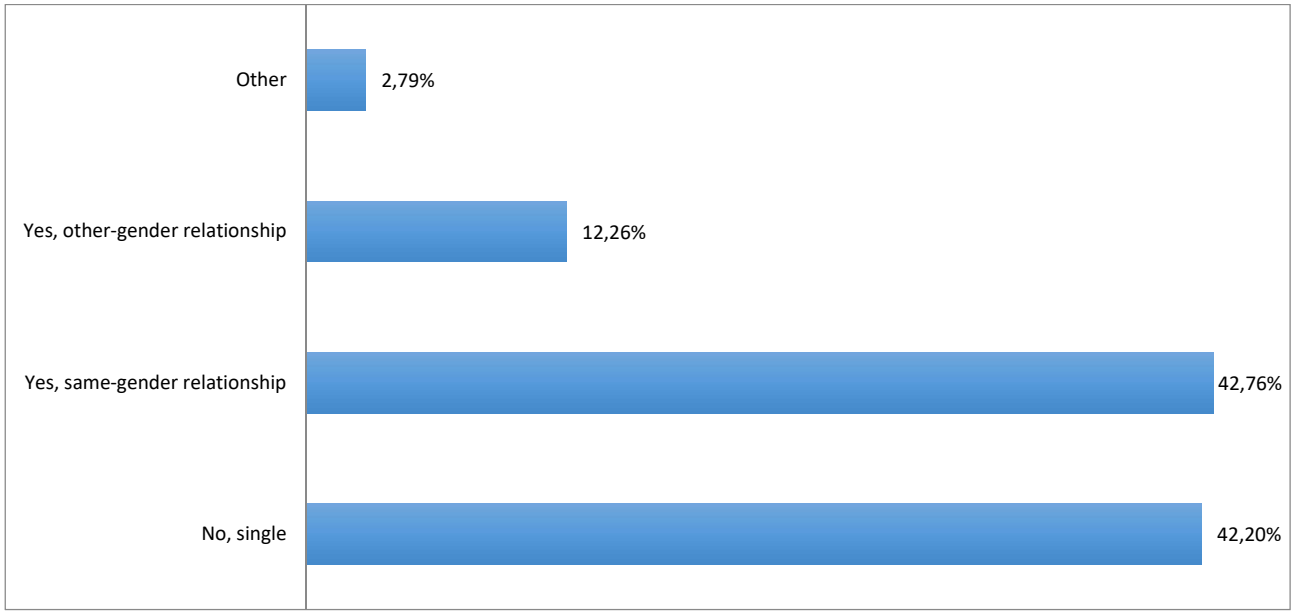


Fig. 9 – Respondents distribution by relationship status

12.26% of respondents declared to be involved in a same-gender relationship. Not necessarily a homosexual relationship (in the next chapter it will be shown that only 8.1% of respondents self-define themselves as heterosexual) but not a strictly homosexual relationship, which shut out pansexual relationships and all the relationships of those who avoid labels, both for people and relationships. Finally, 2.8% of people answered with the open option "other", specifying the

type of relationship in which they are involved: these modalities are often early flirtation, habitually with same-gender people, sometimes open relationships or polyamory. More than once people answered "I am not binary" to the question about the relationship, as if to say that non-binary could be more than a gender identity: non-binary refuses every categorization, also those of relationship status, which oppose the often simplistic dichotomy single Vs. engaged.

The relationship status of interviewed people, jointly with their age and their civil status, strongly influences their living situation (who do you live with?) as shown in the following chart:

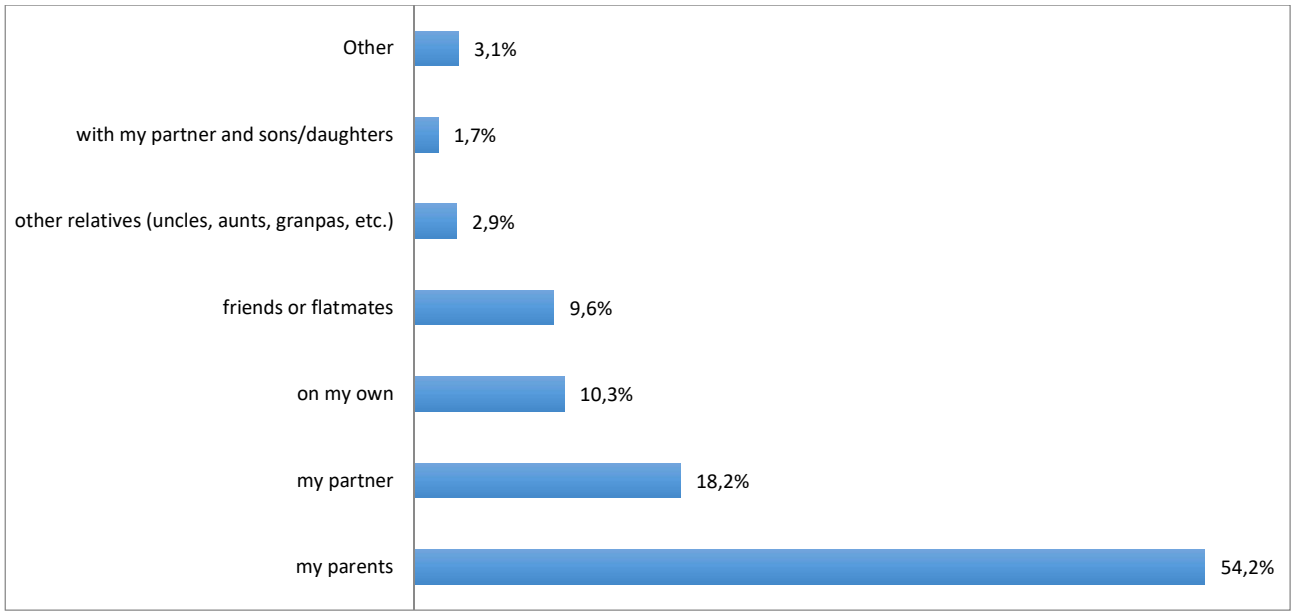


Fig. 10 – Respondents distribution by who they live with

More than half of the respondents still live with their parents (54.18%). 18.25% lives with their partner, 10.31% live on their own, as many share a flat with friends or roommates. The remainder of people lives with other relatives (2.92%, often brothers and sisters), with their partner and sons or daughters (1.67%) or others (3.06%).

From what we have seen, it can be said that the survey intercepted many unmarried people, who usually are still living with parents and who still have not contracted civil union.

This characteristic could be really useful for social research on this issue, because one of the most important surveys that are going to be planned in Italy about LGBTQ, will be the Istat census of civil unions (paragraph 4.3). This survey

will accurately represent the situation of all those same-sex couples who have united, according to Italian law since 2016. However, this census will structurally overlook who is still unmarried, both because they do not want to, or because they cannot get united, for many reasons.

The OTR sample (although with all its weaknesses, given by the auto-selection of users who come out on Instagram and by their use of certain hashtags) caught an extremely variegated target population(s), which would stay out of most surveys about LGBTQ subject, because of their age or because they have not yet united.

From what we saw, it can be said that OTR respondents are often young, sometimes very young, usually still living at home with their parents. They come from all over Italy, but they more frequently tend to live in the metropolitan cities of the central northern Italian regions. The study, is aspiring to a high-school diploma or a degree, and/or they work, almost homogeneously across the many labour market sectors. Some are single, others are in a relationship. But all of them wanted to collaborate with this research project, entirely conducted on Instagram, a social space initially projected for other purposes, but which turns out to be valid and efficient also for social researching.

Unfortunately, geographical and age information cannot be crossed, as stratification variables, because the moderated sample size, which entails privacy and representativeness problems: analysing the survey variables, coming from the questionnaire, by age classes, regions and type of municipalities, of course, would allow understanding many sociodemographic differences about the LGBTQ issues; however, splitting the sample size in all these covariate patterns would generate too low sample sizes, and empty cells too. Generally, in social research, and statistics, the estimation level is a compromise issue between informative needs and privacy and representativeness constraints, which in this research project are really strict, due both to the sensitive questions and to the number of respondents.

8 Self-perception: to define, or not to define, themselves

As introduced in the first chapter, gender studies, starting in the 70s, have redefined and articulated the definitions of biological sex, gender identity, gender expression and sexual orientation. First gender studies (Butler, 1990; Haraway, 1991, 1997; Barad, 2003, 2007), and Queer studies (Hughes, 2006; Ahmed, 2006) to follow, have deeply elucidated these complex concepts far beyond the existing dichotomies, which none the less still resist in our societies. According to Barbagli and Colombo (2001, 2007), these self-definitions have huge sociological importance: «they allow [to LGBTQ people] to integrate needs, experiences, ideas, to find a place in the social space and to make sense out of people's life. At the same time, sexual identities end up prevailing on all other identities: social class, job, ethnicity, religion, citizenship, etc. Many researchers have wondered what these identities are: are they the expression of the real human nature of a person, or are they just illusions? Moreover, are they cages which imprison and choke all of us, denying every individual difference, or are they life opportunities and personal stability conditions? ».

To answer these complex questions, we asked the 718 respondents of the OTR survey to define themselves. Or to not do so. Giving them some default classifications to choose from in the questions about biological sex, gender identity and sexual orientation, but always leaving an open modality, which gave everyone to the possibility to describe themselves in the freest, accurate, personal and sensitive way.

8.1 Biological sex: the third option

At the end of the socio-demographic part of the questionnaire, deeply analyzed in the previous chapter, sex was asked, as it happens in most social surveys, but focusing on biological sex. This label, as already seen in chapter 1, is given at birth based on medical factors, including hormones, chromosomes, and genitals. Biological sex can be described according to the two most common names, male or female, but also with a third option: intersexual. Intersex people, often neglected, have body characteristics that fall outside the strict male/female

phenotypic binary. There are many different reasons why one is considered intersexual, such as peculiarities in chromosomes, hormones, or genitals characteristics. The intersex population size estimation is actually a daunting task, both due to the vastness of the definition and to the characteristics which define this hidden (sub)population. In 2000, Blackless, Fausto-Sterling et al., estimated that 1.7% of human births might be, in some way, intersex. This number was strongly criticized but many scholars and academic researchers (as it often happens when gender studies cross other disciplines, medicine in this case): it was judged too high for a "residual" biologic phenomenon. Despite awareness of the existence of intersexuality, acknowledges during the past 20 years, still few surveys give the possibility to answer questions about biological sex an answer which is not male or female. In the OTR survey biological sex has been asked according to 3 modalities: male, female, intersex, as shown in the pie chart below:

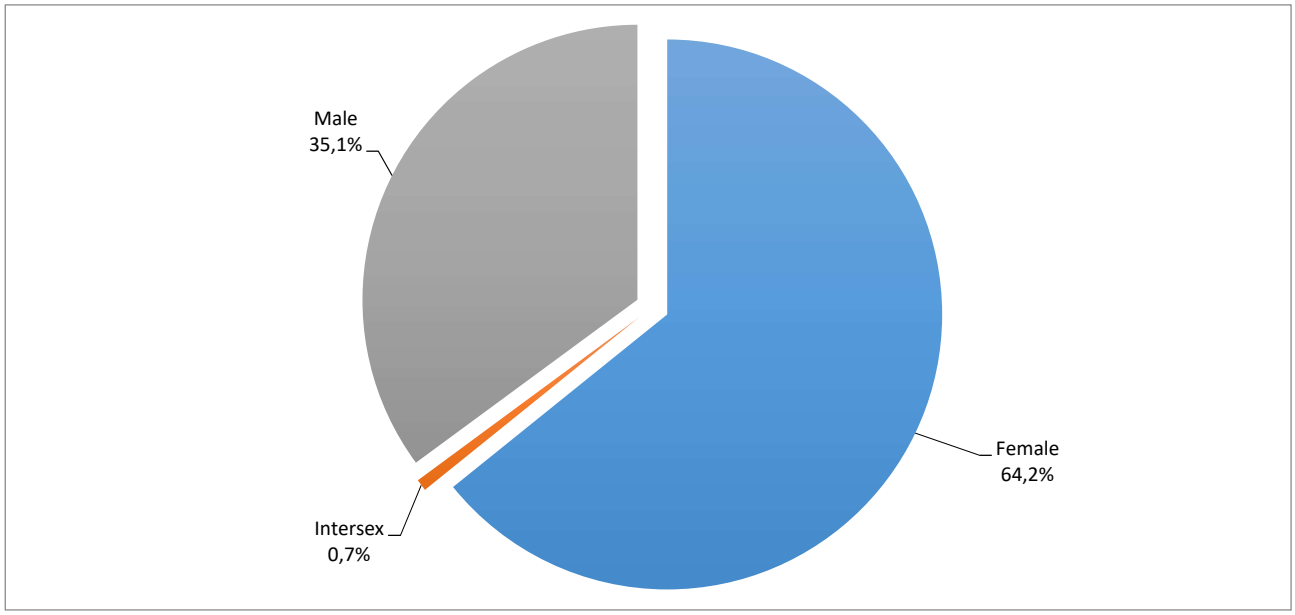


Fig. 11 – Respondents distribution by biological sex

The majority of respondents (64.3%) self-define their own birth sex as female. Male respondents are 35.1%. Intersexual respondents are 0.7%.

In absolute terms, 5 intersex people on 718 respondents do not represent a negligible number of people, although related to a very rare phenotypic phenomenon. OTR sample, analysed by biological sex, is unbalanced, however, this is not necessarily a survey weakness: studying a hidden population, on which it is impossible to strictly make an inference, we can consider the higher

collaboration of women as important feedback about data collection, but also about the web-scraping phase. Indeed, as it has been previously noted most of the posts collected by the web-scraping procedure, according to the hashtags contained, were those relative to male keywords, such as #gayitalia, #gayitaly, etc. (see Tab.1). However, most of the respondents were females. This difference, between the theoretical sample and the realised one, points out a higher collaboration from women than from men. In addition, it can be interpreted as higher use of the other hashtags collected (#lgbtitalia, #lgbtqitalia, #lgbtyitaly, etc.) by women than by men; as if to say, non-heterosexual men, define themselves as gay, non-heterosexual women define themselves as part of the LGBTQ community, at least depending on their use of keywords. This deduction could be explained based on more variegated gender identity of women, as will be illustrated in the next paragraph.

8.2 Gender identity: who people feel to be

In the first chapter, we saw how gender studies focus on the differences between the concepts of biological sex, gender identity e sexual orientation (Zevallos, 2014), definition matter on which we based our target population recognition.

After having asked, to interviewed people, about their biological sex, gender identity is asked to understand how society defines people and how individuals define their identities, being male, female, transgender, queer, and all the other gender definitions. Many other gender options, which increase without ever being exhaustive:

A recent Australian survey about sexuality, conducted by the Queensland University of Technology (QUT), has listed 33 different options as the answer to the question "Which of the following terms do you feel best describes your gender?". The options were: *female, male, transgender man, transgender woman, trans person, trans man, trans woman, female to male, male to female, transsexual, cisgender, cis female, cis male, gender non-conforming, none gender, non-binary, neutrois, genderfluid, genderqueer, demigender, demigirl, demiboy, agender, intergender, intersex, pangender, poligender, omnigender, bigender, androgyne, androgyny, third gender, trigender*. Regardless the QUT researchers had to include an open option to leave people free to describe their gender identity in a text box.

In the OTR survey, we preferred a concise format, mainly because the survey is mobile-oriented and so it should be easily accessible, also through the smartphone. Answer options proposed were precisely

- male,
- female,
- transsexual,
- queer,
- open text form.

This open modality has been deliberately considered to guarantee everyone the maximum freedom level of self-definition, but also to understand if the proposed modalities could be enough or if they leave unexpressed a part of the target population. It should be expressly specified that considering transsexuality as a gender identity could be considered a bit of a stretch or even a mistake. Anyway, this choice was a compromise between the questionnaire conciseness and the quality of the information collected: indeed, a transsexual person could reasonably define its gender identity with the one they hold at the end of their transition, without declaring that they are transsexual. This issue was one of the thorny arguments of semi-structured interviews carried out during the drafting phase: an interviewed person FtoM suggested considering transsexuality as gender identity, in light of the long and complex path undertaken by transsexual people which could likely, define themselves in this way, also after the end of their transition.

We explicitly referred to transsexual people in the answer options, and not to transgender people, because it could be easier to have estimation benchmarks from other surveys on this population, mainly obtained by national health system data. Transsexual people are those who psychologically feel that they belong to the opposite sex, and carry out a transition from one sex to another, usually through gender reassignment surgery and hormone therapy. Transgender people, on the contrary, prefer not to carry out any invasive transition, but still identify as a different gender. Finally, Trans is an abbreviation that usually includes both transgender and transsexual people.

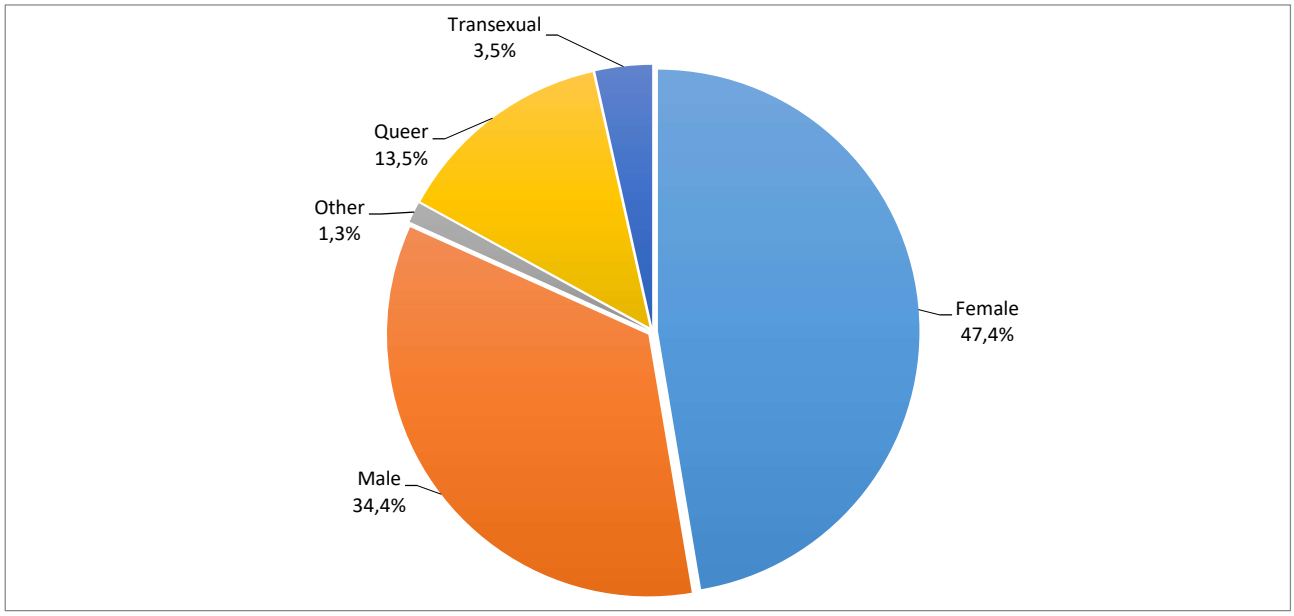


Fig. 12 – Respondents distribution by gender identity

The difference between the concept of sex and gender emerges in its full force comparing the two last pie-charts: 14.8% of respondents do not categorize its gender as female, nor as male, preferring the modality "queer". Where 3.5% carried out (or is carrying out) a transition procedure of sexual reassignment. 1.3% gives qualitative answers attributable to the "other" modality, which include uncertain responses (e.g. "I do not know"), hybrid (such as "androgynous woman", "female man"), and finally categorical refusal of pigeonholing the gender identity.

The male gender identity percentage (34.4%) fits with the male biological sex percentage (35.1%), but not all of them are cisgender. Respondents gender identity often does not coincide with their own sex. The queer umbrella term collects all those people who refuse the bipartition male/female (Macintosh and Bryson, 2007; Marcus, 2005). These people declared more frequently female biological sex (10.03%) than male (3.2%), as shown in the following graph:

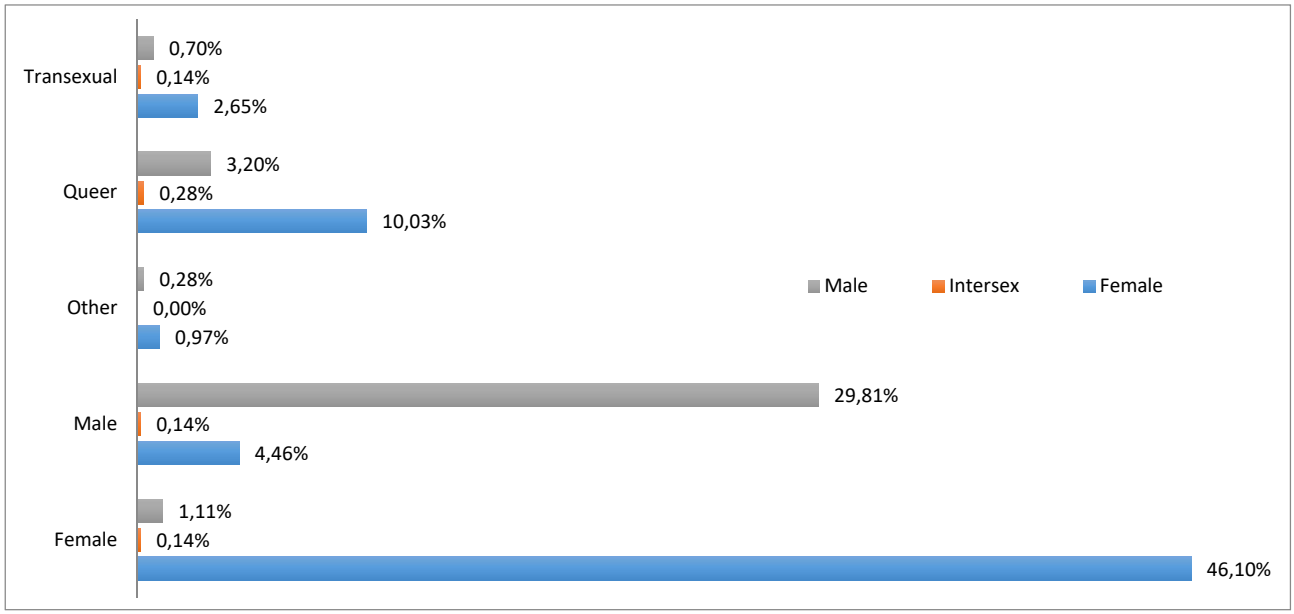


Fig. 13 – Respondents distribution by gender identity and biological sex

The chart shows respondents percentage distribution of gender identities by biological sex. 4.46% respondents who feel male were born woman. On the contrary, 1.11% of those who feel female were born man. We can define this 5.57% respondents definitely transgender people, maybe transsexuals. Indeed, 3.5% recognize its own gender identity as transsexual, respectively 0.7% MtoF, 2.65% FtoM, 0.14% was born intersex.

Cisgender respondents represent 29.81% of male identities and 46.1% of female ones.

From what has been seen and experienced so far we can understand how often the gender role can diverge by the biological and phenotype characteristics of sex (Connel, 2013). Gender identity is manifold, complex, fluid and often interviewed people refuse standard definitions, often preferring the queer option or other modalities which constitutes a "caesura" with the traditional polarised contraposition between male and female.

8.3 A multitude of sexual orientations to define who people love

After asking about biological sex and gender identity, sexual orientation is asked to OTR survey respondents. Some "default" modalities have been proposed, honouring what literature on this subject suggests but also taking into consideration what emerged during the semi-structured interviews carried out.

The result is a wide range of sexual orientations, which depicts the idea of how wide the classification of a person's pattern of emotional, romantic, and sexual attraction to other people can be.

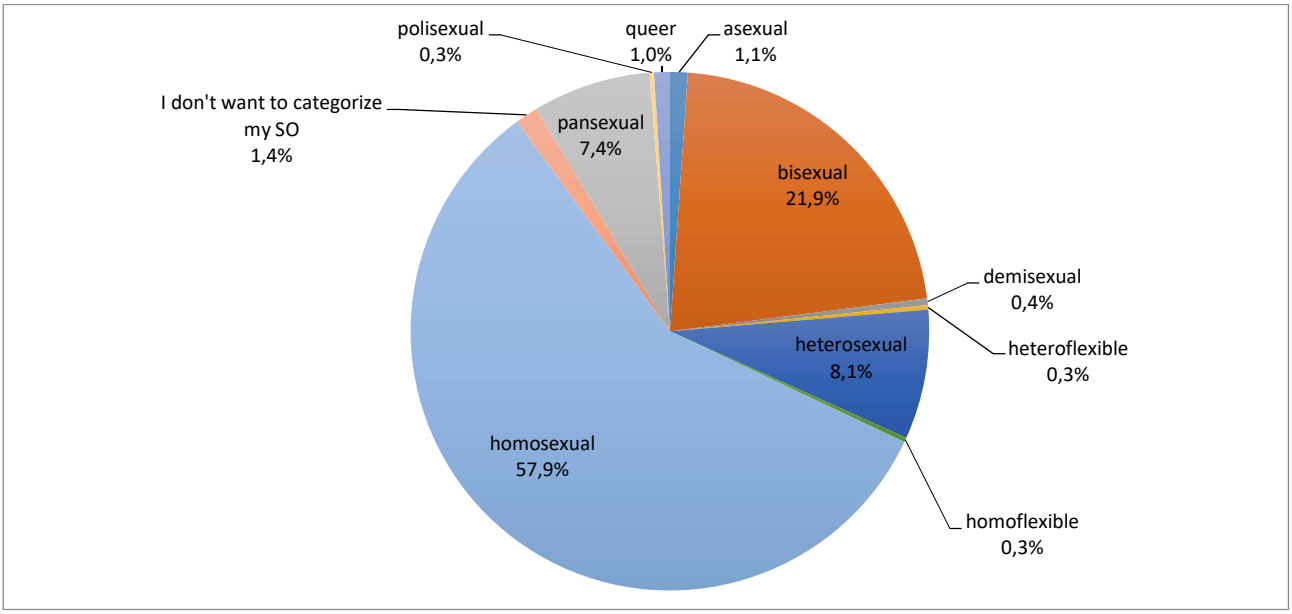


Fig. 14 – Respondents distribution by sexual orientation

The pie chart shows how many sexual orientation modalities were included in the questionnaire, through the open modality. In addition to the 57.9% of homosexuals, 21.9% of bisexuals, 8.1% of heterosexuals, 7.4% of pansexuals, interviewed people also defined themselves as queer, demisexual, heteroflexible, homoflexible, asexual or, in 1.4% of cases, people expressly did not want to categorize their orientation in any way.

The chart also shows how widespread non-monosexual orientation is, firstly bisexuality and pansexuality, although they still result hard to fully understand because of the binary classification by which sexuality is often and mistakenly studied. Bisexual, pansexual or demisexual people, although they represent a numerically relevant cut of the LGBTQ population(s) (also in this sample), still often consider a "niche of a niche". Besides, they are sometimes accused, both by the homosexual community and by heterosexual people, of being "non-declared homosexuals or heterosexual people who want to experiment with new situations" (Matsick and Rubin, 2018). These behaviours generate a particular stigma, generically called *biphobia*, which tends to make invisible the bisexual component of the LGBTQ community, branding bisexual, pansexual and demisexual people of being confused, indolent or inadequate (Klein, 1993).

However, the pie chart points out that bisexuality (21.9%), but also pansexuality (7.4%), does not represent a residual part of the community. Lots of people deliberately choose to not set any boundaries to the genders of people there are willing to be intimate with.

The chart below analyses the percentage distribution of sexual orientation of respondents by biological sex, so that we can understand if the distribution of sexual orientations is independent or not among men and women:

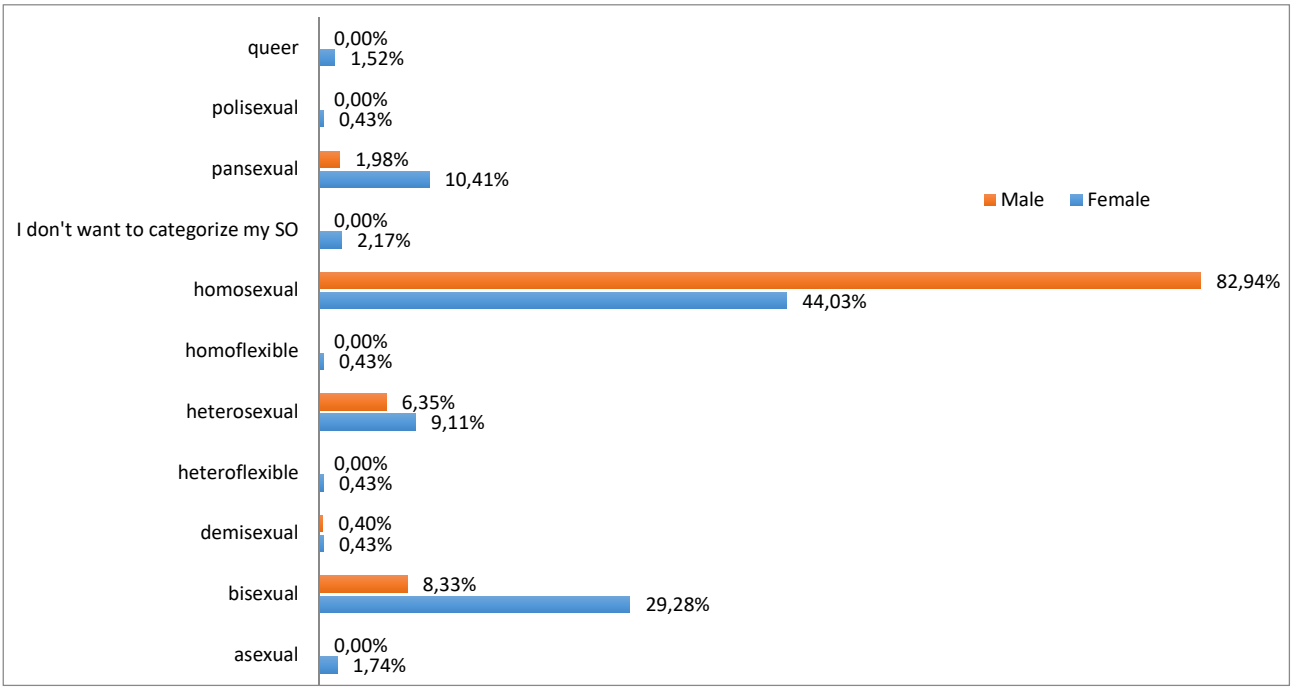
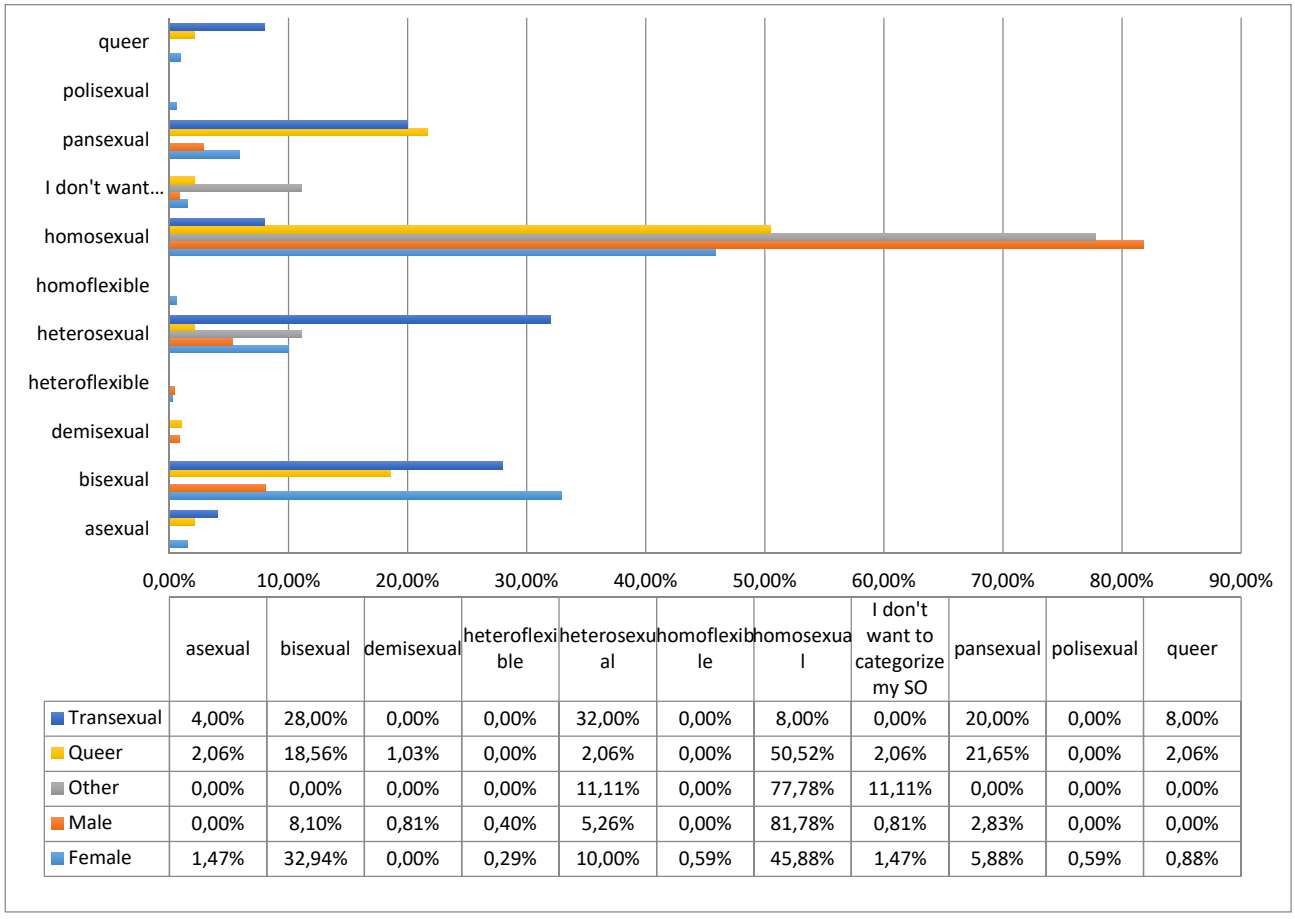


Fig. 15 – Respondents distribution by sexual orientation, and biological sex

Survey respondents, identified according to their use of Instagram hashtags, define their sexual orientation depending on their biological sex: with an equal number of men and women (both male and female percentages sum to 100%), men define themselves more frequently as homosexuals, while women choose often pansexual and bisexual answer alternatives. Heterosexual people[20] were usually born women, which can mean that they actively support the LGBTQ community, through their posts, as a kind of social media endorsement, or that they could have a gender identity differing from their biological sex (e.g. FtoM).

[20] From now on, all the analysis carried out, will be referred to non-heterosexual AND non-cisgender people, to focus only on target LGBTQ population(s); except for the analysis reported in Figures 30 and 31, in which the sense of belonging to a community to which, by definition, heterosexual-cisgender people do not belong

Again, among females, there are no male respondents who define themselves as asexual, queer, or who expressly do not want to declare their own orientation. Women seem to have a more multifaceted sexual orientation than the common binary homosexuality/heterosexuality. To confirm these observations, we can see the percentage distribution of sexual orientation, by gender identity:



	asexual	bisexual	demisexual	heteroflexible	heterosexual	homoflexible	homosexual	I don't want to categorize my SO	pansexual	polisexual	queer
Transexual	4,00%	28,00%	0,00%	0,00%	32,00%	0,00%	8,00%	0,00%	20,00%	0,00%	8,00%
Queer	2,06%	18,56%	1,03%	0,00%	2,06%	0,00%	50,52%	2,06%	21,65%	0,00%	2,06%
Other	0,00%	0,00%	0,00%	0,00%	11,11%	0,00%	77,78%	11,11%	0,00%	0,00%	0,00%
Male	0,00%	8,10%	0,81%	0,40%	5,26%	0,00%	81,78%	0,81%	2,83%	0,00%	0,00%
Female	1,47%	32,94%	0,00%	0,29%	10,00%	0,59%	45,88%	1,47%	5,88%	0,59%	0,88%

Fig. 16 – Respondents distribution by sexual orientation, and gender identity

Also at an equal number of gender (every identity sums to 100%), it is clear that homosexual people are more frequently male (81.78%) than female (45.88%).

Among transsexual people heterosexuality is the most common orientation (32%), so they are attracted by people having a different gender from the one they have acquired. Although, there are also transsexual people who define themselves as bisexual (28%) and pansexual (20%); the latter modality means an attraction towards people, regardless of their sex or gender identity, so towards other transsexual people as well, both MtoF and FtoM.

Analogously, people who refer to themselves as queer profess all the considered sexual orientations, including heterosexuality. The term queer

represents in this way, not only a gender identity but also a definition with a strong political value, which characterizes the rift with the male hetero-normal hegemony, even though it sometimes does not result in a non-heterosexual orientation (Marcu, 2005; Ahmed, 2006).

8.4 Coming-out: unveiling oneself in real life and on social media

Having intercepted the target population of the survey on social media, through the content of posts, makes this survey particular from the coming-out point of view as well.

Having posted a picture with an LGBTQ-related hashtag allows us to assume that the non-heterosexual orientation of a said person is known to most. In this context, the "social network coming out" becomes the last step of the progressive and concentric revealing process, which begins with the closest people, relatives and friends, then it enlarges to colleagues and schoolmates, until it gets to acquaintances and, finally, followers and virtual friendships (Norton, 2016).

In the questionnaire, we asked, those who declared a non-heterosexual orientation, who was the very first person which they told about their own sexual orientation. Barbagli and Colombo, in their research project "Omosessuali Moderni", concluded that the person to talk about such a sensitive issue is chosen according to two main and opposite criteria: from one side, who comes out looks for a close, dear and kind interlocutor; on the other side, those coming out do not want to ruin fundamental and irreplaceable affective relationships. Consequently, "chosen relationships", main friends, are often the first people chosen to come out with. This statement is confirmed by data coming from the OTR survey, illustrated in the graph below:

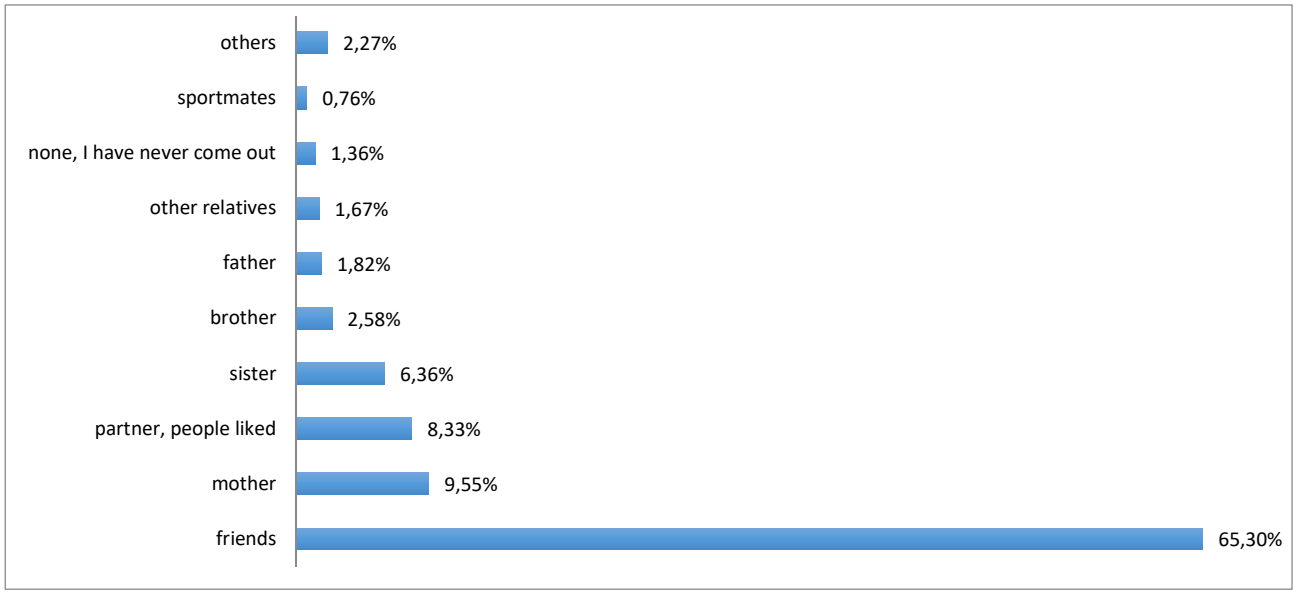


Fig. 17 – Respondents distribution by first person they came out with

The graph shows a strong predilection to firstly come out with friends (65.3%). The factors influencing this choice are easily recognizable in same age groups, in the profound knowledge of the interlocutor, mainly in youth, but also in the multitude of friends which oppose to the singularity of traditional family figures (one mother, one father, and often one brother and/or one sister).

Moreover, it emerges that 8.3% of respondents came out directly to the first same-gender loved person, declaring the availability of a homosexual and homo-affective relationship.

Focusing back to the household, it is important to note a certain gender disparity: people come out more often with their mothers (9.5%) than with their fathers (1.8%), with their sisters (6.4%) than with their brothers (2.6%), with female cousins than with male cousins, with grandmas than with grandpas (in other). Women are though the preferred interlocutors for the very first coming-out, likely due to their more sensitive attitude or, maybe, for a lesser vehemence of reactions.

To the following question "Who knows your sexual orientation today?", 98.6% answered "everyone". It confirms the truthfulness of the social network coming-out process: people intercepted through web-scraping effectively belong to the targeted population and, although young, they have already faced the coming-out phase, first in daily life, then on the web.

Social media, which will be deeply analysed in the next chapter, works both as an interlocutor and as a "sounding board" to unveil themselves to the world,

to meet similar people, to build up a community, and to approach people far away, otherwise difficult to get in touch with, in real life.

8.5 Self-perception according to gender study concepts

In the last question of this part of the questionnaire, a Likert-scale question, about self-perception of gender identity and sexual orientation, was proposed, to understand how the LGBTQ population(s) lives these concepts. The following graph shows the percentage distribution of agreement (and disagreement) for each statement regarding the self-perception issue.

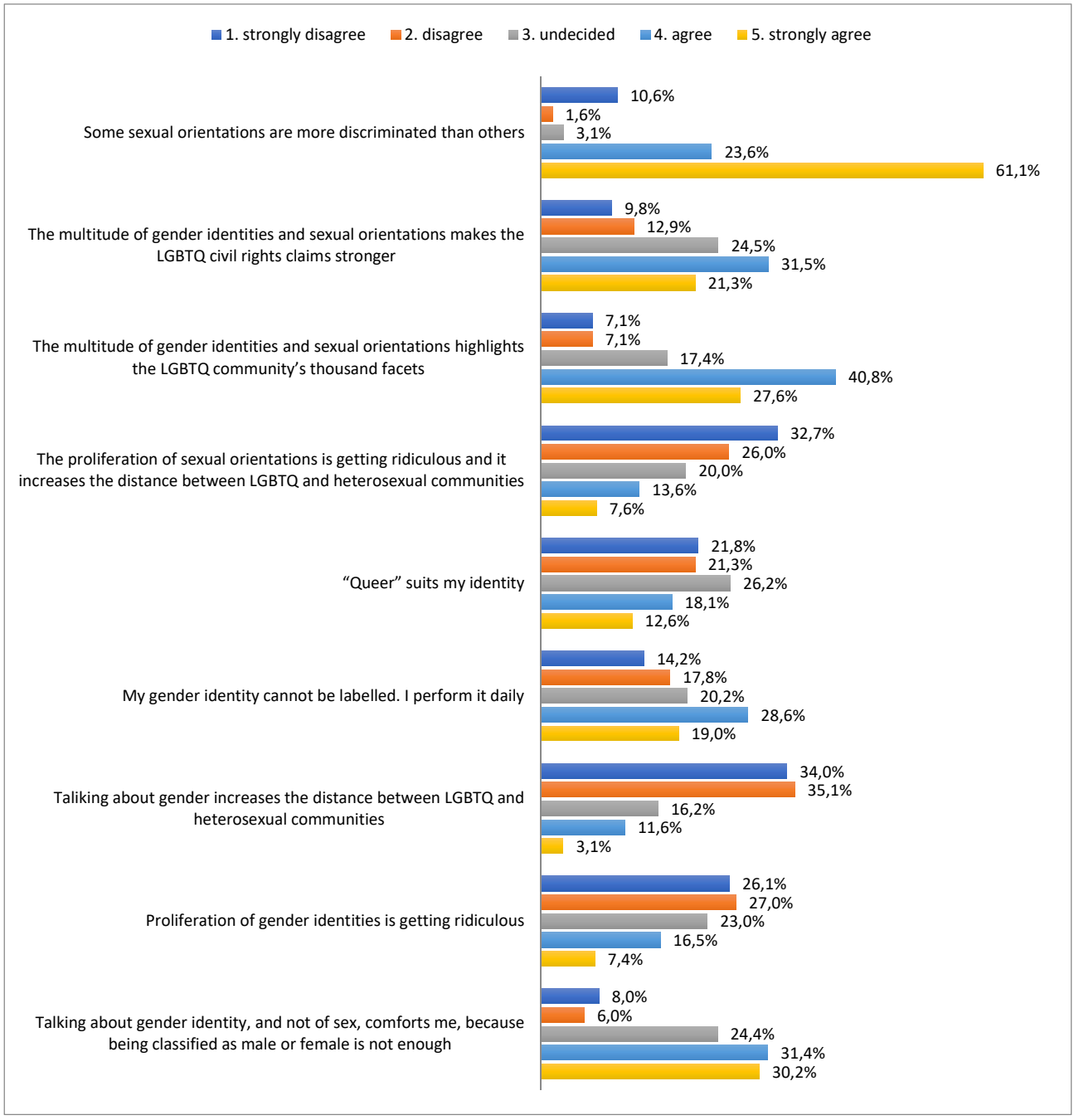


Fig. 18 – Respondents distribution, by agreement scale, about self-perception statements

The vast majority of respondents (84.7%) agrees that some LGBTQ sub-communities are more discriminated than others. It likely means that most of LGBTQ respondents understand that the occurrence of discriminating behaviours and events can be very varying into the "community": for instance, analysing discrimination through an intersectional approach we could assume that lesbians may be more discriminated than gay men, but at the same time, we could also suppose that bisexual people suffer from a sort of discrimination also

inside the LGBTQ community (Klein, 1993), as already seen before. Additionally, during the semi-structured interviews carried out before the survey, it emerged that transsexual people run into discrimination problems more often than LGB people, mainly in the working field, simply because of their appearance (Whittle, 2007). This figure suggests that studying sexual discrimination, accurately, is a duty that should be analysed separately, taking into account the single and different LGBTQ sub-populations, to better understand the causes of this phenomena and correctly plan contrast measures against them.

The second statement pointed out that the multitude of identities and sexual orientations, which the community is made up of, not always strengthens the civil rights claim. Indeed, only 52.8% respondents agree with the sentence, confirming an issue that emerged also during interviews: a kind of lack of cohesion, inside the LGBTQ community, has worked against the civil rights claim process, in the last years. Gay men, lesbian women, bisexual people and transgender can fight for a common empowerment process, but they could have also their own stances, which should not slow down the common human and civil rights claim process.

The queer term reflects only 30.7% of respondents, mainly women. 48.6% of respondents say that their identity cannot be labelled and that they prefer to "perform" it daily. The choice of using the verb "perform" is not casual at all, but it explicitly refers to Butler's notion of gender performativity, introduced in Gender Trouble (1990), where she proposes that people should conceive gender not as a set of free-floating attributes, but rather as "an incessant and repeated action of some sort".

From what we have seen so far, also according to the use of hashtags, men define themselves as gay, while women prefer more free, open, and fluid definitions (or "non-definitions") to define their own sexuality and to be part of the community.

According to 69.1% of respondents, Gender studies should not distance the LGBTQ community from heterosexual people. Gender studies should just be considered a theoretical tool to framework civil rights for the people, irrespective of their sex, gender or sexual orientation, and so to address social progress. However, many people oppose the widespread of gender studies. Those people want to defend traditional conceptions of masculinity and femininity, in contraposition to homosexual and transsexual people, but they also want to

contrast women empowerment, the rainbow family and the secularity of the state (Bernini et al., 2015).

Gender studies are not integralist theories that want to subvert the establishment, as their detractors say. LGBTQ have internalized gender studies main concepts, and 61.6% of respondents appreciate talking about gender instead of biological sex, especially if a great part of them are cisgender.

9 Social media, apps and digital technology among LGBTQ population(s)

In chapter 5 we have discussed how social media are used more and more frequently, transversally by many people all over the world, and for lots of different purposes. We have also analysed how social media have increased the showcasing opportunity of the LGBTQ community, helping it to get away from the "shadow cone" in which it was, according to four main aspects, such as socialization, information, community and self-representation (see paragraph 5.3). Furthermore, we have talked about social research and how it could, and should, benefit from the many potentialities of these new data sources, which represent a full-fledged *multi-purpose research tool*. From these methodological reflections, applied to the complex context of hidden populations, the basic idea of the present research project was born: conducting a survey exclusively on Instagram.

To argue in favour of (or against) these theories, a relevant part of the questionnaire has been dedicated to the relationship between the LGBTQ community and digital technologies. In the following paragraphs, we will see which social media are used, by whom, and through which devices. Moreover, treasuring the issues emerged by the semi-structured interviews, we will deeply investigate the spread of dating apps among LGBTQs. We will try to understand how widespread they are, who uses them more, how efficient are they in terms of dating, and for which (other) reasons people started to use them, or, on the contrary, uninstalled them. By doing so, we should be able to understand the strengths and the weaknesses of the methodology used, looking for unexplored details which link digitalization issues, one of our theoretical frameworks, to the LGBTQ population(s), is the target of this research project.

9.1 The impact of social media on the LGBTQ community's daily life: a powerful expression, informative and socialization tool

Social media apps like Facebook, Instagram, and Twitter have had a huge impact on society since smartphones became a widely used technological accessory. These media platforms have revolutionized the lives of all of us. However, literature on this issue suggests that the LGBTQ community mostly profited from the social media spread, to overcome the oblivion in which it had been for a long time, in many different ways (Dara, 2017; Pullen and Cooper, 2010; Ross et al., 2014). The advantages which social media bring to the LGBTQ community can be defined as transversal and multi-purpose, both because they involve all the components of the LGBTQ population(s), and also because they have effects on many different aspects of the online activity of LGBTQ people.

These purposes are, however, not equally pursued among different providers. Each of them has different goals, managers, corporate policy, and target users; so, people use them in different ways, as shown in the graph below:

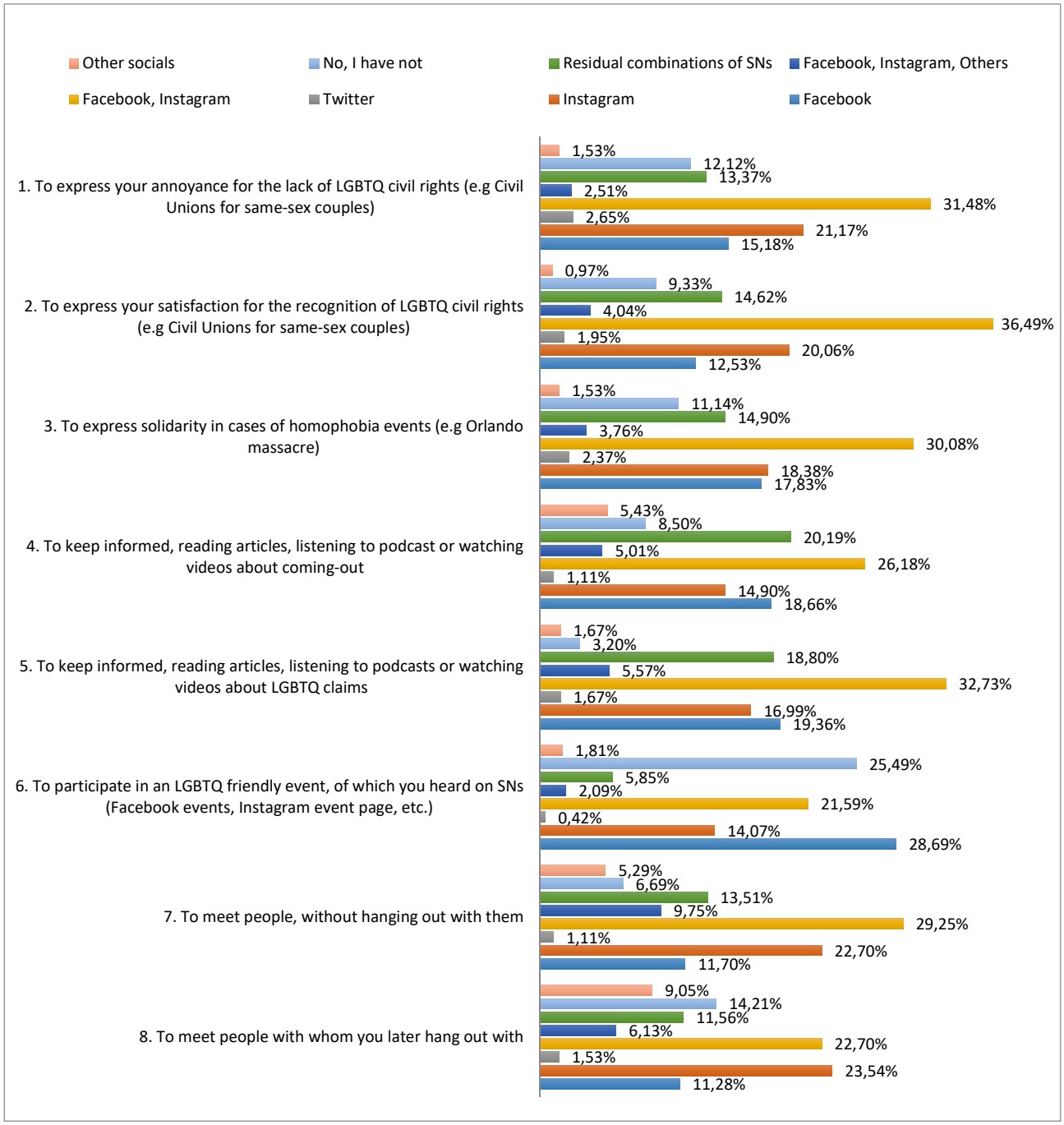


Fig. 19 – Respondents distribution by purposes pursued on every social media platform

Facebook and Instagram are the most widely used social media, both singularly and jointly, but for very different purposes. Every social media has its own user-base, even though they sometimes "overlapping". Facebook is used to read articles and to watch videos (4.: 19%, 5.: 19%), but, mainly, to be informed about LGBTQ-friendly events (6.: 29%). Facebook events, which actually remains

the last and most unique functionality that we can find only on this social media, seems to be a utility very appreciated by interviewed people.

Instagram, on the other hand, is used not only to share pictures but also for meeting new people, both for hanging out or not (8.: 24%, 7.: 23%). Instagram, in fact, among the various social media, is the one that most facilitates meeting new people. The app provides several utilities, such as geo-localization, direct messages and stories. Jointly, Facebook and Instagram are widely used as an endorsement tool, to express annoyance for the lack of LGBTQ civil rights (1.: 32%), or to express satisfaction towards the recognition of these civil rights, for instance, the promulgation of the law about Civil Unions for same-sex couples in Italy (2.: 37%), or, finally, to express solidarity in cases of homophobia events (3.: 30%).

Other social media represent a minority stake, often a residual among social media used for the purposes mentioned above. Twitter, especially, proves to be down, as anticipated in paragraph 5.4. User percentages, of those who use it, if only to read and keep informed, the main purpose of Twitter, are very low (4.: 1%, 5.: 2%) and lower than percentages relative to other social media.

According to several digital sociologists our society is going for the visualization of thought, ideas, concepts and bodies (Lupton, 2005; Marres, 2012) and that could be the reason why Instagram is actually more popular than other social media: it is more visual and less textual than Twitter and Facebook, and it integrates just the positive aspects of them.

From what we have seen so far, expressing themselves and their own ideas (1., 2., 3.) are the social media purposes that users prefer, mainly on Facebook and Instagram. That is why many users consider social media a safe place, where they say what they think, in a constructive, or indignant, or sympathetic way. The social media safety net is confirmed by the fact that most of the interviewed people have a profile that, more or less clearly, shows their sexual orientation, as shown in the graph below:

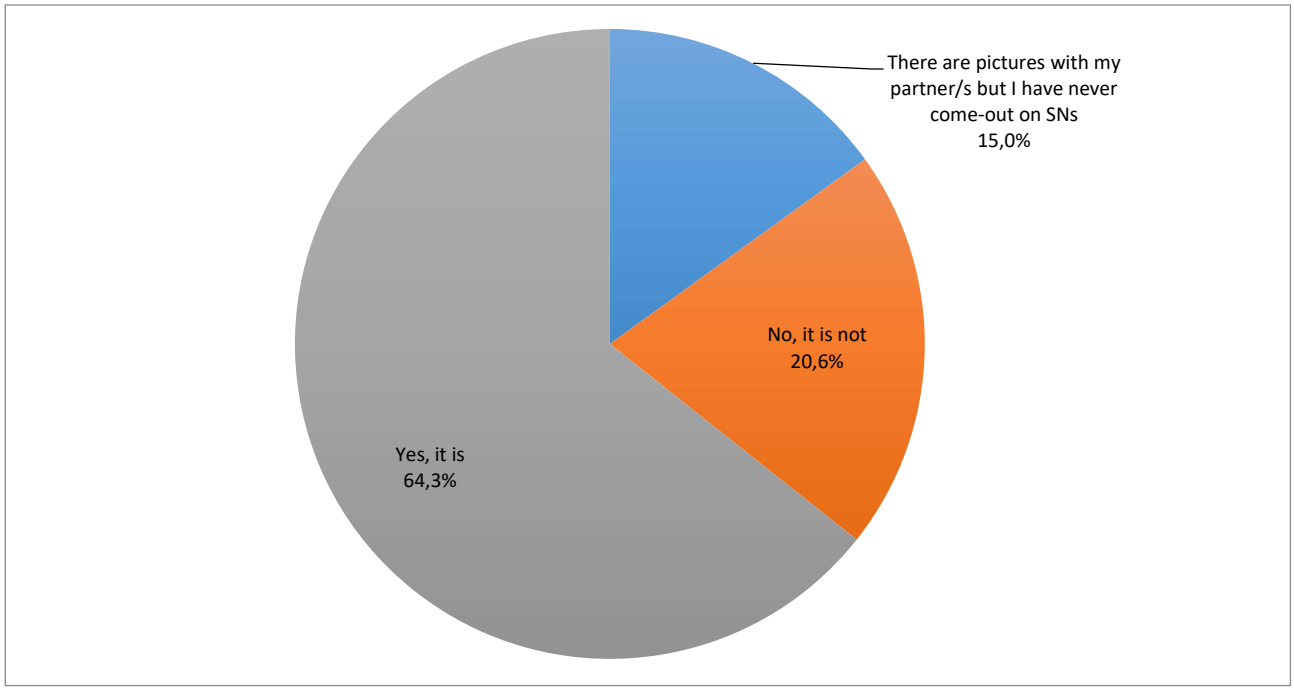


Fig. 20 – Respondents distribution by purposes pursued on every social media platform

64.3% of non-heterosexual people have a deducible sexual orientation by their social media profiles, with the addition of 15% of people that have nevertheless pictures with the partner. This result, somehow, confirms that finding the target population on Instagram could be an efficient methodology: listing people through web-scraping, and asking them for other usernames to interview in turn with the snowballing, could have given a wrong list, but fortunately it did not. Only 8% of interviewed people define themselves as heterosexual (but some of them are transsexuals), and almost 80% of non-heterosexuals have a profile that expresses their sexual orientation. Social network coming-out, discussed in the design phase, is effective and real: who speaks its own sexual orientation, or publically endorses the LGBTQ instances, on the web often already did it in the real life.

Moreover, a study conducted by the journal Vocatin in 2020, even claimed that members of the LGBTQ+ community are more likely to first come out online before formally coming out to friends and family members: 1 in 5 LGBTQ people interviewed came out online, whether that's in a YouTube video, or a Twitter post, or an Instagram story. The results of the study show just how integral of a role social media plays for those in the LGBTQ population(s). While coming out in person can be “nerve-wracking” when people are not sure how friends and family will react, coming out online offers a place to trial-run the conversation.

In particular, for people belonging to Generation Z, which identifies more as LGBTQ and Digital Natives than any other previous age group, social media can be a way to spread the word of an LGBTQ identity and avoid traditional vis-à-vis coming-out. 75% of Gen Z respondents in Vocatin's survey came out on an online platform or closed group first before coming out to their friends or family (Leskin, 2019).

Patrick M. Johnson, in his book named "*Coming Out Queer Online: Identity, Affect, and the Digital Closet*" (2020), asserts that through social media, LGBTQ individuals have sought new ways to forge communities and increase their visibility. This higher visibility provided individual means to seek out and distribute information to help in the coming out process, and this is definitely positive in acting as an intervention for LGBTQ suicide rates and to prevent homophobic behaviours. However, Johnson also underlines negative side effects of coming out online: the author also contends that it has vastly re-centred and prioritized white, cisgender, masculinity, creating potentially dangerous environments for women, transsexual individuals, and gay men who do not meet high standards of masculinity, imposed by society; the latter is a subject really close to Connell's hegemonic masculinity, which we will also broach in the next chapters when talking about discrimination.

In this way, if on one hand technology strongly helps the coming out process, on the other it, unfortunately, foments disparities and new ways of discrimination, that only an intersectional sociologic approach can detect and understand.

Anyway, what we have seen so far confirms that social media is strongly rooted in the LGBTQ community's daily life. This consideration was taken into account during the design phase, and therefore the proposed questionnaire was deliberately mobile-oriented, namely that it can be filled out on smartphones as well to respect privacy, facilitate the interview and reduce refusals.

To evaluate the applied methodology, in the survey, a question asked which device (or devices) do people most often use for social networking:

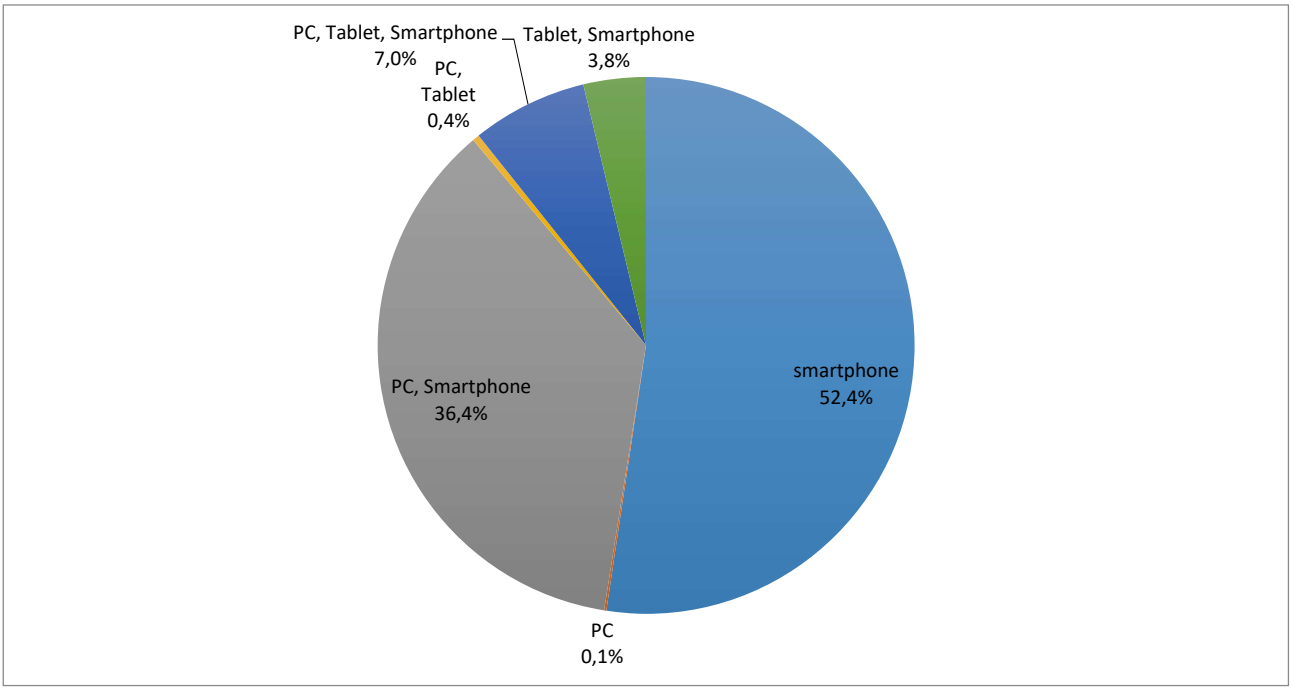


Fig. 21 - Respondents distribution by the device(s) used for social networking

More than half of the respondents only use smartphones, 36.4% use both PCs and smartphones. Tablets are used for social networking only jointly with some other device (11.2%). Only 0.1% of respondents use PCs as a unique device. It confirms the importance of submitting a mobile-oriented survey, characterized by slender questions, simple response options, etc. sometimes the detriment of answer precision. Planning a survey, mainly about the LGBTQ population(s), is a compromise between response burden, namely the statistical disturbance caused to interviewed people, and the meticulousness of the answers (Gobo, 2015).

From what we saw, it can be asserted that social media have revolutionized the daily life of the LGBTQ community, becoming a powerful expression, informative and socialization tool. Not least, social media have radically changed the important and delicate coming out process, sometimes overthrowing the hierarchy of the people to tell their own non-heterosexual orientation. Obviously, all that glitters is not gold, and social media entail negative side effects too. One above all, the possibility of continuing and intensifying discriminative behaviours against the most fragile parts of the LGBTQ population(s), such as transsexual people, men and women who do not respect femininity or masculinity imposed standards, or who have no access to digital technologies.

LGBTQ people, however, are conscious of the bad and dangerous aspects which social media entail, as shown in the graph below:

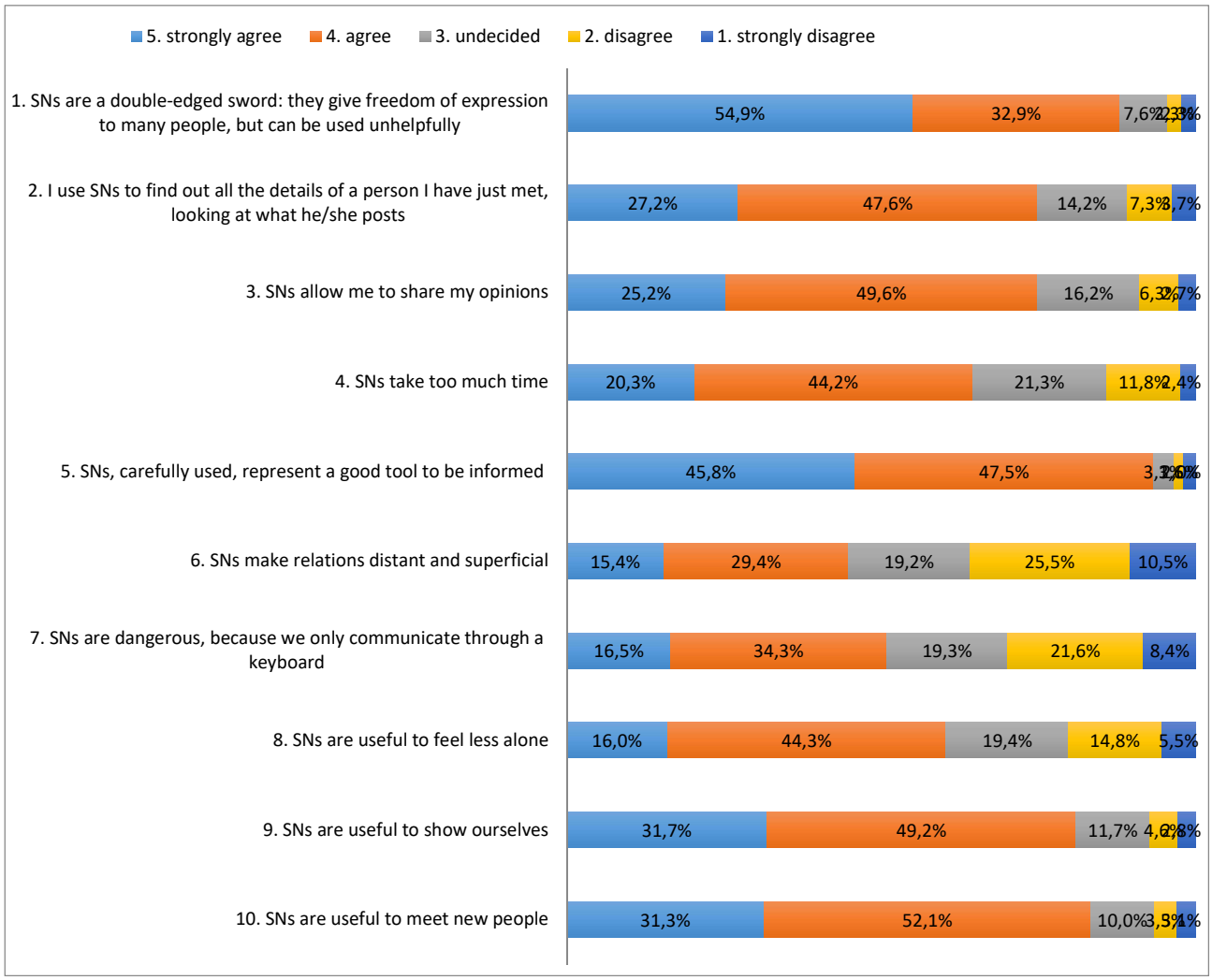


Fig. 22 – Respondents distribution, by agreement scale, about social network (SN) statements

Interviewed people appreciate social media because they are useful to extract information about new people met (2.: 75%), to share opinions (3.: 75%), to keep informed (5.: 93%), and to let somebody show themself (9.: 81%).

Nevertheless, respondents understand the "dual nature" of the tool: 33% of respondents agree, and 55% strongly agree, with the sentence "SNs are a double-edged sword: they give freedom of expression to many people, but it can be used unhelpfully". Moreover, 92% of respondents recognise that social media have to be used carefully to keep informed, without running into fake news, or fraud.

9.2 Gender differences and individuals' safety perception in the use of dating apps

In the last decade, an increasing number of smartphone applications have come to market to facilitate dating, such as (in strict order of the number of users) Tinder, Grindr, Her, OKCupid, Bumble, and many others, less used.

Semi-structured interviews, carried out during the questionnaire drafting phase (par. 6.2), pointed out the crucial importance of dating apps among LGBTQ people. Interviewed people told us that these apps are very wide-ranged, that they have forever changed the LGBTQ way of meeting, for better or worse. Each person interviewed told us their personal experiences with these applications, avoiding enthusiastic or defeatist tones, but just listing their strengths and weaknesses, according to users' biological sex, being cisgender or transgender, and age.

Given the importance of the issue, we asked "have you ever used dating apps?" in the survey:

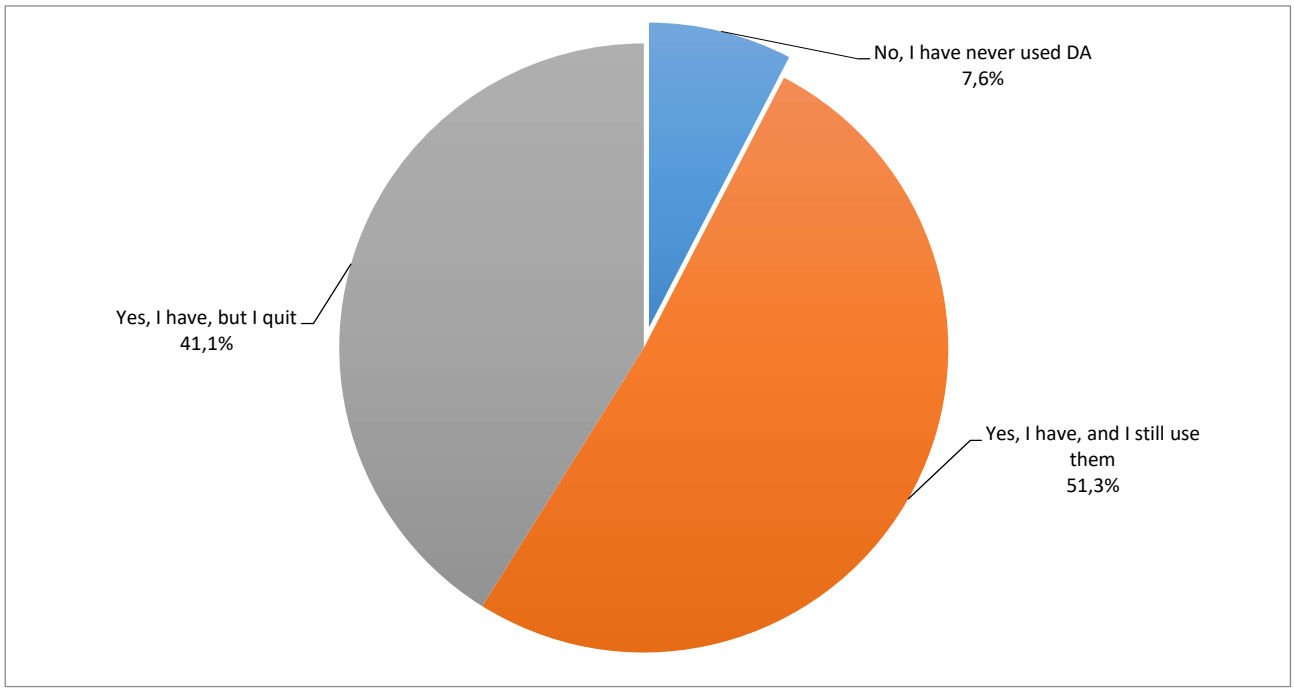


Fig. 23 - Respondents distribution by use of dating apps

71.3% of non-heterosexual respondents have used a dating app at least once.

Several very recent studies have suggested that LGB people are more likely to get a profile on a dating website and to initiate romantic relationships online compared to their heterosexual counterparts (Sumter and Vanderbosch, 2019;

Zervoulis, 2019; Albury et al., 2019; Pew Research, 2020). The lower levels of openness to communicate, and their difficulty in locating partners are the main reasons that literature suggests as the cause of the LGB higher use in mobile dating apps.

Anyway, talking about these applications without taking into account sex, or gender identity, of the people, could be extremely misleading: for instance, the use of these applications is strongly variable between male and female[21], as shown in the pie charts below.

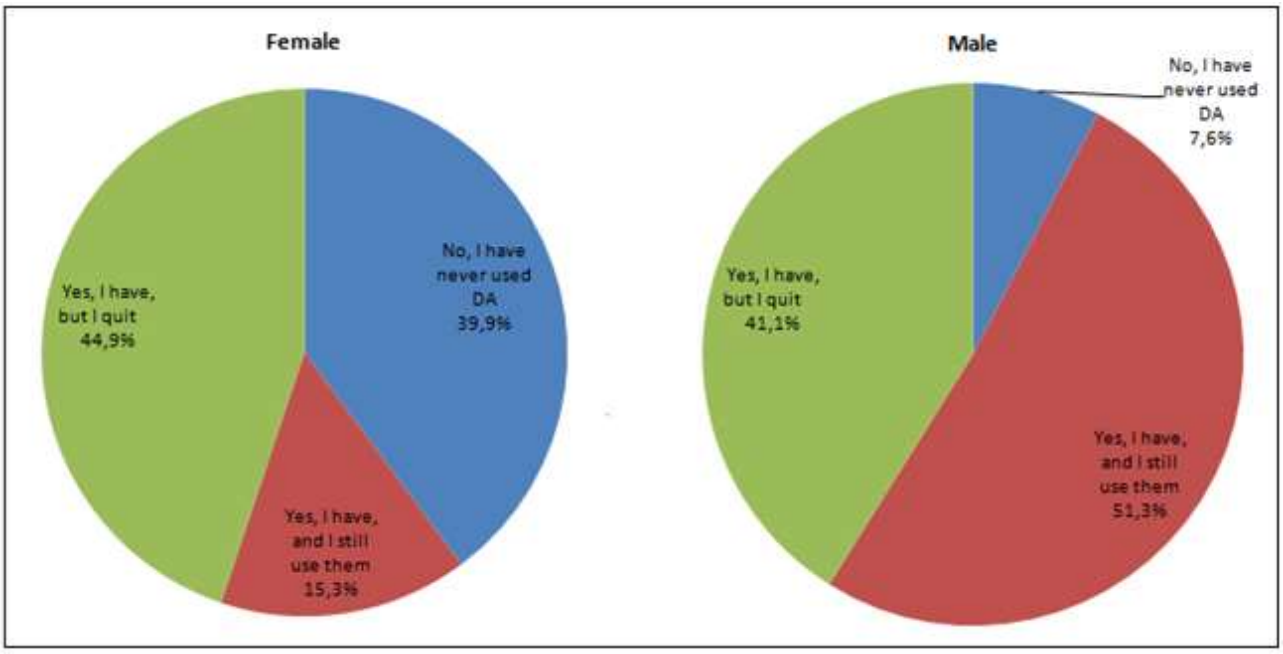


Fig. 24- Respondents distribution by use of dating apps and biological sex

92.4% of men have used a dating app at least once. Though only 60.1% of women have.

A 32% difference can only be statistically significant, mainly observing that percentages of those who quit dating apps are similar (F: 44.9, M: 41.1), but percentages of those who still use them are radically different (F: 15.3%, M: 51.3%).

[21] Analysing the use of dating apps, by biological sex, intersex people has been omitted from this analysis only because of the small sample size of the intersexual population, which could have privacy implications.

There seems to be a very different way of using dating apps, between gay men and lesbian women. A difference not yet investigated, Sumter and Vanderbosch (2019) explicitly state one in their paper: "the literature hints at various relationships between gender, sexual orientation, and dating app usage and motivations: however, for several relationships, empirical evidence is missing".

Many surveys on this subject investigate sex, gender and sexual orientation on dating apps users, but without crossing this demographic information. An intersectional approach, rather, can lead to understanding the real difference of use across the different segments which compose the multifaceted LGBTQ population(s). In the following four pie charts the use of dating apps, by gender identity, is shown:

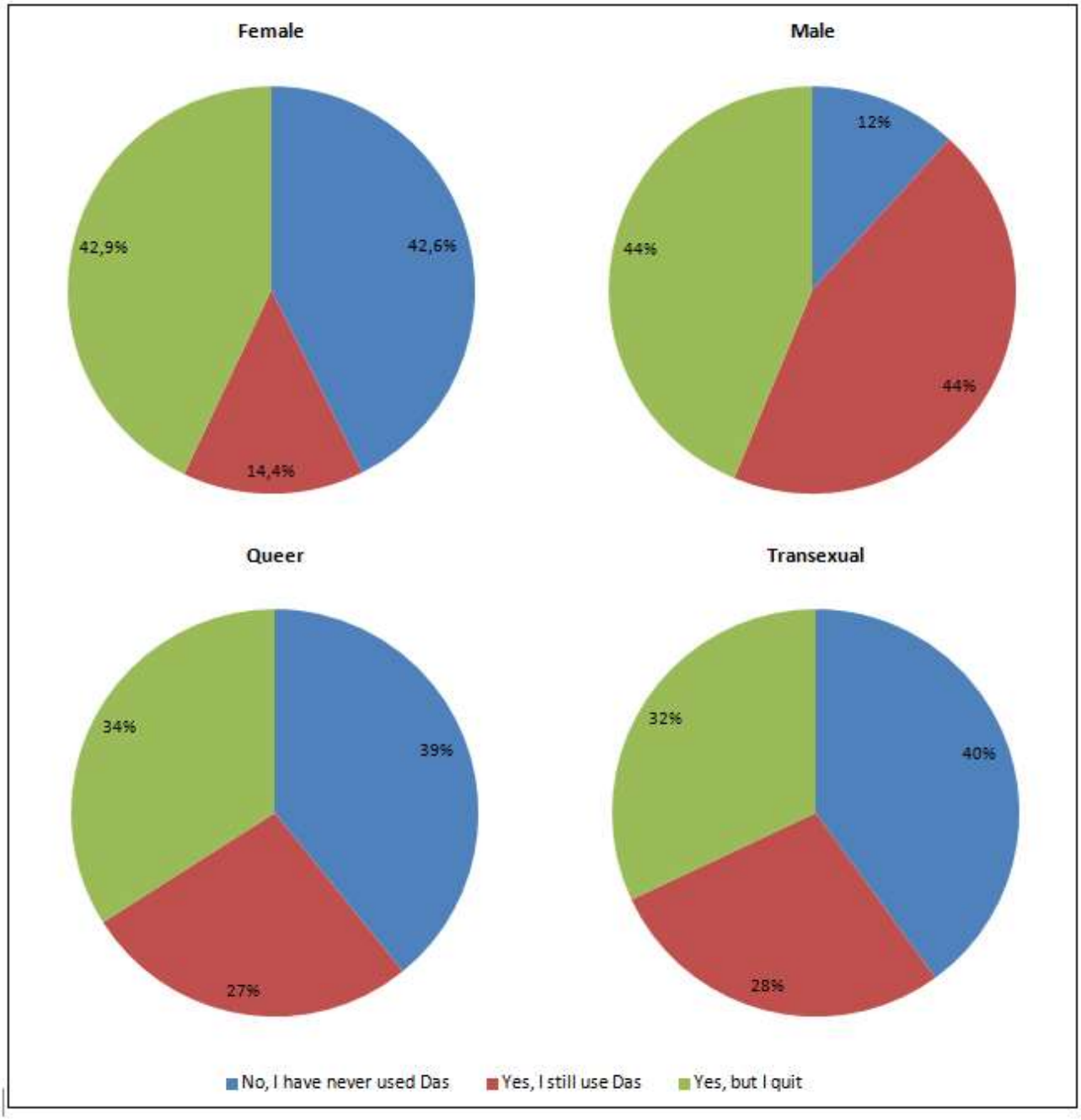


Fig. 25 - Respondents distribution by use of dating apps and gender identity

The question arises: why is the use of dating apps so different between men and women? There could be multiple reasons, and they could also be combined with one another. Beyond all the personal and micro-sociological motives, females do not have as powerful of a tool as their male counterparts: almost all gay men who use a dating app, that is Grindr. Female users are divided into a lot of apps, downloadable on Apple and Play store: firstly, Tinder (largely used by heterosexual people too), but also Her, Wapa, OKCupid, Bumble and many

others. A multitude of apps allocate, in little parts, a smaller catchment area than the male ones, and that therefore, inevitably, make these apps less efficient for female users. The size of a dating apps catchment area is crucial, because, like social media, they only work if having a massive number of users. That is the reason why Grindr represents a monopoly and an institution for the gay community: all male users gravitate around one single app.

Despite what could be thought, queer and transsexual people, instead, use dating apps more frequently than females; moreover, they tend to uninstall these applications less than both female and male respondent people. This result belies those who said that mainstream dating apps were unsuitable for queer needs, because they were "not even capable of properly accommodating non-binary genders". Certainly, dating apps construct gender as a rigid category that has more to do with matching profiles than with identity, but, at the same time, these gender constructions can result functional and helpful to queer people, and not necessarily a discriminating technology.

What is common to all gender identities is the high percentage of those who used dating apps in the past, and then quit them (F: 43%, M: 44%, Q: 34%, T: 32%). These types of applications arouse interest, but they often leave their users dissatisfied. The use of dating apps seems to be discontinued, or even just an attempt. In the survey we asked the reason why these apps left people unsatisfied and why they uninstall them:

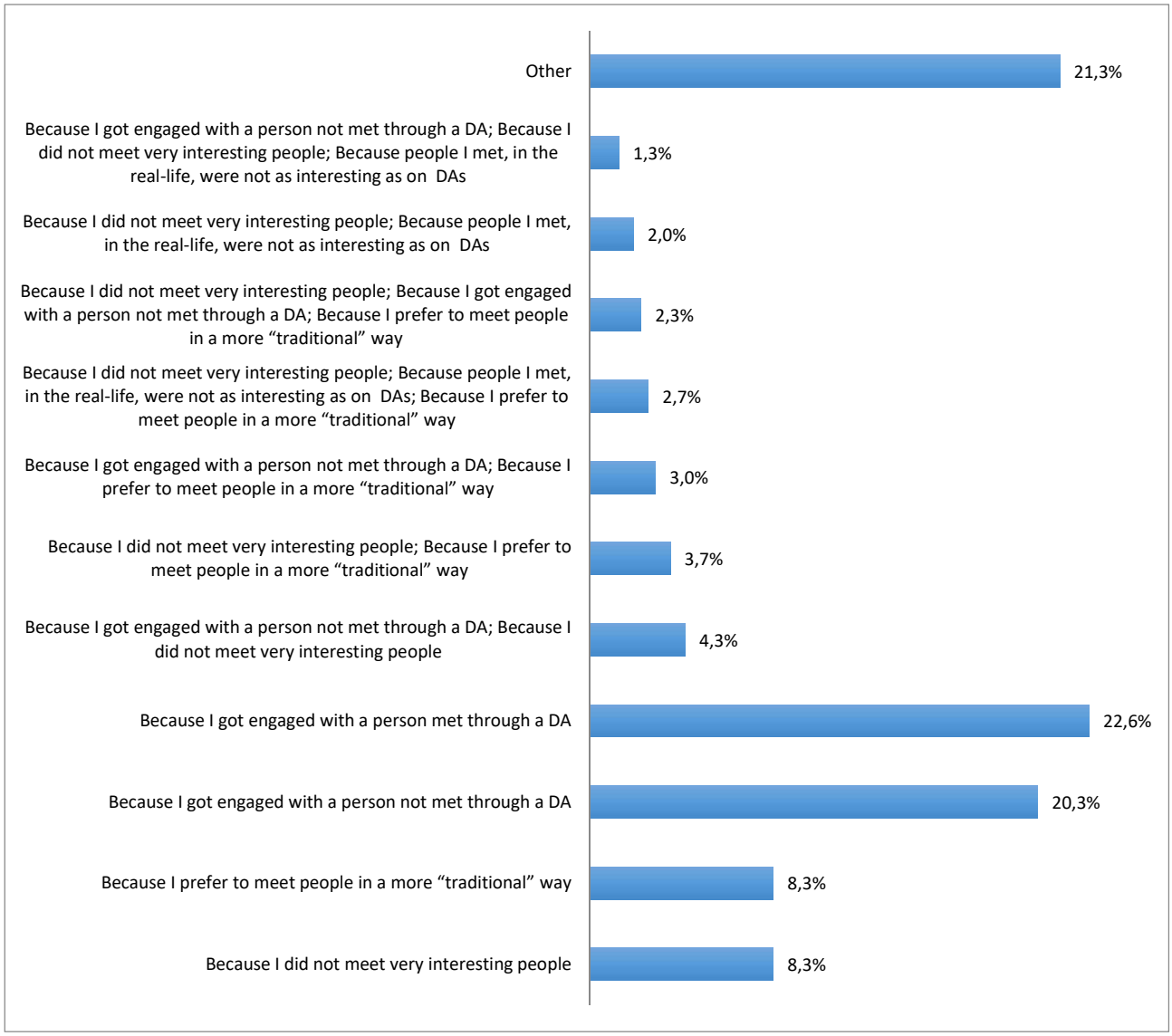


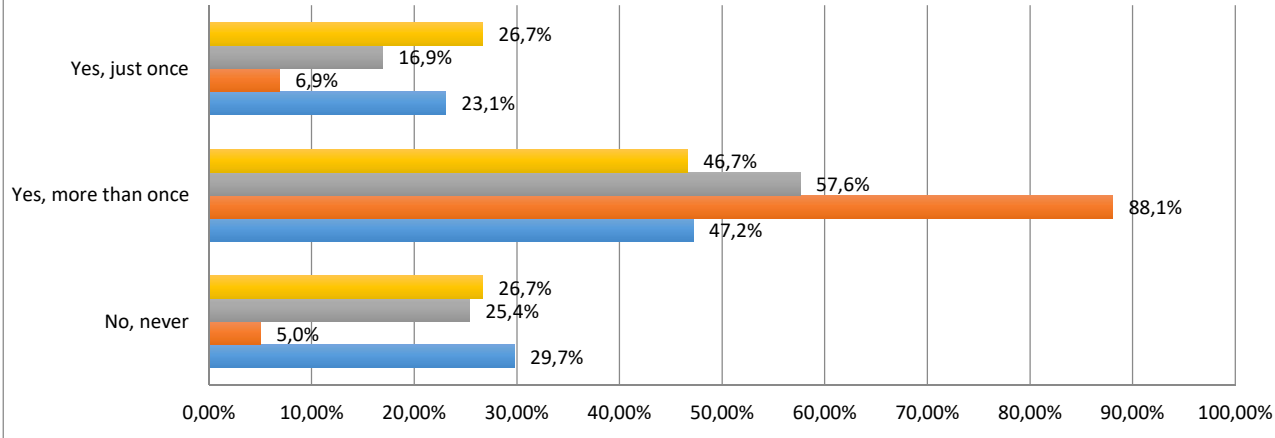
Fig. 26 - Respondents distribution by use of dating apps and gender identity

The question "Why did you stop using dating apps?" left the possibility of answering with more than one response modality, but the great part of the respondents chose just one option: 22.6% of respondents uninstalled dating apps because they got engaged to a person met through the same app, 20.3% because they got engaged to a person without the help of any apps, 8.3% rediscovered the traditional way of meeting people, and the other 8.3% just met non-interesting people and so they quit. These reasons are the same independently of gender identity and biological sex, but no charts will be reported for reasons of space and readability of this paragraph.

What emerges from these results is the strong discontinuity in the use of dating apps. Many respondents delete apps when they establish a monogamous

relationship (and probably reinstall them when, and if, they break up), but also as a self-care or self-regulation strategy: reasons for deleting dating apps included a desire to focus on themselves (for instance, a student user who wants to focus on studying during university exam sessions), but also due to the disillusion given by the dates they had. Some women spoke (in the open modality response) about deleting apps due to feelings of frustration and disappointment towards the low number of matches or messages they received. On the other side, men uninstall dating apps less, and not for lack of other users, but for safety and privacy issue: in fact, dating app use, is sometimes, interconnected with the use of other social media platforms, for instance, Tinder links to Instagram and Facebook. Participants deliberately connected (and disconnected) their dating app with their preferred social media platforms to manage a sense of safety and visibility in encounters with friends and strangers (Albury et al., 2019).

That said, we asked LGBTQ people if they have ever hung out with someone met on a dating app. In the bar chart and table below, the respondents' distribution, by gender identity:



	No, never	Yes, more than once	Yes, just once
Transexual	26,67%	46,67%	26,67%
Queer	25,42%	57,63%	16,95%
Male	5,05%	88,07%	6,88%
Female	29,74%	47,18%	23,08%

Fig. 27 – Distribution of respondents who have (or have not) hung out with someone met on a dating app, by gender identity

The bar chart above points out that most of the dating-app users have met up with people they met online in person, but with strong differences by gender: 88.1% of male respondents have met other dating app users personally more than once; this percentage decrease to under 60% for queer people (57.6%), and under 50% for female and transsexual people (respectively 46.7% and 47.2%). Most of the people who met, just once, other online people in person are transsexuals (26.7%) female (23.1%), and this data could be a wake-up call because, usually, who only met up with someone once, probably had a bad experience. Analogously, only 5% of male users have never hung out with someone met on it, against the 25.4% of females, 26.7% of transsexuals and 25.4% of queer people.

These results could be easily interpreted as individuals' different perception of safety: dating app users feel more or less safe, at the moment they personally meet someone met online, depending on their own gender identity. This interpretation can be confirmed by the Australian Research Council, who published the final report of the Linkage Project "Safety, Risk and Wellbeing on Digital Dating Apps", in which they state that most women want to have the first date in public, "both as a safety precaution and as a means to avoid the potential 'awkwardness' of changing their mind about hooking up" (Albury et al., 2019).

In the following bar chart gender identity and sexual orientation have been crossed regarding their experience of hanging out with people met online:

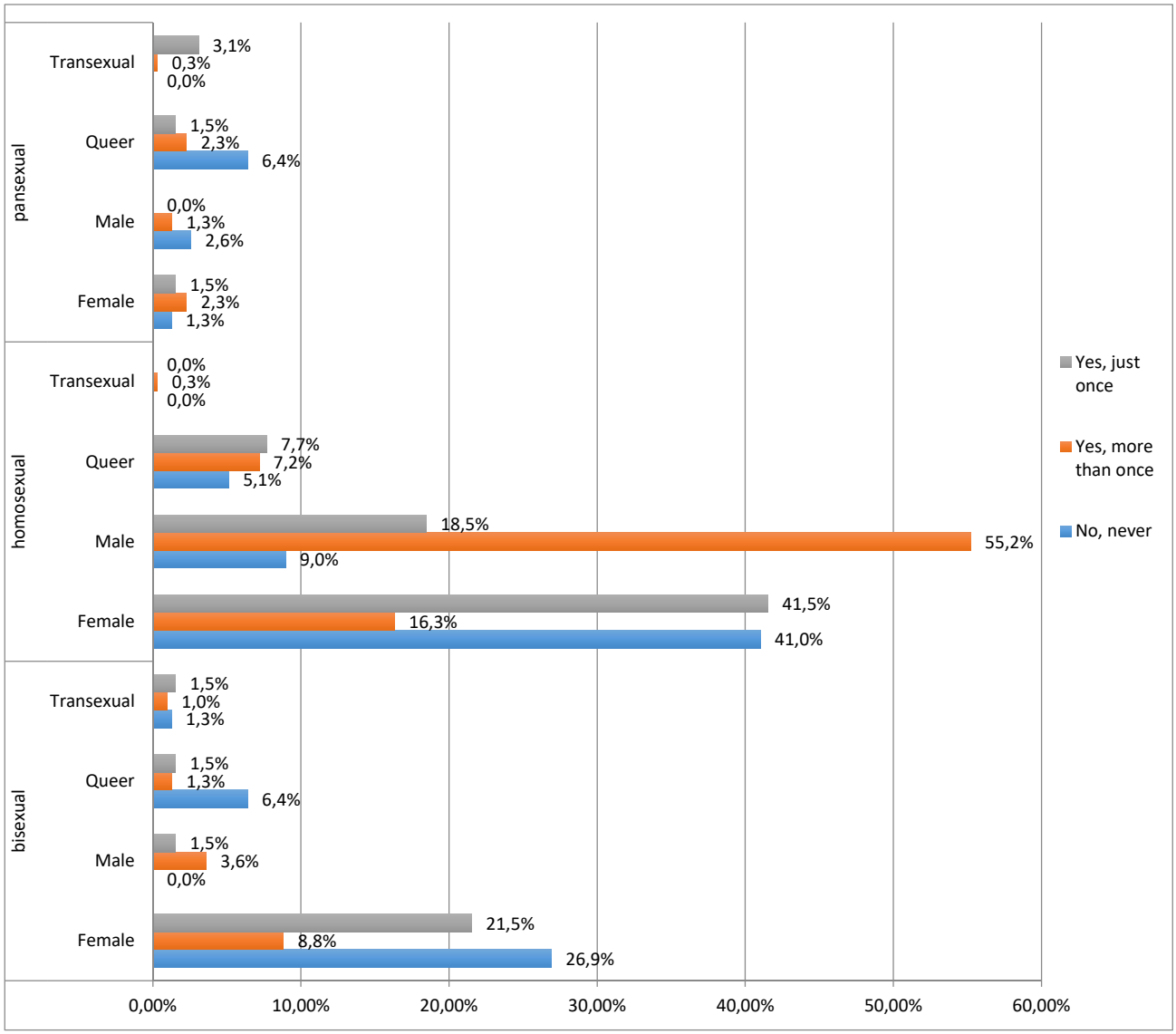


Fig. 28 – Distribution of respondents by gender identity and sexual orientation, in having (or not having) hung out with people met on a dating app

Individuals' safety perception does not depend on gender only, but likely on sexual orientation too. For instance, the majority (71.1%) of those who have hung out with someone met online more than once is homosexual; 55.2% of them are, in turn, gay men. On the contrary, 41% of lesbian women have never gone out with a woman met online, 41.5% have just once.

Dating apps, which aim to be an innovative tool of interaction and socialization, often re-proposes traditional heteronormativity schemes, in which females, queers and transsexual people are more afraid than gay and bisexual men of going out.

Concluding, we can say that dating apps reach their purpose, and represent a resource for the LGBTQ population(s). In the beginning, dating apps facilitate the research of new people to meet. However, after a first phase, they introduce, in a very different way in each subpopulation of the community, problems of self-construction in a computer-mediated environment and embodiment in the digital age (Mowlabocus, 2010). A direct consequence of these problems is that some gender identities, and some sexual orientations, may feel more uncomfortable than others in these digital spaces, and they could also become more exposed to discrimination, harassment, and other negative behaviours.

10 Being an LGBTQ person and being part of the LGBTQ community

Since the introduction we have said that referring to LGBTQ as a population, thus overlooking the heterogeneity of the groups identified by this acronym, can be strongly inappropriate. For this reason, in previous chapters, we have referred to LGBTQ adding an "s", within brackets, following the world population. We consider the LGBTQ population(s) our target one, but always considering the inner many-sidedness of what we usually call community.

In the fourth part of the questionnaire, we asked the people involved with the survey, to talk about their relationship with the community. We investigated the *pride*, understanding both self-respect for being a member of a socially marginalized group, and the Pride, namely the parade which aims to celebrate LGBTQ social and self-acceptance, commemorating every June the 1969 Stonewall riots, the first worldwide milestone for the LGBTQ rights movement.

Moreover, the acceptance of a set of sentences, regarding the relationship with LGBTQ associations, has been tested through Likert-scale questions.

Every questionnaire item will be deeply analysed through the usual magnifying lens of intersectionality, the chosen perspective to explore how interactions between social characteristics, such as gender, sexual orientation, age classes and geographical region, shape an individual's, or group's, life experience.

10.1 New and old socialization exclusive spots

When Barbagli and Colombo wrote, in 2001, Omosessuali moderni (the mixed-methods research book, deeply analysed in paragraph 4.1), they dedicated a whole chapter to "sex places". Two authors noted that during the 80s and 90s, homosexual meeting places were radically changed, passing from being spontaneous and mixed places, to organized locations which attracted just homosexual users: "battuage" spots, initially for gay people, such as pornographic cinemas and saunas, left the field to cruising bars, often located in

entertainment districts, exclusive places with an almost completely gay clientele, and then to local circles and, especially, associations. The research of casual sex gave way to the creation of collective identity: the gay identity, which, then, will concentrically enlarge to lesbian first, then bisexual people, transgenders and transsexuals, followed by all the other subpopulations, such as asexual, intersex and queer people. Mixed places gave way to exclusive ones, based on the LGBTQ identity.

But now, almost twenty years later, socialization has most likely further changed, and, needless to say, the blame (or the credit) goes to digital technology. One of the semi-structured interviews carried out during the survey design phase, to a member of the roman LGBTQ association board named after "Mario Mieli", alerted us to a negative correlation between digital technologies and community socialization. In fact, it was said that social media and dating apps brought decreasing participation in all those activities, mainly sponsored by LGBTQ associations, whose aim was to promote socialization between people sharing the same sexual orientation, such as cineforum, disco events or soirées in gay-bars. A strong hypothesis concerning a side-effect of new technologies introduction on associations, confirmed also by a Swedish study (Ross, Tikkanen, and Berg, 2014) which explicitly said «those who use the Internet extensively are less likely to be involved in other aspects of the community».

We tried to investigate this issue in the survey, asking respondents if they frequent these kinds of places to meet people:

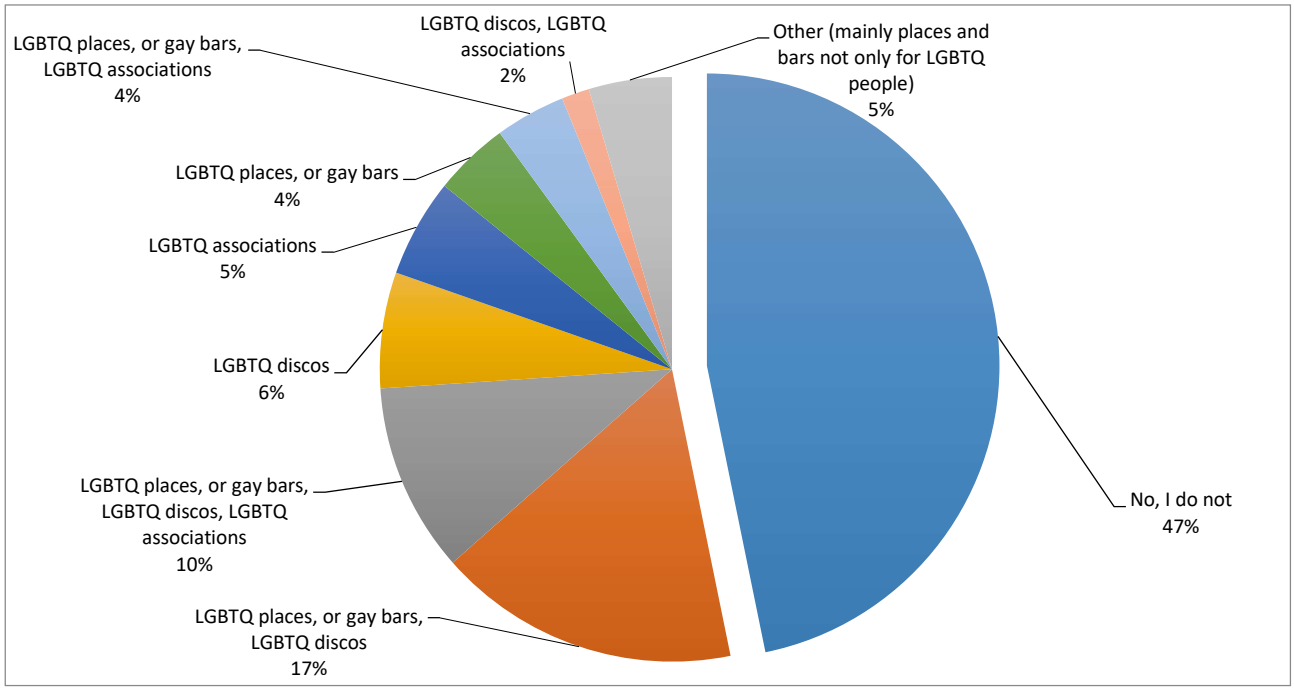


Fig. 29 – Distribution of places frequented by respondents to meet people and hang out

21.2% of respondents frequent LGBTQ association to socialize, often jointly with LGBTQ places, or gay bars and clubs (10%). On the contrary, 78.7% of respondents do not go to any associations, 47% of which frequent none of the above places.

Analysing the same percentages by gender identity we can note that 17.6% female respondents frequent associations, against 23.1% of males. Higher percentages can be observed for queer (33%) and transsexual people (36%). Contextually, 55% of females do not go to any of the above places, against 36.4% of males, 44.33% of queer, and 44% transsexuals. It would therefore appear that female respondents are less likely to gravitate towards LGBTQ associations than male, queer and transsexual people.

The differentiation of the results divided by age class, in regards to frequenting LGBTQ associations is more significant. Only 12.4% of responding minors frequent associations; the percentage slowly increase alongside the respondents' age: 19.9% in the 18-23 age group, 27,6% in the ages 24 and 29, 25% of 29 to 34 year olds, 24.3% of 35 to 39 year olds, 20% of 40 to 44 year olds, 23.1% of 45 to 49 year olds, and finally, 44.4% in the 50s age group. LGBTQ minors, particularly, seem to be uncatchable, if not through the internet: 78.4% of respondents under 18 do not go to any proposed LGBTQ socializations spots, actually twice the percentage observed in the other age groups (40.4%).

Conversely, significant differences based on the geographical area or the residence city size, are not observed, unexpectedly, and percentages among the regions and city size groups are similar and close to the national average.

These percentages lead us to make some essential methodological considerations:

- most of the reviewed survey (PART 1) massively leans on LGBTQ associations, but noting how many respondents never frequent them (on average 78.7%, peaking at 87.6% in the youngest age class) confirms the "misleadingness" of continuing to carry out surveys exclusively through associations.
- LGBTQ-friendly bars and clubs, for example, are more frequented than associations (25% against 21.2%), even if often affiliated to them.
- However, the most significant data not to be overlooked is the 46.8% of respondents who do not frequent any meeting places. Thus meaning they would be completely invisible to any traditional survey. Of course, this data is strongly influenced by respondents' age and gender distribution (deeply analysed in chapter 7), but, isolating correlation effects introduced by these socio-demographic variables, points out that LGBTQ bars and clubs are more appreciated by gay men than lesbian women, and by adults than youngsters.
- Moreover, it results that transsexual (36%) and queer people (33%) are more likely to be part of an LGBTQ association. This figure helps us understand the importance of these associations nowadays: as we will see in paragraph 10.3, they represent a supporting authority (psychological, social-sanitary, and social-assistance) that transsexuals probably need most; but also a strong political institution, which struggles for emancipation and equality values.

10.2 More than a "ready-made" community

The acronym LGBTQ, is often followed by the word "community", in newspaper articles, in politicians' speeches, and also by LGBTQ people themselves. This term, LGBTQ community, is right to want to create an inner cohesiveness but, at the same time, it risks not catching the differences and complexities of dissimilar understandings and experiences. For instance, it could

wrongly suggest some form of sharing, which for some people can be annoying, because it would ignore some experiences of inequality, or discrimination, within the LGBTQ community.

In 2017, Eleanor Formby tried to explore "LGBT Spaces and Communities, Contrasting Identities, Belongings and Wellbeing", through a study which involved over 600 LGBT participants all over the UK, deeply described by a book of the same name. This research pointed out, through an interdisciplinary approach, several discriminating experiences lived by participants from other LGBT people relating to their age, body, disability, ethnicity, faith, HIV status, or perceived social class. So, what has always been, superficially, considered as a ready-made and harmonious community, might actually not be: belonging to a community is not a fact just because many different people share a gender identity and/or a sexual orientation. Using the term "LGBTQ community" indiscriminately, as a saying expression, could alienate some people and even risks deterring LGBTQ (and other, such as intersexual and asexual) people from joining some important civil rights claims.

In our survey we tried to investigate the same issue through a quantitative point of view, directly asking respondents: "do you feel part of the LGBTQ community?". Results are shown in the pie chart below:

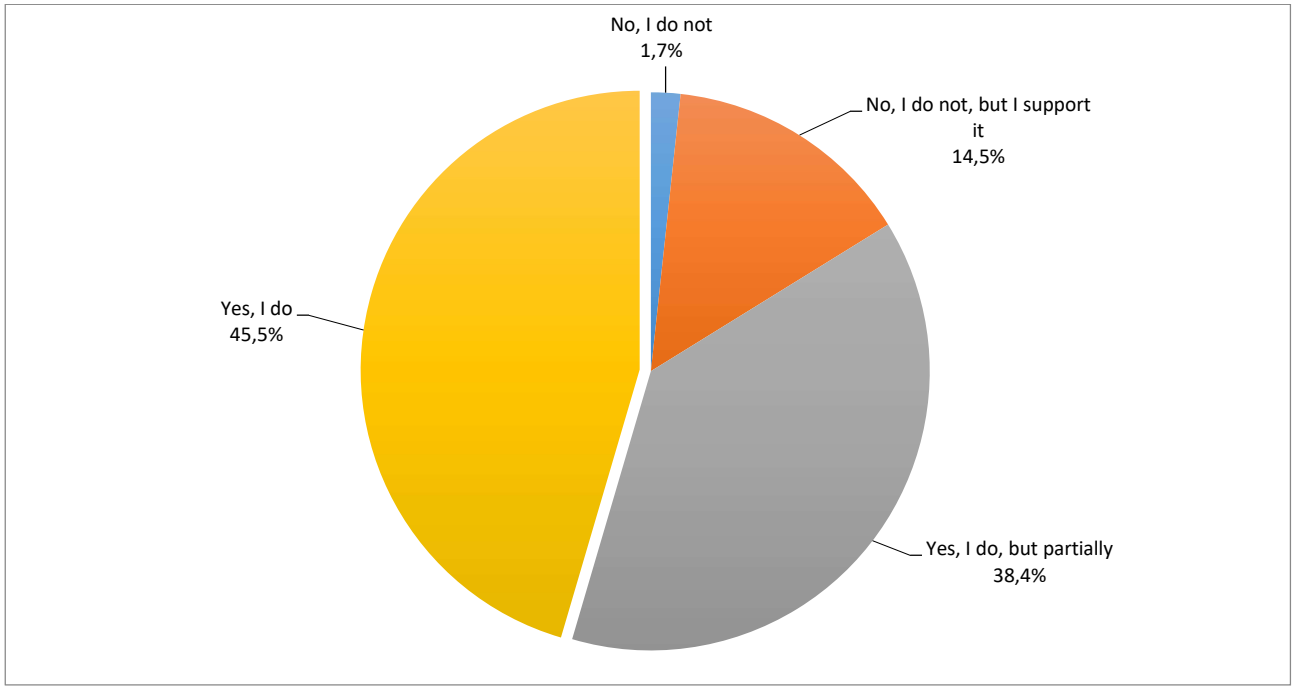


Fig. 30 – Distribution of respondents by the perceived degree of belonging to the LGBTQ community

45.5% of respondents unconditionally feel part of the LGBTQ community. 38.4% have some reticence, and feel themselves "partial belonging" to the community. 14.5% of respondents do not feel part of the LGBTQ community, despite their gender identity and sexual orientation, but they support it. Lastly, 1.7% do not believe that they belong to the community at all.

Results seem to argue Formby's thesis of a conditioned belonging to a so multifaceted community: people support it, feel part of it, but often not unreservedly. On the other hand, the number of people who feel like they do not belong to the community at all is very few (1.7%).

These percentages could be very different if analysed regarding people's characteristics. In the chart below they are presented according to respondents' sexual orientation:

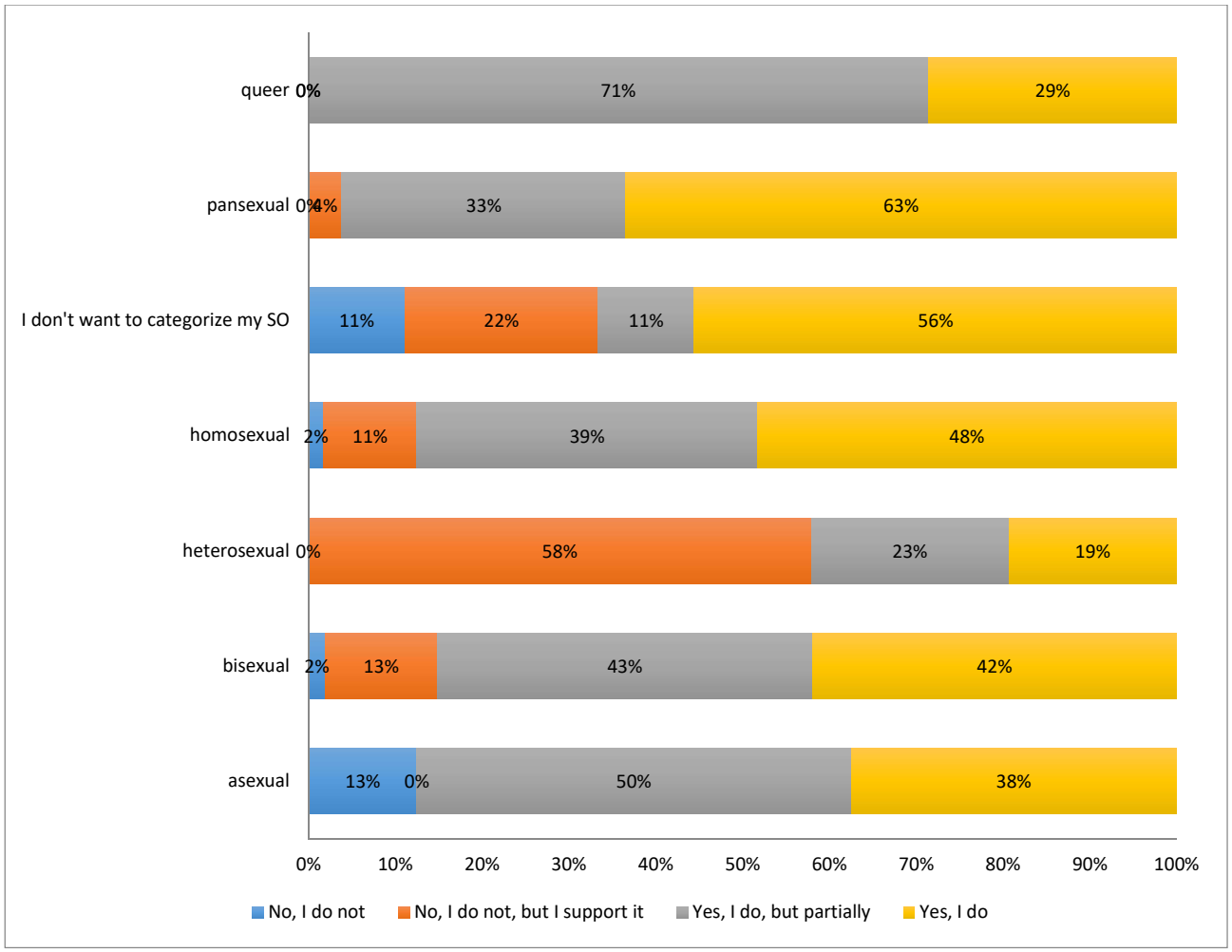


Fig. 31 – Distribution of respondents by the perceived degree of belonging to the LGBTQ community, by respondents' sexual orientation

The degree of belonging to the LGBTQ community is significantly variable depending on sexual orientation: pansexuals (64%), homosexuals (48%), and people who do not want to categorize their own sexual orientation (55%) are most inclined to give their adhesion to the community. Conversely queer (71%), bisexual (43%) and asexual people (50%) more likely to feel to belong to the community, but partially and with some constraints.

In this, and the next, analysis we also considered heterosexual cisgender respondents, to test the sense of belonging to a community to which, by definition, they do not belong: 58% reached self-defined heterosexual people answer to not feel part of the community, but to support it; likely they are cisgender respondents included in this analysis. Other heterosexuals (42%) feel instead to belong to the community, in some way. None says not to be a part of it, confirming their interesting point of view on the survey issues.

The same analysis has been carried out by gender identity too, but results were not as significant as by sexual orientation.

The recurring conditional adhesion of surveyed people to the LGBTQ community, should make all of us think about the wording we usually use. A community is definitely more than a group of people having a particular characteristic, such as gender identity or sexual orientation, in common. In sociology, the concept of community introduced in the late 19th century by Ferdinand Tönnies (as *Gemeinschaft*, counterposed to the concept of society, *Gesellschaft*) requires precise constraints:

- membership,
- influence (of the people, both in the group, and by the group),
- participation for the fulfilment of need,
- community's members share common values, history, beliefs, or behaviours

(McMillan and Chavis, 1986).

It is easy to note that these community definition requirements are quite strict, and probably not so simply suitable to the LGBTQ community case, because of its noteworthy inner diversity, and because of the different degrees of integration and emotional connection with each other. Consider for instance the history of the fight for civil rights of the individual groups which compose LGBTQ: it is easy to note that gay men have a path of struggle different from that

of lesbian women, or, in the same way, transsexual people have a longer experience of discrimination than asexual people.

This is not to say that all of us should avoid the phrase "LGBTQ community" altogether, but often using "LGBTQ people" would be more accurate, and would not risk alienation felt by an already (at times) marginalised group (Formby, 2017).

So far, we have highlighted the differences within what we usually call the LGBTQ community, to emphasise the fairness of talking about individuals' experiences rather than macro analytical groups, which risk debasing the discussion. However, there are several things that (almost) every LGBTQ person agrees with. The most important of these things is the Pride, namely the parade held in June in many cities all over the world to celebrate, precisely, the pride and the history of the movement. Pride is common to almost all, and participation in the event is transversal and generalized, across the boundaries of gender identity and sexual orientation which individuate the LGBTQ population(s), as shown in the graph below:

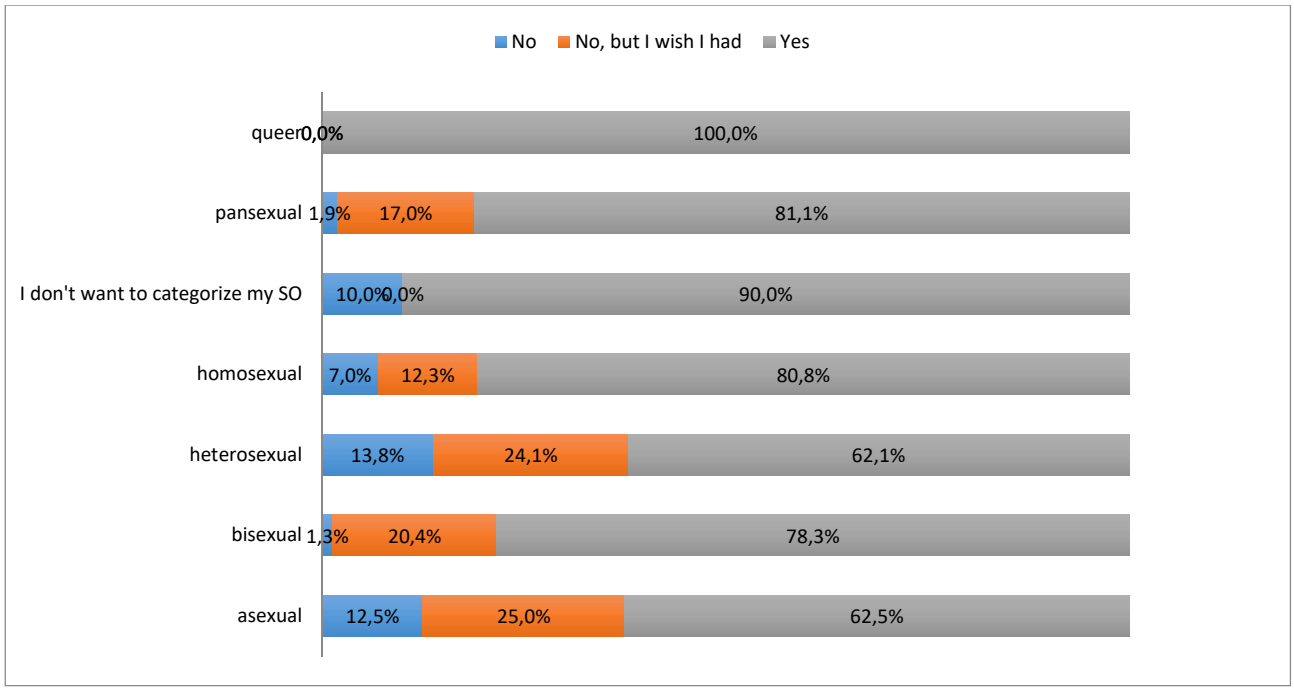


Fig. 32 – Distribution of respondents by participation in a Pride march, and by respondents' sexual orientation

Averagely, 78.8% of respondents have been to a Pride march; a high and transversal percentage, since almost every sexual orientation reports a value higher than the average: queer (100%), those who do not want to categorize herself/himself (81.13%), pansexuals (81.13%), homosexuals (80.77%), bisexuals (78.34%). 62.07% of heterosexual respondents have been to a Pride, and this data confirms that surveyed people largely support the community, even when they do not properly belong to it. The same percentage for asexual people, who still have to find their dimension into the community, just because of their non-sexual orientation, and because they could be considered "latecomers". Anyway, 25% of asexuals have never been to a Pride march, but they wish they had, and this could be seen as a good sign of their consideration amongst the community. Also among heterosexual people, the percentage who would want to attend is significant (24.14%) and confirms how this parade is getting increasingly massive and open to all those people who want to support the LGBTQ instances, independently of their identity.

The wide participation of the Pride parades could entail some methodological "good intentions" for further investigations: both qualitative, such as ethnography and participant observations, and quantitative methodologies, through a survey which use snowballing methods, it could employ the massive participation to this kind of event to recruit people to interview, by diversifying and enlarging the sample.

10.3 The LGBTQ organizations key role in the virtual collective identity construction

In previous chapters, we mentioned the change of the LGBTQ daily life due to the spread of digital technology. We have seen how social media, nowadays, play an important role in uniting LGBTQ community members and allowing them to organize community advocacy to further their impact. Furthermore, by providing a platform for people all around the world to connect onto, social media is fulfilling much of its potential as a tool for advocacy and information sharing. This helps people from around the world stand together and demand justice when an unethical act occurs to a specific individual (Wallace, 2019).

This epochal change may seem like a drop in power, prestige, or appeal of LGBTQ associations and organizations. So in this survey, we asked participants

to express their agreement (or disagreement) about certain sentences on this topic, through a Likert scale item.

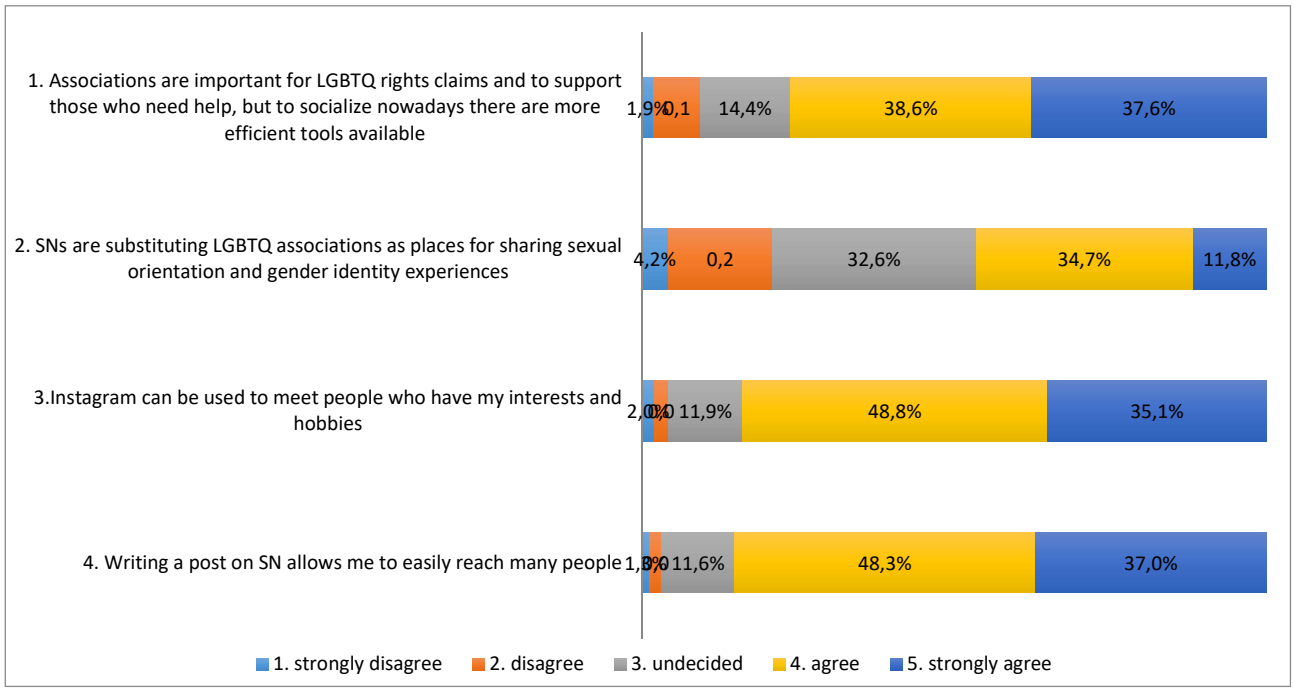


Fig. 33 – Respondents distribution, by agreement scale, about LGBTQ associations and organizations statements

The graph shows, again, how social media, specifically Instagram can be used to meet kindred people (3. 35.1% strongly agree + 48.8% agree) and to easily reach many people with a post (4. 37% strongly agree + 48.3% agree). However, when asking if social media are substituting LGBTQ associations, as places for sharing sexual orientation and gender identity experiences, many people remain undecided (2. 32.6%), or even disagree (16.7% + 4.2% strongly disagree). Greater appreciation is received by the sentence "Associations are important for LGBTQ rights claims [...] but to socialize nowadays there are more efficient tools available" (1. 37.6% strongly agree + 38.6% agree).

Survey data confirm somehow one of the most important themes which emerged during interviews: LGBTQ associations and organizations are still essential as political pressure institutions on the front of civil rights recognition and to fight homophobia, biphobia and transphobia (Barbagli and Colombo, 2007). Associations, organizations, and communities are not dying but just in transition, as it has happened to all other institutions involved in political activism, such as political parties or trade unions (Simon Rosser et al., 2008).

Respondents seem to agree that the individualistic drift still does not work in a political struggle context, where each individual political subject, becomes "us" in public (Butler, 2015). Some of the most profound social changes of the last century have been promoted through a combination of research, public education, advocacy, legislation, and litigation fostered by non-profit organizations. Moreover, after the Brazilian Resolution[22] and the Yogyakarta Principles[23], the global LGBT field has become more strategic, both for public institutions and private enterprises. This sector is slowly beginning to professionalize, creating new opportunities and challenges for the people and organizations working within it. The global LGBT movement's success in the next several years will largely depend on its members' abilities to grow their own organizations, collaborate with partners, and develop strategies to overcome the formidable obstacles and opponents in their paths, inevitably bound to technological progress.

Anyway, digital technologies do not entail only negative side effects, and, online social media are not only used to organize encounters, but also to correspond and sometimes to endorse positions, by sharing it: for instance, 83% of respondent people have used a form of social media to endorse or support the LGBTQ community, and this percentage remains high across the several groups of people composing the LGBTQ population(s). So, social media not only help the LGBT people to feel a sense of "normality"; it also helps to normalize this community to outsiders who may not have much exposure to LGBTQ people in their day-to-day lives. Some journalists and researchers, calls this one the "post-gay era" (Collard, 1998), in which activists are not motivated anymore to draw boundaries against members of the dominant group, but to build bridges toward them: no longer a collective identity construction using an oppositional "*us versus them*", to underline all the differences with heterosexual people, but an inclusive "*us and them*" to emphasize the similarities with them (Ghaziani, 2011). Platforms

[22] The Brazilian Resolution is a document concerning human rights in the area of sexual orientation and gender identity presented to the Economic and Social Council of the United Nations in 2003. It affirmed the equal and inalienable rights of all people, and promotes respect for diversity in societies.

[23] The Yogyakarta Principles is a document concerning human rights, specifically in the area of sexual orientation and gender identity, published as a final report of the International Meeting of human rights groups, held in Yogyakarta (Indonesia) in 2006.

like Facebook, Instagram, and Twitter help unite communities and encourage groups to take action, by helping them feel empowered and strong together, also with heterosexual supporters. Associations can thus take advantages of this new empowerment, digitally conveyed.

Summing up, most of the surveyed people exploit the potential of a new socialization process through social media; in the same way, they answered to have been to a Pride march. The square, which physically fills up once a year, becomes a virtual square, daily full of people who want to meet up, do networking, hang banners (in the form of posts and tweets), and voice their endorsement. However, when we asked the people surveyed "do you think social media are substituting LGBTQ organization?" the great part of the answers were negative. Associations remain, and it's likely to remain this way forever, the official community institution of political and legislative pressing. The fact that associations no longer have the leadership of socialization does not need to cause alarm: before them, it happened to political parties, and then to trade unions. LGBTQ organizations have to make, and are making, a great job in updating as to catch the collective identity which people build and perform daily, in both real and virtual environments.

11 Beyond sexual discrimination: old and new forms of homophobia, biphobia, and transphobia.

11.1 Hate crimes against LGBTQ people: still an emergency in Italy

Most of the surveys studied for projecting the OTR research, and reviewed in the first part share the main goal of the study, from a quantitative point of view, the discrimination for sexual orientation and gender identity. Many of these surveys aim to be usable, and useful, to policymakers who are responsible for planning efficient diversity inclusion policies.

In Italy, discrimination based on sexual orientation and gender identity is still an emergency, despite the positive steps taken, for instance, the civil unions for same-sex couples. Italy does not have adequate laws to oppose hate crimes against LGBTQ people, therefore, in 2014, a document titled "*Italy: the state of human rights of LGBTQI people*" was submitted to the United Nations by a coalition of Italian organizations. In the document, all the wrongs that the Italian state commits against LGBTQs were listed, such as incomplete legislation on this topic, incitement to hatred, also by politicians and institutional figures, no recognition for sons and daughters of homosexual people, no regulation for homophobia and transphobia, no refugee status for LGBTQ migrants coming from countries in which homosexuality is still considered a crime. Studying discrimination among the LGBTQ population(s) means understanding its causes, contexts in which it happens, and by who, paying the utmost attention to single discriminating events, but, at the same time, by grouping them into clusters theoretically interpretable, and numerically quantifiable.

Obviously, quantitative methods suffer some limitations because asking such sensitive questions in a survey, entails a risk of information loss, and a low accuracy of collected data. To avoid these quantitative weaknesses, we asked participants if they had suffered the several forms of discrimination, emerged during semi-structured interviews. These five discrimination contexts will coincide with this chapter's paragraphs, and they are:

1. household,
2. work,
3. school and university,
4. social media,
5. sport.

We take all possible care to distinguish old and new discrimination forms, trying to understand if non-heterosexuality, and non-cisgenderism, are still considered a dishonour in the household, a flaw at work, a shame at school, a source of troubles online, and an embarrassment in the locker room of sporting centres.

11.2 Within the household: when the family heteronormativity breaks down

In paragraph 8.4 we have seen 98.6% of the survey's respondents came out, both in real life and on social media, but only 21.7% of them came out to a family member. Unveiling themselves, within the household, still represents a hard step to make for many homosexual people: consequences of coming-out can be unpleasant, or even oppressive. Homophobia, as well as biphobia and transphobia, is not just an unknown person who offends and assaults an LGBTQ person in the streets; it could also have the face of a person living under the same roof.

For this reason, we asked those surveyed if they have ever suffered discrimination events in their household, and, if so, of what kind: 54.9% of respondents experienced some sort of negative behaviours towards them within the household. Differences among different sorts of discriminations and by gender identity are significant:

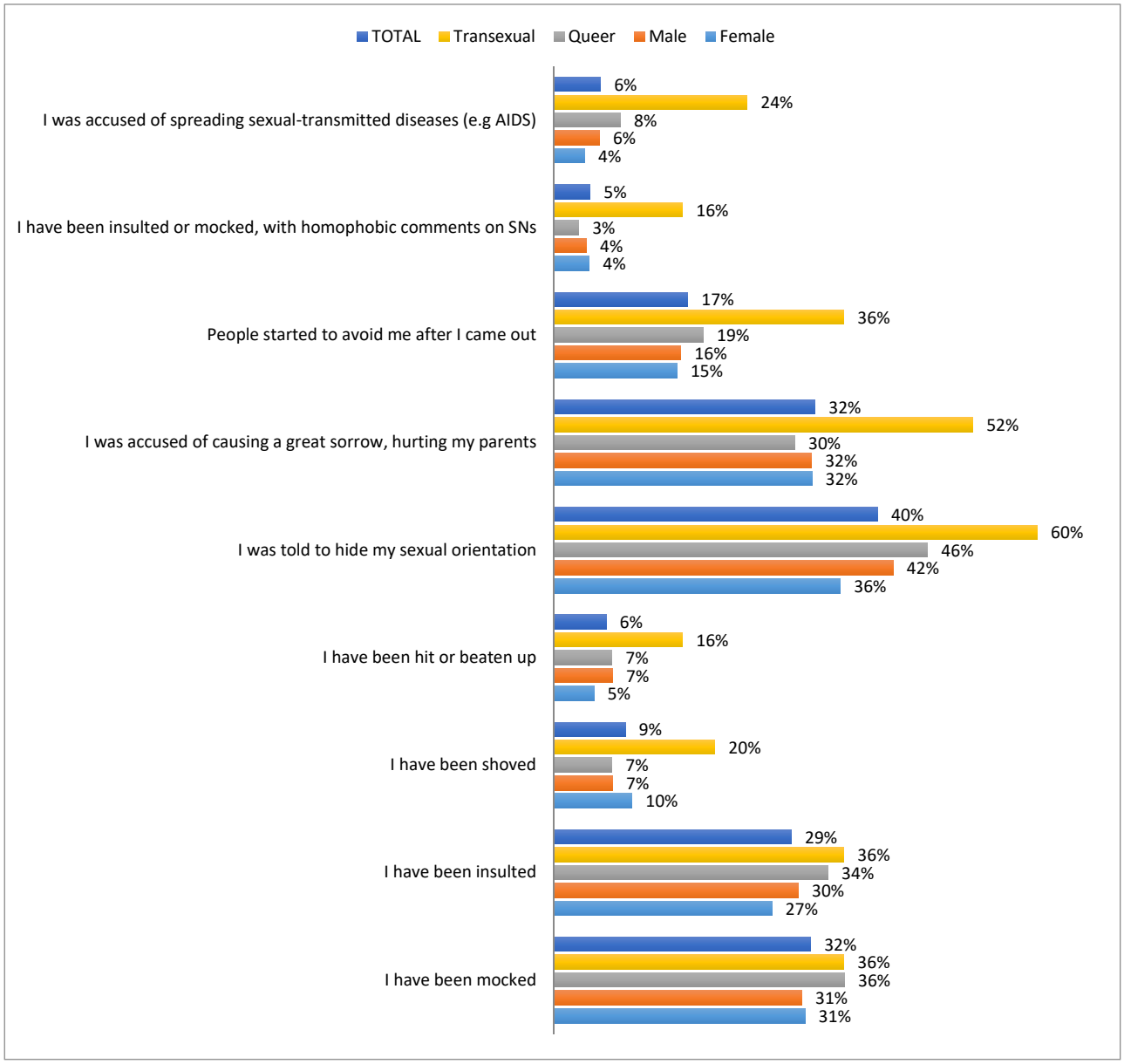


Fig. 34 – Respondents distribution, by discrimination reported in their household, and by gender identity

Discriminating behaviour in households can be the most variable, often less violent and dramatic than we thought, but more silent and sneaky: 40.2% of respondents affirm they were told not to show themselves, 32.4% were told that they upset their parents, 16.6% have been avoided but their relatives after the coming-out. Nevertheless, the amount of discriminating events through more direct forms of discrimination, both physical (6.5% of respondents got beaten up, 8.9% have been shoved) and verbal (29.5% have been insulted, 31.9% have been mocked, 5.8% have been accused of spreading diseases) are sadly remarkable.

Transsexual people are far more than likely than others to report every kind of discrimination in their household: 60% of transsexuals are told to hide themselves, 52% to have caused a feeling of great sorrow to their parents, and 24% have been treated as a plague spreader by a relative. People most discriminated in the family, after transsexuals, are queers. We found little percentage difference between the male and female gender.

Social media result as a safe place, at least from family: only 4.5% of respondents have been publicly discriminated on the web by a relative. Unfortunately, dirty linens continue to be washed in private spaces. Anyway, saying "discrimination within the household" can be really generic. That's the reason why another question of the survey was "which one of your family members discriminates you, and in which way?"

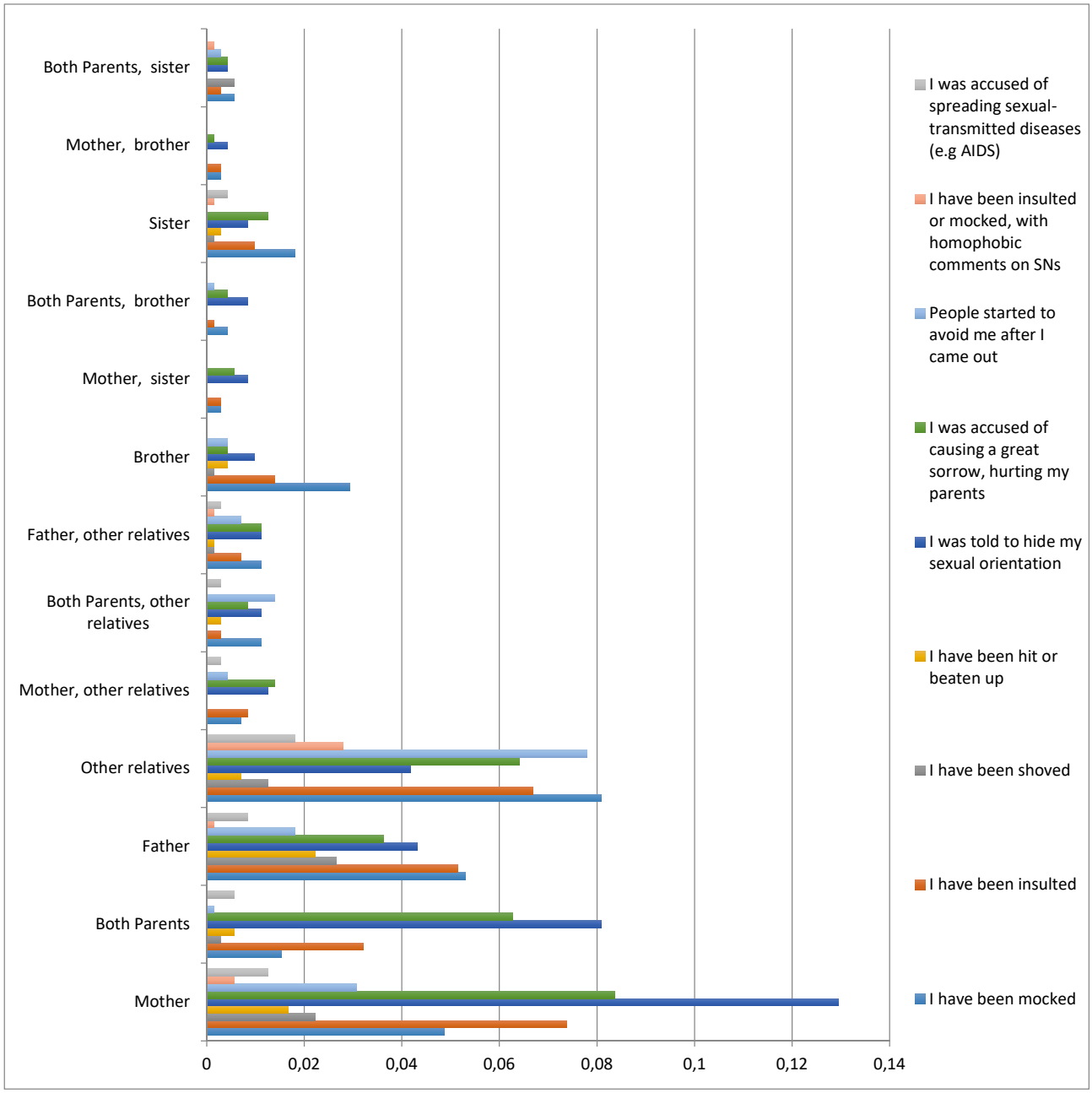


Fig. 35 – Respondents distribution, by discrimination reported in the household, and by which family members

In paragraph 8.4 we stated that women are the preferred interlocutors for first coming-out to, likely for a more sensitive attitude or, maybe, for a lesser vehemence of reactions than men. However, in the light of last graphics we have to say that mothers are those who most likely discriminate their sons and daughters, although in a less violent way: 13% of respondents' mothers suggested their children should hide their sexual identity, 8.3% accused them of causing a great regret in the family, 7.3% has even insulted their LGBTQ sons and

daughters. Fathers have lower percentages but they have harder discriminative behaviours, such as insults (5.3%) and mockery (5.2%). Derision is also the most frequent type of discrimination by other relatives, such as uncles, aunts, cousins, grandparents, etc. (8%), they also report a high percentage for ghosting (7.8%), namely the situation in which relatives disappear from the life of a person who comes out (Spitale, 2015).

These behaviours are interpreted in literature (Graglia, 2012; Field, 1995) as the break of the expectations of heteronormativity that parents and relatives have towards their sons, daughters, brothers, sisters, nephews, etc. Homosexuality revelation shatters these families' dreams more than anything else. All of us are affected by the overwhelming message that families are "good", "right" and "normal": housing, welfare, advertising... everything is based on the assumption that most of the people live, or aspire to live, within the edifying structure of a family. Homosexuality, for instance, is an unforeseen circumstance, because it betrays the normality imposed by society. Emotional reactions to this unveiling are very intense, such as guilt, unbelief, negation, sense of betrayal, concerns regarding health and social inclusion, anger, sadness, disgust, shame, feeling of not really knowing sons and daughters.

Anyway, there would be no discrimination towards LGBTQs if people stopped to conceive, as a traditionalist and conservative axiom, that there is only one way of being a family. Parents and other relatives, certainly need time for elaborating, cognitively and emotionally, sons' and daughters' coming-out (Graglia, 2012), but, first of all, they need to understand there is not a unique manner of performing sexuality, gender, and then family idea.

In the next paragraph, we will focus on fields in which action by policymakers could be more easily manageable and fruitful. In-depth, we will examine the work, scholastic, virtual and sporting fields. However, it is important to specify that these are not the most dangerous and unsafe fields. Most violent, direct and unpredictable, discriminating behaviours remain, frequently, prerogative of complete strangers: 2.4% of respondents have been hit or beaten up by an unknown person, often in the streets, 4.9% have been shoved, 19.8% insulted and 17.1% mocked, again all by unknown people.

Data confirms that homotransphobia is a well-rooted behaviour, which winds between many people in our county. Percentages collected are remarkable and unveil the size of the most difficult homophobic phenomena, which remains unpunished according to Italian jurisdiction. For this reason, both the United

Nations and the European Parliament solicited our country, to legislate on anti-homotransphobic discrimination, because, from a civil rights point of view, we are behind other Western-Europe countries, such as Denmark, France, Iceland, Norway, Holland and Sweden, where homophobia, is understood as a violent act or hate speech is a full-fledged crime.
Italian LGBTQ activists have been asking for this law for more than twenty years, but unsuccessfully, because our legal system is really meticulous on the interpretation of laws.

The latest proposal would add to discrimination for racial and religious reasons (art. 604), those for the sexual identity of discriminated people. The differences on which it struggles, is that if against Shoah, is a crime also if only propaganda, against LGBTQ people, the only instigation would be a crime, and not the propaganda. This discussion, long-lasting and sterile, between LGBTQ organizations and catholic conservatories, led to a stalemate, which has lasted for years, and disadvantaged victims. A greater openness, in Italy, has been the adhesion to the International Day Against Homophobia, Transphobia and Biphobia (17th May), a day to raise awareness of LGBTQ rights all over the world, and celebrated mainly in schools, universities and at work, places in which social policies seem to be more feasible.

11.3 At work: mobbing LGBTQ people for undermining their careers

During the questionnaire design phase, amongst the people interviewed, the interest in discrimination in the workplace emerged. Obviously, discrimination at work can come in a lot of different forms, such as mobbing, sexual harassment[24], pay-gap, etc., from the most concealed, to the most evident and violent ones. Mobbing, however, has the peculiarity, of undermining the career possibility of an LGBTQ person differently from how it would undermine one of a heterosexual and/or a cisgender person. Mobbing is defined as the situation in which a person or group of people in a work environment exercise extreme,

[24] Sexual harassment includes a series of attacks ranging from suggestive comments, jokes or remarks about the appearance or sexual orientation of an employee, verbal or written comments of a sexual nature, or excessive and unnecessary physical contact or sexual advances and requests, to serious physical and/or sexual abuse.

abusive and unfair psychological violence over another individual. This may be done systematically and recurrently over a prolonged time, to break down the victim's communication networks, destroying their reputation, undermining their self-esteem, disrupting the performance of their duties, deliberately degrading their work conditions and, finally, forcing the affected person to leave their job, producing continued and progressive harm to their personal dignity. So, we asked respondent workers if they have ever been mobbed in the workplace, due to their sexual orientation and/or gender identity:

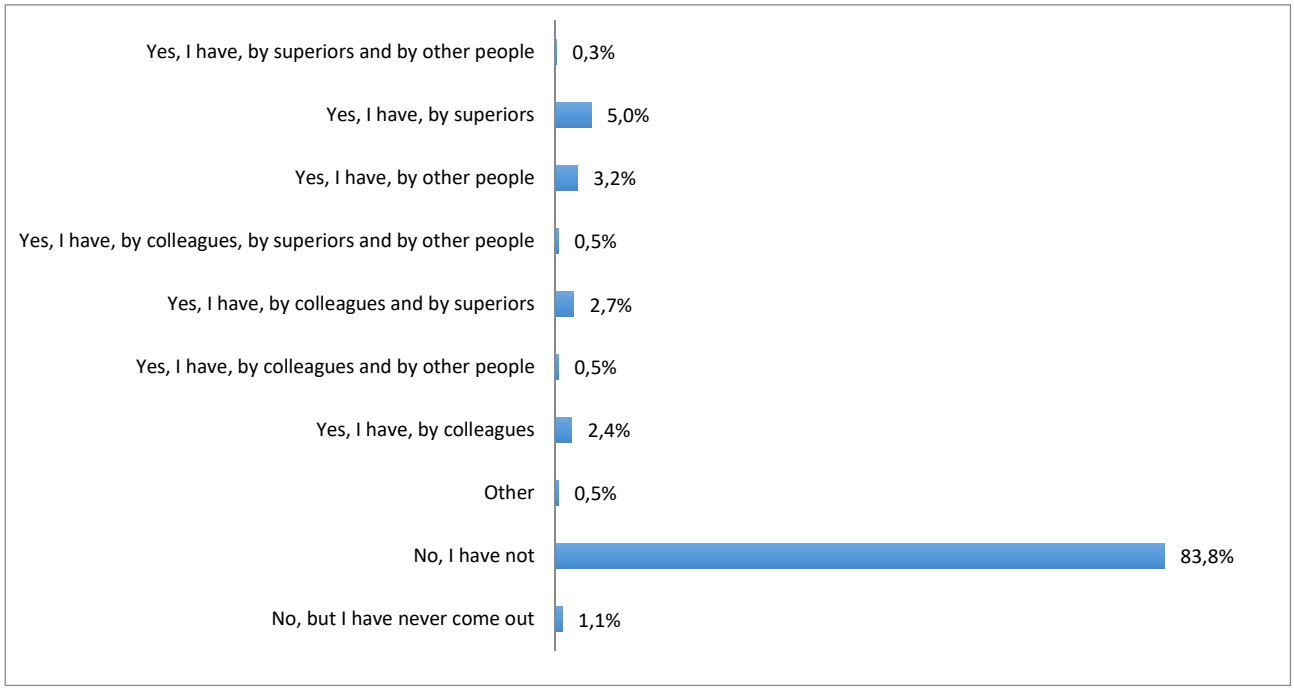


Fig. 36 – Respondents distribution, by mobbing experience in the workplace

At work, those who have suffered mobbing for sexual identity reasons represents 15% of working respondents. Unfortunately, the one responsible for these oppressions is often the boss, who wields the power of impeding the professional growth of an LGBTQ person. Colleagues, in fact, are more compliant than superiors (respectively, 2.39% vs. 5.04%), although mobbing behaviours are intolerable even when committed by a peer. 1.06% of workers say they have never suffered mobbing, but (in the open modality) they say that they have never even come out, "to avoid any kind of problems": this issue is recurrent and important because if on one side coming-out is a right and not a duty, mainly at work, on the other side, avoiding coming-out due to fear of retaliation, is a full-fledged form of discrimination, even if indirect and precautionary.

Percentages registered are honestly lower than those obtained by other surveys on the same subject (Istat, 2011; Lelleri, 2011; D'ippoliti and Schuster, 2011) and, sincerely, we do not feel like drawing optimistic conclusions based on the OTR survey data, or generalizing results on the whole LGBTQ population(s). Furthermore, it is important to say that mobbing occurrence is not even significantly different by gender and sexual orientation, and percentages by subgroups are close to the average.

This issue highlights a methodological piece of advice: for studying complex phenomena, related to the labour market, and giving reliable estimates on several domains, such as gender, sexual orientation, type of job, etc. survey sample sizes have to be consistent, mainly for the most under-represented groups belonging to the LGBTQ population(s): according to literature (Graglia, 2012; Whittle et al., 2007), transsexual people, particularly, are the most defenceless victims of a recruiting business system, and data about them is often lacking and inadequate.

In Italy, difficulty in data collection makes it hard to help policymakers plan working inclusion policies, already in force for some years abroad: article 27 of the Italian law 198/06, which prohibit discrimination in the labour market, makes no mention of sexual orientation, or transsexuality, as causes of discrimination, leaving LGBTQ people in a condition of "legal vacuum", and lack of protection. Abroad, conversely, sexual discrimination is an issue of great sensitivity for social researchers and economists. Some of them would quantify the "exclusion price": the failure by enterprises to hire LGBTQ people is not only a personal problem, of those who suffer from this kind of discriminations but also irrational wastefulness of human resources, which entails economic and social damage unfairly paid by our societies (United Nations, 2017).

11.4 At school and university: being bullied for sexual identity

If the smaller sample size of workers' respondents allows us to draw cautious conclusions about mobbing behaviours, the very large share of (current and past) students, allows us to examine discriminative dynamics of the classroom, with greater awareness.

In the survey questionnaire, after having deliberately used the word "mobbing" to study discrimination at work, we have used, just as consciously, the word "bullying", to investigate discrimination, specifying as its cause the

sexual orientation and/or e gender identity, at both school and university, and by whom.

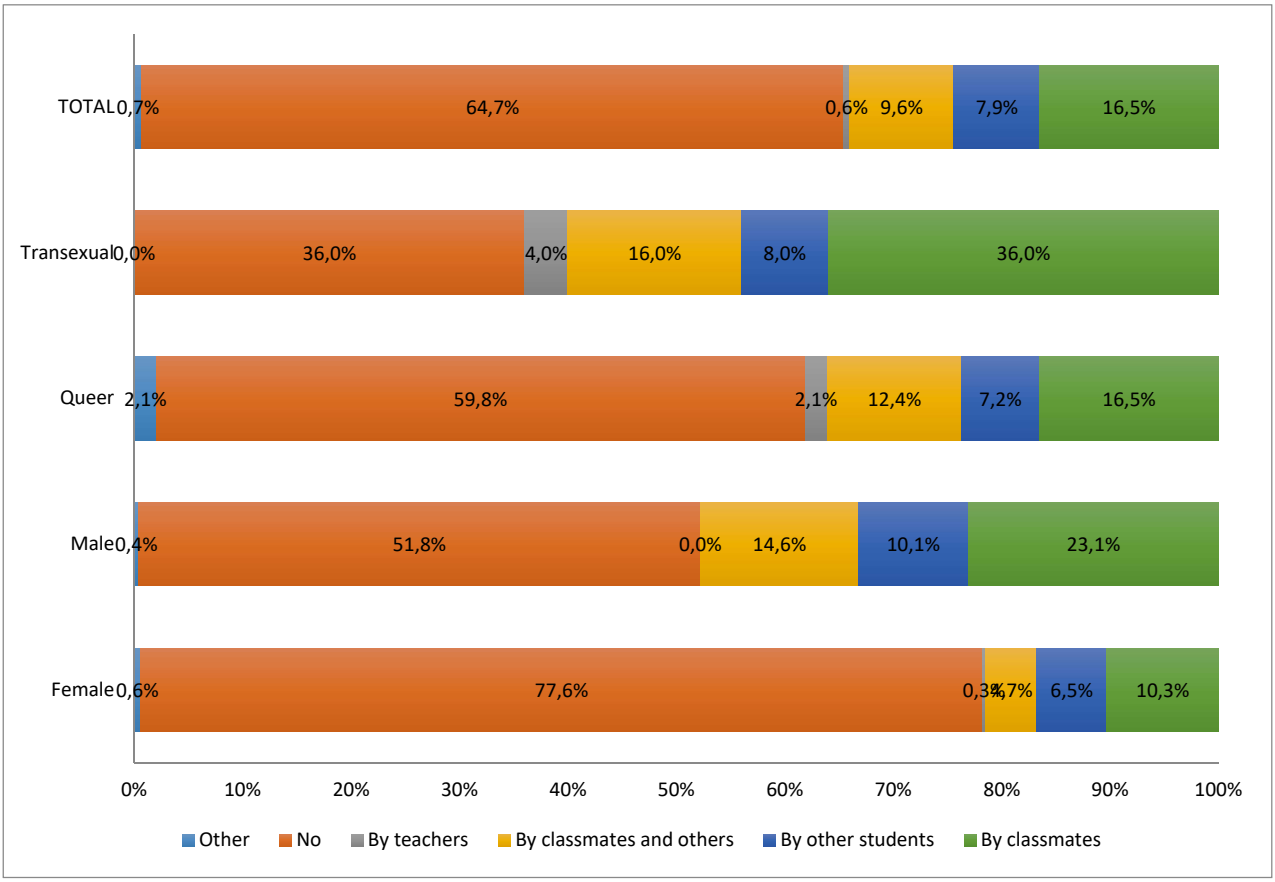


Fig. 37 – Respondents distribution, of bullying experiences at school and/or university, by gender identity

School results to be a less safe place than the workplace: the percentage of students, altogether, bullied for their sexual identity is around 35%. Usually, bullies are classmates (16.43%), a student belonging to another class (7.8%) or both (9.61%). Fortunately, only 0.56% of teachers committed these discriminative actions towards their students.

In the graph, we have reported percentages divided by gender identity because differences among genders are sensational: if 77.65% of females have never experienced bullying acts due to their sexual identity, this percentage drops to 59.8% for queer people, 51.82% for males and even 36% for transsexuals. Classmates are those who most likely bully LGBTQ students: 23.08% of male, and 36% of transsexual, respondents have been bullied by one or more classmates.

School should be, according to Graglia (2012), the ideal context to reach the largest number of adolescents to deal with crucial issues regarding different

sexual identities, where young heterosexual people can learn respect towards difference, and where non-heterosexual people can learn to be who they are. School, first, and university later, instead they represent a difficult context in which to act on, because of a set of reasons, nested and deeply rooted, firstly, reluctance to introduce sexual education and gender studies in ministerial programs (Bernini et al., 2015). These reasons mean that, in our country, not enough has been done to prevent bullying based on sexual orientation: reducing bullying would seem to be a goal that everyone could support, but the acceptance of sexual orientation as a reason for bullying is objectionable to some policymakers. Being a straight teenager is difficult enough, but being a teenager struggling with sexual identity brings its own set of issues, including the potential for harassment (Russel et al., 2010).

LGBTQ students, left to find a safe place on their own, seem to search for them also through particular formative trajectories, both in secondary school and university. The graph below shows student respondent distribution by type of diploma they are earning, and distribution of graduates of Italian high school, in 2015[25].

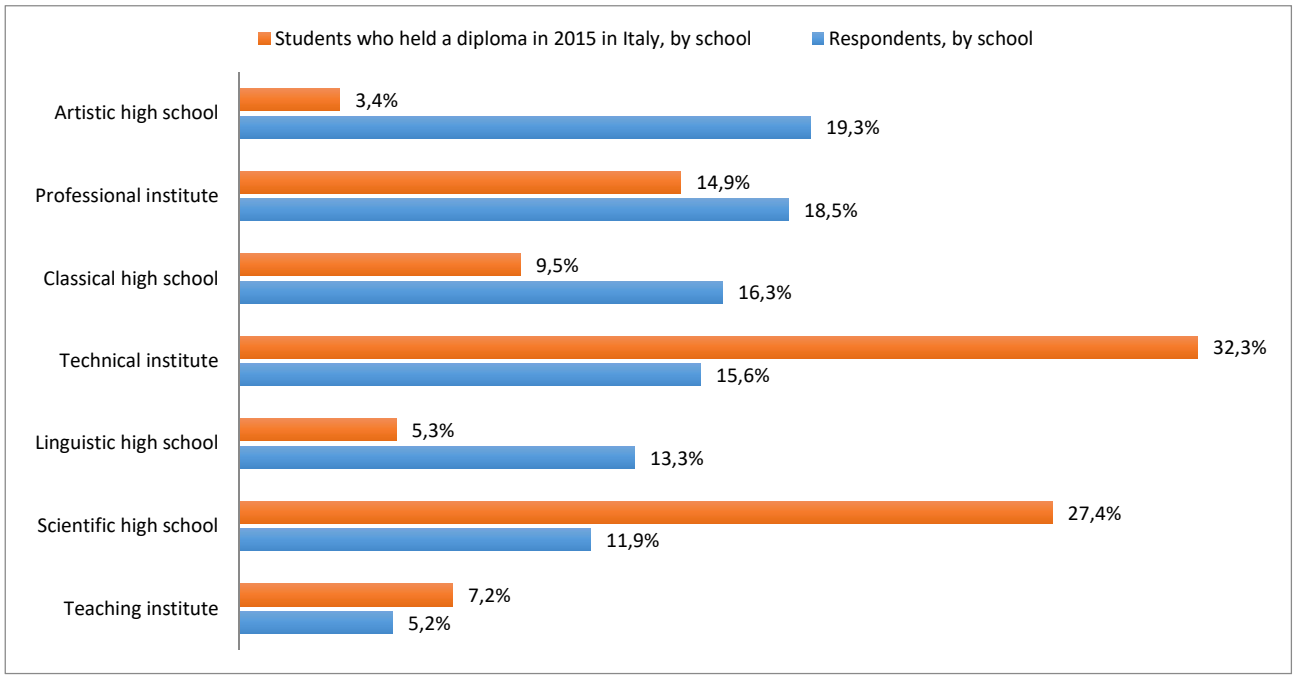


Fig. 38 – Distribution of respondents, and graduates in Italian high school (2015), by type of high school

[25] Istat's Survey on educational and professional routes of upper secondary graduates

Respondents graduating class is homogeneously divided by school type, but with a strong bent for artistic (19.26%), professional (18.52%) and classical (16.3%) instruction. Although homogeneous, this respondent repartition among school type does not reflect the national student distribution. Globally, the most frequented schools are technical institutes (32.3%), scientific high schools (27.4), professional institutes (14.9%), classical high schools (9.5%), teaching institute (7,2%), linguistic high school (13.33%), and finally artistic institutes and high school (3.4%).

It seems that most of the high-school respondent students, almost all belonging to the LGBTQ population(s), chose "formative trajectories" quite different from what was chosen by their peers. Artistic education is a clear example: only 3.4% of graduates in Italian high schools attended an artistic institute, but in our sample, 19.26% of respondents are going to that kind of high school.

Also, university students respondents often frequent certain faculties, which are not the most frequented by the whole university population. In the graph below, we report respondents' distribution by academic area, compared with the distribution of students in Italian universities in the academic year 2018/2019.

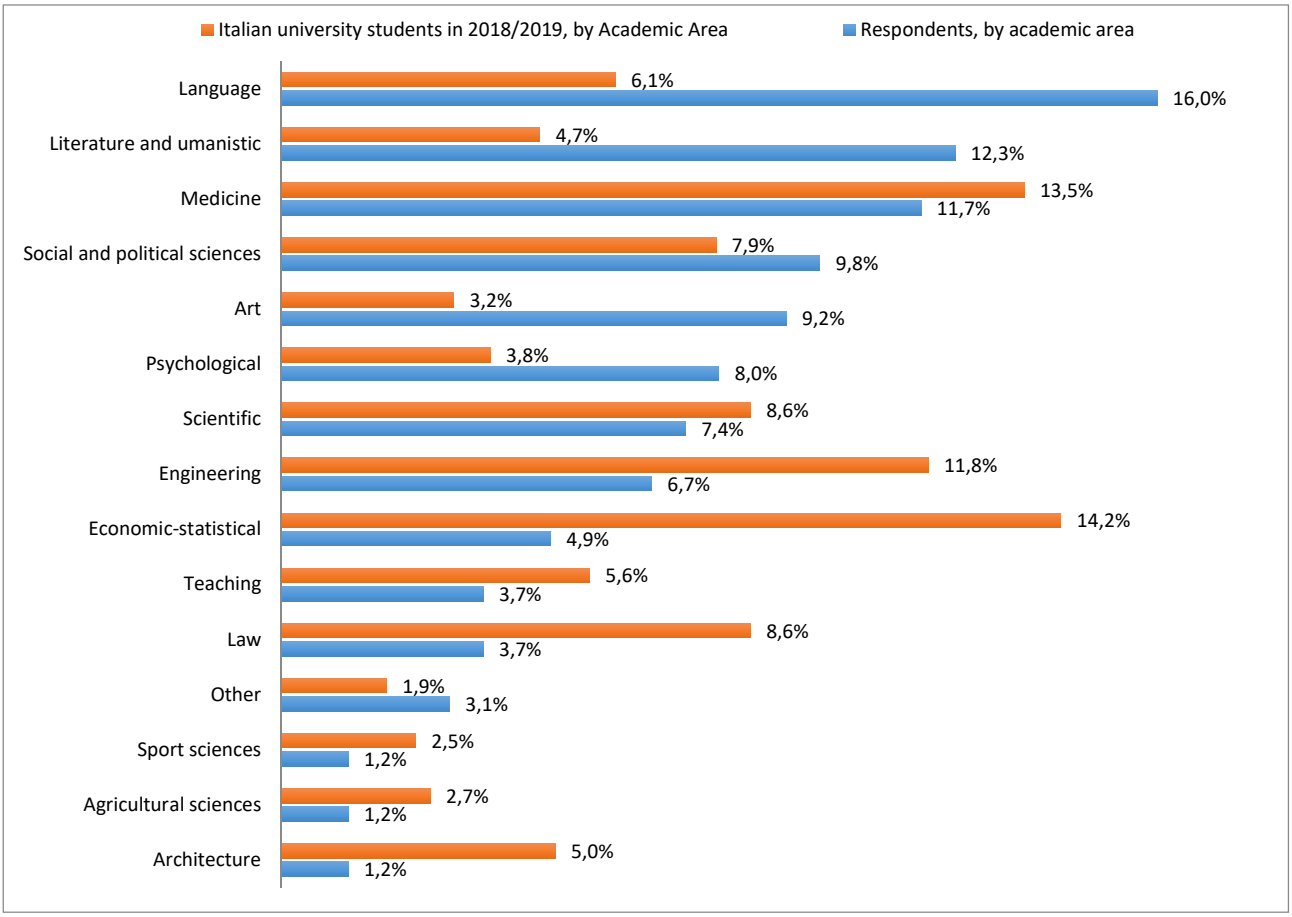


Fig. 39 – Distribution of respondents, and graduates in Italian universities (2018/2019), by academic area

40% of respondents are preparing for a university degree and we asked them on which academic area, according to the classification used by Istat and MIUR. Respondents prefer humanities faculties: such as languages (15.95%), literature (12.27%), medicine (11.65%), political sciences (9.81%), art (9.2%), psychology (7.98%).

The same academic areas, considered on the total of all Italian university students, report very lower percentages: languages (6.09%), literature (4.71%), political sciences (7.92%), art (3.15%), psychology (3.81%).

Differences between observed percentages on the realized sample and those on enrolled students at a national level, both in schools and in universities, are significant. In previous chapters we did not hide that sample result were unbalanced, by age, gender and digital alphabetization; however, the imbalance towards some disciplines seems to be more than a mere sampling bias. It looks more like young students, who become aware of their own identity, choose a particular formative curriculum, coherent with their needs. Sociological literature highlights a persistence of a relevant imbalance in the choice of the field

of study both at a secondary school and at a university level (Gasperoni, 1996; Benadusi et al., 2009; De Vita and Giancola, 2017). These differences are explained through gender dynamics which address formative and then professional choices (Bocchiaro and Boca, 2002). These dynamics could address, in a similar way, the choice of the field of study of many LGBTQ people, who choose high schools and faculties where they tend to feel safer, free of being who they are and accepted. Anyway, leaving many young LGBTQ students to find their own safekeeping, without any kind of support, can be considered as an institutional fault: to ensure that school learning environments are supportive of all students, it is crucial to implement specific policies and procedures that support LGBTQ youth, affirm their identities, promote safe and healthy learning environments, and advance equity and respect for all, both in schools and in universities.

11.5 On social media: the deadly combination of cyber-bullying and homophobia

In the ninth chapter we looked over the influence that social media has had on LGBTQ people and their community, in the last two decades, affirming and performing in a new way gender identity and sexual orientation on the web; moreover, previously, in chapter 5, we saw how social media, and more generally big data, introduced new possibilities in social research, also in the sensitive field of gender studies. The prospective taken so far, although not rhetorically optimistic, highlights many positive aspects of social media, without hiding the existence of criticalities in the use of these technologies and tools.

Talking about discrimination we asked the people involved in the survey "have you ever been teased on a social network?", and by whom, testing the safety of social media and the incidence of all those behaviours generically called *cyber-bullying*.

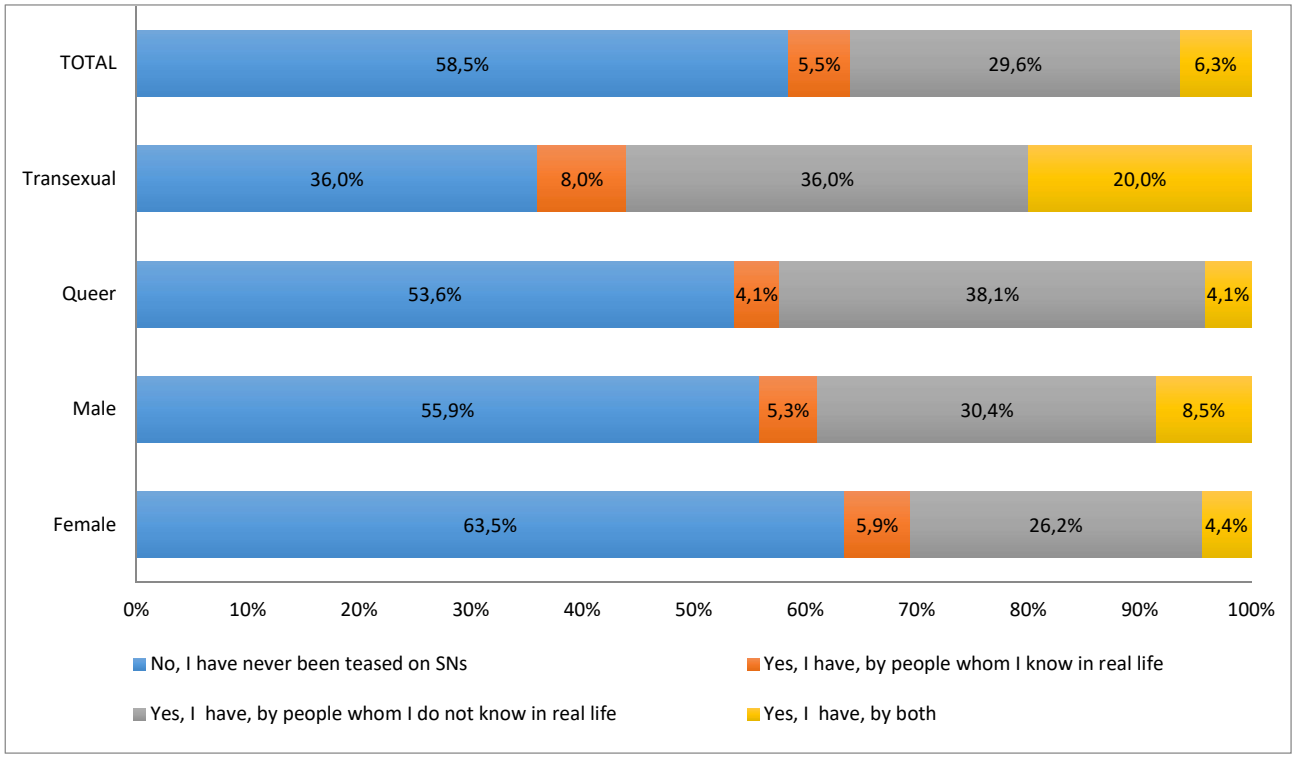


Fig. 40 – Respondents distribution, by cyber-bullying for sexual identity experience, by gender identity

58.53% of respondents have never been teased on social media for their sexual orientation, and/or gender identity. When it happens, usually, the guilty person is a stranger (29.62%). Discrimination on a social network, as that one at school (Fig.32) is often against transsexuals (64%), queers (46.39%), and males (44.13%). Transsexuals turn out to be, also in the virtual context, the most harassed people, usually by strangers (36%).

If cyber-bullying becomes the technological evolution of bullying, social media enables bullies to torment people from a distance, often unknown too. Homophobia and cyber-bullying, together, become a "deadly combination", which sometimes push young LGBTQ people to commit suicide (Wiederhold, 2014; Blumenfeld and Cooper, 2010).

Bullies are usually sure to be protected, because victims often do not denounce what happens: they usually leverage shame, both in real life and on social media. Moreover, the LGBTQ victim fears retaliation denouncing, mainly if they are not yet out with friends, parents or relatives.

In 2018, the most important Italian LGBTQ association (ArciGay) carried out a European project, called "Accept" to identify and contrast homotransphobic hate spread by social media. The project aimed to recognize, through a textual analysis procedure run on Twitter's tweets (Daas and Puts, 2014; Ceron et al.,

2013), sets of words used in some contexts to harass a person or a whole community. The project divided a sample of tweets about the LGBTQ community into positive, negative, and ambiguous; successively the most frequent words used in these three categories have been listed. While words used in negative posts are clearly homophobic, such as *faggot, paedophile, repugnance, depraved*, etc., words used in ambiguous posts are often ironic (*privilege, opinion, heterophobia, joke*) and used in tweets which of course are also discriminative, but subtly. This textual analysis project, as well as being an interesting example of social research on big data, shows how, and how well, words can be brutal towards LGBTQs on social media: discrimination should not be thought (only) as a person hit, because they are gay, or lesbian. Discriminative phenomena are often much more complex, articulated and concealed than we usually imagine.

While on one side we should take action, mainly among young people, to educate for the respect, diversity, and inclusion, annihilating any bully and homophobic behaviour, on the other side many people directly turned to social media staff, demanding them not to be means of dissemination of hate and discrimination messages. The majority of social media agreed to the request of users who stood up and introduced different tools to prevent and block bullying. Facebook, which is probably the most unsafe website from this point of view, developed a "Bullying prevention hub", aimed at blocking all those users who do not respect the netiquette about cyber-bullying towards other users.

Social media, as already said, are new full-fledged social space subjectifying, self-expressive and free, though subject to their netiquettes (Snee et al., 2016), and they have to be treated as such, with their strengths and their weaknesses. Many efforts are being made by social media companies to tackle the cyber-bullying issue, for instance trying to not give any space to homophobic posts, comments and virtual negative behaviours. However, the cyber-bullying solution seems to be indissolubly linked to its previous stage, traditional bullying, which often takes place at school, and does not only affect an individual during childhood, but can have a lasting effect on their lives well into adulthood. By effectively preventing and tackling bullying, schools should create safe, disciplined environments where students can learn and fulfil their potential. If that happens, there will be no more space both for bullying and for homotransphobia, not in real life, at least not on social media.

11.6 In sport: beyond traditional sporting hegemonic masculinity

Sports and homotransphobia are two deeply linked issues, although they can seem far apart. Some current tragic events led journalists to deepen these links, but sociological literature, instead, runs low. Some researchers of the University of Napoli (Amodeo et al., 2017; Cuccurullo and Casolare, 2013), collaborating with the SINAPSI university listening centre, carried out very interesting research on this subject, giving a sociological and anthropological interpretation to the interconnection between sport and homotransphobia.

According to Casolare (2017), the human being can be considered a social animal, based on Aristotle's traditional theory. Sport, like other social institutions, follows this theory and, collaterally, accepts gender stereotypes which inevitably led to sexist and homophobic stereotypes. Sport becomes, based on this point of view, a sexualized environment, and sexuality its "backbone". That is the reason why it is difficult to organize, into sports societies, awareness and information campaign against genderism, sexist and homophobic prejudice. The whole idea behind these researches is the same which emerged in one of the interviews carried out during the design phase of the survey: the shortage of LGBTQ athletes, mainly males and in some disciplines. So we have asked people involved in the survey "have you ever been discriminated in sport, due to your sexual orientation and/or gender identity? If so, by whom?", successively we divided response distributions by gender, as shown in the chart below:

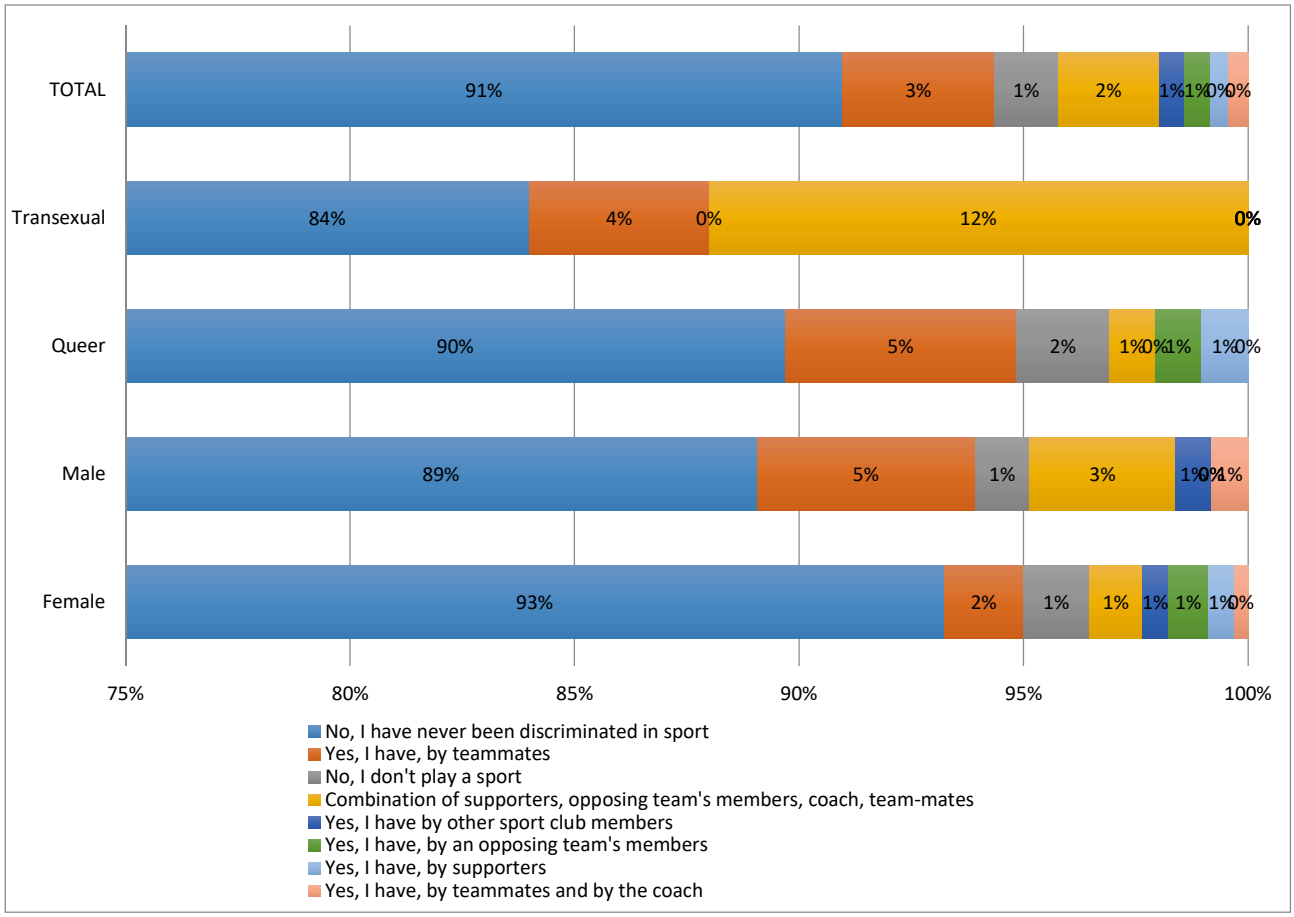


Fig. 41 – Respondents distribution, by discrimination for sexual identity in sport, by gender identity

The graph shows that the sporting environment seems to be safer than school, work, virtual reality and family. 91% of the survey's respondents have never experienced harassment, relative to their sexual orientation or gender identity, playing sports. Those who have suffered these awful experiences, did so because of a team-mate (3%), their own age and a peer. Discriminations are instead less frequent by supporters, coaches, team members, and adversaries. Unfortunately, this information does not take into account if respondents came out in this sporting environment, or not. Differentiating would be substantial and could explain subtle differences reported among gender identities: if on average, as already said, 91% of respondents have never experienced harassment, this percentage rises to 93% for females, and drops to about 89% for males, 90% for queers and 84% for transsexual people, often discriminated in sport, by more than one person, among team-mates, supporters, coaches, team-members and adversaries (12%). 5% of male respondents have been harassed by one or more teammates.

Low discrimination percentages could mean that the sporting environment is ultimately safe, or that people usually do not come out, sometimes for fear, sometimes for inadequacy. We can explain this interpretation through the theoretical frameworks of stereotyping and gender roles: the male gender is identified with the traditional cliché of the strong and virile man, exasperated in some sports such as football, boxing, rugby, basketball, hockey, etc. A prototype to fully embody (otherwise the marginalisation) in contraposition with what is not typically as manly and masculine: women and homosexuals. This concept represents the transposition of Connell's "hegemonic masculinity" (2005) to the sporting field. This hegemony can be considered as a stereotypical notion of masculinity that shapes the socialization and aspirations of young males (Ricciardelli et al., 2010), and is often represented in literature as a pyramid; on the lowest rung of the pyramid, we can find women and homosexual men. From this biased point of view, female homosexuality is not considered so weird, because the stereotype of the lesbian woman, androgynous and graceless, fits with the figure of the agonist girl. Besides, the hegemonic masculinity, distinguished in *male sports*, carried out by "real men" and "lesbian women", such as above mentioned football, boxing, rugby, basketball, hockey, and *female sports*, carried out by women and gay men, the most obvious example of which is ballet. Transsexual people are never considered. This strict sexist bipartition is usually called "sex-typing sports" (for instance, artistic gymnastics disciplines, some for men only, others for women only), and it is widely accepted by the crowd, unaware that this genderization inevitably leads to stigmatizations, discriminative dynamics and strong prejudice. In manly sports, coming out thus becomes an inconvenient choice for gay men, which likely would be marginalised by a radically male and heterosexual system; on the other hand, in the same sports, coming out for lesbian would be considered pleonastic, in the same measure as it would be considered superfluous the coming out of a male ballet dancer. In doing so, prejudice causes taboo and triggers so creating a vicious circle from which there will be no escape until someone makes a stand against it, not only coming out, but challenging the whole system which caused it.

Concluding, we can specify that sporting discriminations percentages are very low, but usually referred only towards athletes who do not come out, for fear, as gay men in manly sports, or because their homosexuality is already taken for granted, as woman footballers. However, fear and prejudice are in turn discriminative behaviours. The sporting environment thus could be much less safe than data tells us, subject to the hegemonic masculinity which perpetuates the traditional and rusty bipartitions of men vs. women and heterosexual vs. homosexual.

Survey results: concluding remarks

Survey data collection pointed out some remarkable aspects, deeply analysed in the previous paragraphs. These aspects can be divided into methodological features (linked to adopted survey technique and data collection) and contents, gleaned from the survey's answers. Some of them can be considered unattended, whereas others confirm the results of other studies carried out on LGBTQ issues.

Regarding methodological issues raised, can be worthwhile to think about what it meant to carry out a survey entirely on social media, by posing some questions to ourselves: can these platforms be used for further social researches? Who are the respondents? Do the results confirm the reflections of the drafting phase? Let us analyze these aspects singularly:

- The trend of monthly survey responses (between 1st April 2019 and 31st January 2020) was decreasing, mainly due to Instagram's anti-spam policies, which blocked temporarily, and repeatedly, the messages sent by the OTR account, suspected of being a *spammer*. This feature confirms an important methodological aspect: the complete dependence on social media. Obviously, the messages sent were not spamming, but blocks are like "sentences without any appeal". We do not know who labels the messages as spam, if a person or an algorithm. Anyway, there is no interlocutor with whom to talk with and to ask for explanations. This criticality cannot be ignored if we plan to carry out social research, on social media, contacting a massive number of users.
- Despite anti-spam blocks having been a big problem during the data-collection phase and having caused a response number lower than expected. The response rate, given by the ratio of 718 valid responses out of 4854 contacted users, is 14.8%. For being a CAWI survey entirely carried out on Instagram the response rate can be considered satisfactory, both in absolute and relative terms. Moreover, the high propensity of responding people who provided us with other 638 usernames to interview in turn,

under the snowballing scheme, is remarkable and a good signal of the availability of respondents towards the aim of the survey.

- The population(s) involved in the survey is young, emancipated, more frequently female, and spread all over the Italian territory. Being our sample non-probabilistic at all, we cannot strictly talk about representativeness. Anyway, the characteristics of people caught can be considered a result too. For instance, the higher collaboration from women is important feedback about the web-scraping phase: in fact, also if most of the collected post by the web-scraping script (according to the hashtags contained) were male-oriented, most respondents are female. Concluding, it is important to say that many people would have been uncatchable, if not on the internet: for instance, 46.8% of respondents do not frequent any typical LGBTQ meeting places, such as bars, associations, etc, and so they would have been completely invisible to any traditional survey.
- The attention paid, during the drafting phase, in studying an efficient questionnaire paid off: respondents exploit the open modality answer to define themselves; moreover, phenomena emerged during the semi-structured interviews, the extensive use of dating apps, gave interesting results, both from a quantitative and from an interpretative perspective.

The survey, being both a methodological experiment and social research about LGBTQs, does not show just methodological tips, but also important empirical evidence, as already said interesting and often unexpected:

- social media has had an epochal impact on LGBTQs daily life, redesigning the social process of self-expression, socialization, making community, etc. but dating apps have had an even stronger impact: they have made meeting new people extremely easier, but not only; dating apps have also changed the way of starting relationships, uninstalling these apps, redefining starting and ending points of a love affair. Anyway, it is fair to say that dating apps represent a cumbersome intermediary between two people, introducing problems of self-construction in a computer-mediated environment, and of embodiment in the digital age (Mowlabocus, 2010). Furthermore, these problems are strongly different among LGBTQ subpopulations, and they can depend on different individuals' safety

perception as well, which, in turn, is strongly correlated with the different gender identities: for instance, results show that 88.1% of male respondents have met other dating app users personally more than once, but this percentage decreases under 60% for queer people, and under 50% for female and transsexual people. Analogously, the percentage of who still use dating apps is really variable, according to gender identity: it goes from 44% of males, to 28% of transsexuals and queer people, to 14.4% of females.

- LGBTQ socialization could be at a turning point: no more only "exclusive spots", such as cruising bar, and local circles or associations, often located in an entertainment district, frequented by a clientele almost completely LGBTQ. Just under half (48%) of respondents still, usually go out in those places. Nowadays, socialization may have further changed, due to the spread of digital technology, which increases the possibilities of meeting people.
- Also the associations' galaxy could be at a turning point, but their institutional centrality remains crucial, especially for those who recognize their support authority (psychological, social-sanitary, and social-assistance), and their strong political weight of struggling for civil rights of all LGBTQ people. That could be the reasons why transsexuals (36%) and queer people (33%) are more likely to be an integral part of an association.
- In paragraph 10.2 we stressed the concept of community, according to Tonnies's definition, to highlight how using the term "*LGBTQ community*", indiscriminately, as an expression of speech, could alienate some people, often already marginalised. This point of view helps to understand that only 45.5% of respondents unconditionally feel part of the LGBTQ community. 38.4% have some reticence, and feel themselves "partial belonging" to the community.
- Finally, in regards to the discrimination issue, we distinguished five different discriminating contexts: household, work, school and university, social media, sport. Strong gender differences have been reported among discriminated people: transsexuals are more likely to report every kind of discrimination and harassment, within the household (60% transsexuals are told to hide themselves, 52% to have caused great sorrow to their

parents, and 24% have been treated as a plague spreader by a relative), at school (64%), online (64%), and in the sport too (16%).

- Speaking of sports, in the last paragraph, we gave an interpretation for the very low discriminative behaviour percentage (8%), which could mean that in sports people usually do not come out, both for fear, and for incongruity to the athlete stereotype.

So, to recap, survey results give both methodological suggestions for the study of the LGBTQ population(s) and new knowledge about it. These results can be considered as the direct output of an exploratory survey, useful for further more structured and wide surveys, but also as a deepening for all those studies about the LGBTQ population(s) in Italy. Data points out how multifaceted the LGBTQ population is, and, at the same time, how important it is to rethink the methodology used to properly investigate it.

CONCLUSIONS

During the writing of this thesis, throughout 2020, a lot of important things, concerning the LGBTQ community, have happened.

In Italy, in October, the Italian Medicines Agency (AIFA) made hormonal therapy free for transsexual people, guarantying access to indispensable medicines to who is passing from one sex to another through hormone replacement therapy.

At the end of the same month, Pope Francis, in a documentary about his pontificate, explicitly supported the civil unions' law between same-sex people, because "they have the right of being legally protected". This declaration stirred up discontent among the most catholic and conservative people, who, however, did not pay attention to the words used by the Pope (civil unions, and not marriages) who just reaffirmed a position clear by years.

Remaining in Italy, homophobia events have unfortunately continued to happen. Some of them were ferocious and touched the soul of public opinion, sparking a lively debate: the most heinous of these events happened the night of 13th September 2020, in the outskirts of Naples, when a young man rammed the scooter on which his sister was travelling on, Maria Paola, along with Ciro, her transgender boyfriend. In the scooter accident after Maria Paola fell, her brother started to hit Ciro, instead of assisting her, accusing him of being "guilty of having infected" his sister, who died in the meanwhile. This violent piece of news got public opinion talking of homophobia, but, at the same time, showed the huge confusion of journalists towards the concepts of gender identity and sexual orientation, confusing often different words such as transgender, transsexual, lesbian and bisexual.

This increasing attention to the gender issue, anyway boosted the legislative process to draft a law against homotransphobia. This bill specifically contrasts every kind of discrimination and violence for sexual orientation, gender identity and also disability reasons. This draft legislation, named Zan as the surname of its promoter, was approved by the Chamber of Deputies in November, causing widespread uproar among right parties which defined the bill as "liberticide" and invoked the "crime of opinion".

Also abroad many important things, concerning the LGBTQ community, have happened: in Belgium, the Member of the European Parliament Petra de Sutter, became the first European transgender minister. In June, it had happened for the first time in the world in Taiwan, to the minister Audrey Tang, MtoF too.

If some countries recognize diversity as a resource, others regress in the inclusion path of LGBTQ people: one for all, Poland, presided over by Andrzej Duda. The Polish government defined homosexuality as "a perverted ideology, which violates the law". The same government, moreover, encouraged the local authorities and cities to declare themselves LGBT-free, and to emarginate the LGBTQ citizens: over a third of polish municipalities has joined this absurd initiative.

These recent events have rekindled the interest of institutions and policymakers on the need for data to devise targeted measures to ensure the respect of the fundamental rights of LGBTI people (FRA, 2020).

In this respect, it is important to say that in Italy, Istat conducted, for the first time, in 2020 a survey about "LGBT+ diversity management policies in enterprises". To complete the informative framework of official statistics the three following surveys have been planned:

- a census of LGB people who have got into a civil union, since 2016;
- a sampling survey on LGBT+ people (not into a civil union), selected through an advanced snowballing sampling scheme;
- a focus on transsexual people who enjoy health services and help desk, selected through a non-probabilistic sampling scheme.

These surveys on the various LGBTQ subpopulations will be carried out through a CAWI[26], during 2021. Their goal and their methodologies confirm a research trend, explicated in the second chapter, talking about European studies: the proportion of LGBTQ people on the whole population, has lost its appeal, mainly in Europe, in the eye of policymakers. Governments, ministries and all the other institutions principally want data about negative behaviour towards LGBTQ for addressing their policies (FADA, 2017). In this way, LGBTQ research has changed, abandoning the aim of finding the definitive number of people, and starting to quantify the phenomena related to them. For instance, investigating enterprises in which LGBTQ people work.

[26] Computer Assisted Web-Interview

The Over The Rainbow research project aligns itself to the European trend of abandoning the idea of estimating the exact number, or its proportion, of LGBTQ people overall population. Rather, the research project intends to assess the magnitude of LGBTQs behaviours and of society towards them; both good behaviours, such as working inclusivity and family acceptance, and negative behaviours, such as sexual discrimination and harassment. Furthermore, for obtaining results as reliable as possible, both qualitative and quantitative methods have been used, consistently with the mixed methods research, largely used in Italy on gender issues. The present research project, also if with all its operative weaknesses, tried to reproduce the strengths that emerged in the sociological and statistical context, both at the national and international level. On the other hand, the peculiarities of the studies which did not give any reliable results to the community and policy-makers have been accurately avoided, not representing the sociological research purpose of this study. For instance, the clear commercial purposes of the previously mentioned private opinion polls, or the market research, do not give any added value in terms of social science data for policy analysis and policymaking, and they have not been pursued in this research project.

Once framed the epistemological and methodological context, according to the definition matter, illustrated in chapter 1, and the reference quantitative studies (chapters 2, 3, 4) in which this research project is located, the present project particularly focuses on the methodology to use for studying such a characteristic population. The attention paid to the methodological issue, anyway, did not preclude the possibility to obtain interesting empirical data, on subjects often almost unexplored: survey results point out that most of the respondents (46.8%) would never have been involved in a survey, if not through a social media. This result answer the methodological research question posed in the introduction, answering that yes, big data, coming from social media, could be the right tool for studying the LGBTQ population(s), but with the due precautions and reservations, which reside in the cryptic and private nature of social media. Furthermore, this result, besides giving an interesting methodological suggestion, gives important informative content. The massive use of social media and dating apps redesign the border of the LGBTQ socialization, too often incorrectly ascribed to LGBTQ-friendly bars, clubs, associations and organizations: 76.2% of respondents agree with the sentence "Associations are important for LGBTQ rights claims [...] but to socialize

nowadays there are more efficient tools available". At the same time, 53.5% of respondents do not agree with the sentence "social media are substituting LGBTQ associations, as places for sharing sexual orientation and gender identity experiences". The combination of these results, highlights a sort of process of secularization of LGBTQ associations, which do not have to raise concerns. On the contrary, LGBTQ associations remain fundamental as political pressure institutions on the front of civil rights recognition and the fight against homophobia, biphobia and transphobia (Barbagli and Colombo, 2007), as already emerged in the interviews carried out during the design phase. Results suggest that associations are more frequented by people who self-define themselves as queer (33%) or transsexual (34%), rather than male or female gender people. Queer and transsexual people, by definition, give a strong political value to these terms, and this result can show that politics has increasingly permeated the organizations, in the best sense, modifying their goals and their ways of acting the civil commitment, as it has happened to all other institutions involved in political activism, such as political parties or trade unions (Simon Rosser et al., 2008).

All the information resulting from this research project can usually be analysed in a twofold way: the content, linked to the information regarding the respondent people involved in the survey, and the methodology, regarding the modality by which the study was carried out. Every graph, every chart, every percentage in the tables, tell us something about LGBTQ people, but also on how to properly investigate the features of this population(s), starting from the people who are part of it. In fact, the study of a population can never overlook the study of its individuals, and in this context, this research project, leveraging its expertise, want to try to develop *guidelines* for all those researchers and scholars who will further deal with the LGBTQ issue. We want to weave the concluding remarks that the reader found at the end of the three different parts of the thesis, highlighting some fundamental issues to focus on, during the design and the projection of social research about LGBTQs, especially if through a survey. These issues to carefully monitor, are:

1. *Actors engagement*

Create engagement with potential respondents, or their community, is the key factor to obtain a respectable response rate and reliable answers in projecting a survey. The aim of engagement's creation is, firstly, to reassure about privacy, data-safety and anonymity conditions, but also to explain the final aim of the research, persuading people to collaborate.

Working on social media we tried to create this engagement feeling through an Instagram page (@overtherainbowproject)[27]. This page, besides contacting selected people through Instagram direct messages (DM), has been used to explain the aim of the research project, to follow users belonging to the LGBTQ population(s) (both to contact and not), and to post pictures related to the survey's issues, such as Pride month, rainbow flag, reference books about LGBTQ, preliminary results of the survey, and results of other international surveys. These posts reasonably created a feeling of trust and confidence among the users, who started to constructively interact between them and with the admin of the page, in DM: most of the contacted users texted "done!" once filled the questionnaire, often thanking and giving positive feedback about the questionnaire, sometimes asking for clarifications or results of the survey. Furthermore, sometimes, followers asked to follow them back or to share their profiles to meet new users, exploiting the project's page as a "sounding board" for reaching as many LGBTQ people, as possible.

From the engagement point of view, Instagram was a great platform from which to lead a survey, because people were already friendly and comfortable with this app/website.

Instagram's page of Over The Rainbow will be online until the end of the process of dissemination of results, so that interviewed people can receive the direct output of research project they collaborate with, in the form of graphs, charts or infographics.

2. *Clear (but not stringent) definitions*

Of course, definitions are tricky in this field, but defining a proper set of operative and functional definitions is crucial for prolific research. As we have accurately

[27] https://www.instagram.com/overtherainbowproject/

seen, categorizing social gender identity, gender behaviour and sexual orientation introduce noteworthy considerations, both from an epistemological and methodological point of view. Every classification can be considered as heavy and rusty structuralism to avoid: more and more people recognize that sex and gender have separate dimensions and more than two possibilities. Social researchers, therefore, have to propose sexual orientation and gender identity (SOGI) questions, which have to indisputably refer to gender and queer studies, to mediate between the complexity of reality and the feasibility of the social investigation. In this context, the choice of leaving an open modality to self-define themselves, their sexual orientation and their gender identity, can be seen as the operationalisation of the Queer studies' main concept of the "indefinable sexuality": trying to catch and categorise every SOGI shade, in a survey, could be really a hard and unsuccessful job, and it is, therefore, better to avoid heavy and strict classification of such a complex and sensitive concepts.

3. The interdisciplinarity of the research-team

Conducting a quantitative study on an LGBTQ population(s) needs teamwork. A close-knit staff composed of people having various backgrounds, so that they can face off every kind of pitfall: sociologist for theoretically interpreting data, statisticians for designing the sample, IT experts for the data storage and collection, LGBTQ activists for drafting and testing the questionnaire, organizations for promoting the survey and broadcasting results, communication and social media experts for administrating the account which informs and updates users, legal consultants for guarantying privacy conditions... you name it, I put it.

Moreover, the research on this field is progressing towards new applications and new techniques, that will require specific skills and background: for instance, in paragraph 11.5, we introduced Arcigay's European project called "Accept", which exploited a textual analysis procedure run on Twitter, to recognize a set of keywords used in online context to harass an LGBTQ person or the whole LGBTQ community. The plurality of research techniques that will be put in place for studying the LGBTQ population(s), will enhance the knowledge about it, but it will require ever more competencies, informatics infrastructure, and a multi-purpose research team.

Over The Rainbow can be considered as an exploratory survey, useful for further more structured and wide surveys, but also as a deepening for all those studies about the LGBTQ population(s) in Italy. This research project has been a long-lasting "work in progress", in which unforeseen problems have been faced in an "extemporaneous" way, and it has sometimes been difficult to deal with them. That is the reason why we suggest working in a team and as a team, creating an interdisciplinary staff which can deal with every kind of all those problems unavoidably linked with LGBTQs, or other hidden populations.

4. Longitudinality

Periodicity of the studies should be the main outline area for future studies, because only a panel approach can give a longitudinal perspective on all those investigated phenomena involving LGBTQ population(s). In this way, the evaluation could outline how performing gender and sexual orientation have changed through generations, also and especially in regards to the use of social media, smartphones and apps. So far, the approach used by all the surveys, reviewed in the first part, was completely cross-sectional, which means that they investigate the target population(s) just once, at one specific point in time. The only exception, so far, was the Fundamental Rights Agency (FRA), which conducted two EU-wide surveys, at a distance of seven years (2012; 2019). Comparing the results between the two surveys, the FRA highlighted the little progress recorded in terms of improving the social and civil conditions of the European LGBTQ people, although with relevant differences between the Member States.

A longitudinal study, would enable the time comparison and the program evaluation, in order to systematically analyse projects, programs, and policies, particularly about their effectiveness and efficiency. Furthermore, from a theoretical point of view, adopting a longitudinal approach, it would be interesting to understand if, and how, the social acceleration, introduced by Rosa (2010), is a phenomenon destined to last, or will it slow down after a starting phase, focusing the analysis on LGBTQ people. Such an analysis would enable to evaluate the interconnection between gender studies and digitalization, the two theoretical frameworks of this project that we reasonably hypothesised more and

more interconnected, making gender something that everyone does, and says, in its own way, mainly online (Cozza, 2008).

5. Snowballing

Regardless of estimates, which are losing their appeal worldwide, as already said, the LGBTQ population(s) is often defined as "hard to reach" (Hughes et al., 2020) for three main reasons:

- its small proportion on the whole population makes it hard to generate a sampling probability;
- LGBTQ people could be reluctant to participate in survey research; as it happens to many other stigmatized or marginalized groups, such as ethnic and religious minorities or subcultures
- sexual orientation and gender identity are extremely hard to classify and then quantify.

In these cases, statistics employ non-probability samples, because statistical sampling techniques cannot be used: there is no frame of the target population to select a sample from, as it is not possible to calculate the inclusion probability of every sampling unit, and hence the sampling weights. So we cannot extend the sampling results to the whole population in a statistical way, but in methodological literature, several purposive sampling methods are presented as the solution to solve this statistical problem.

Snowballing, is a type of purposive sample, useful for those researchers who are trying to recruit people who are difficult to identify or, rare and elusive, as we defined LGBTQ population(s). Snowballing can be used to ease data collection: find one person who qualifies to participate, ask him or her to recommend several other people who have the traits we are looking for. From there the participant list can grow (Abdul Quader et al., 2006).

Snowballing is of course the right methodological and statistical tool to employ for reaching the LGBTQ population(s) and that is the reason why we also used it in our survey. Anyway, it can be further enhanced, upgrading to the Respondent-Driven Sampling (RDS). This sampling method entails multiple snowballing waves, repeating these waves until the condition of Markov equilibrium is verified (Heckarthorn, 2011; De Rosa and Inglese, 2018). In this way, the

probability of being included in the final wave is independent of the probability of being recruited by the first one, which gets "watered down" as more waves are added. The main advantage of this method is that it asymptotically approximates a simple random sample, but, on the other hand, it requires tracking recruiting IDs of responding people, in contrast with the essential privacy and anonymity constraints.
RDS could effectively be the most advisable methodology for studying such a hidden population, but it is not trivial, and its use needs to be handled with specific and complex issues, such as Markov chains and iterative algorithms, which only an interdisciplinary research group can deal with.

6. Using social media and apps in social research

In this research project the lack of a sampling frame, from which to select people to submit to the questionnaire, was circumvented by listing people through a web-scraping program. This technique allowed us to get into the social media black box, but also to think about what working with (or on...) big data involves in the social sciences field. We have analyzed the different positions of the epistemological debate about the use of big data in social research, and we have proposed an alternative option: working on big data, and not with big data. In this way, we can treasure the data richness guaranteed by the tool power, but without allowing any complex algorithms to address the results and the knowledge about a studied phenomenon.
Relying on big data coming from social media entail pros and cons, which have to be put up on a scale:

- the disadvantage is, of course, working with an API, an interface which receives requests and sends responses, allowing users to interrogate the social media server. This API can change its policies (it happened, and it will happen again), it mediates supply and demand, often in a sibylline way, and finally, it never gives all the data, but just a portion of it selected by the API itself in an unspecified way. Using social media as a data collection tool can also mean receiving temporary or permanent warnings because algorithms often cannot find a difference between contact for research and a spam message.

But, if we trust our data, and their massive amounts then we can profit of remarkable advantages as well:

- we can obtain a list of people very quickly, likely belonging to the targeted population, to interview and who can be the initial "snowflakes" of snowballs, in an RDS context. Furthermore, like in our case, people selected tend to be very young, and often not associated to any LGBTQ organization. So, working through social media entailed us to involve in the survey people that would normally be invisible, and it gave the added value of enlarging the initial sample to other hard-to-reach groups of the population, making results more robust, unbiased and generalizable.

Interviews first, and then surveys, pointed out the widespread use of dating apps, although with notable differences among genders and sexual orientations: many respondents use these tools in a very "disenchanted" way, because, after a first enthusiastic phase, problems of self-construction in a computer-mediated environment often occur (Mowlabocus, 2010).
Anyway, trying to use dating apps, intended as full-fledged social spaces, could be a new frontier in research about LGBTQs. Monaco's study (2018) was pioneering, and one might think to enlarge these studies, for instance using more geolocalization spots, all over the national territory for carrying out wide mixed-method research, not circumscribed to just one local area.

7. Associations & Community

What would be the role of associations if the future of LGBTQ research passed through digital technologies?
In the previous chapters, we have often highlighted the indispensable work carried out by associations to raise awareness about LGBTQ, beyond cliché and stereotypes. At the same time, we have never denied that research should emancipate itself by the support and partnership of associations, through more efficient methodologies which guarantees impartiality and robustness to results. This does not mean that associations have to be excluded by the survey design process, indeed. Activists know better than anyone the issues, problems and people's instances so they will always be important in the design phase, mainly for drafting questionnaires. But, on the other hand, social research should stop

relying on associations to recruit possible interviewees, as often happens. The risk of having a biased sample is too high and data confirmed that many people do not participate in associations' activities. Moreover, we have seen that socialization is at a turning point, also and above all for the reasons linked to digital technologies, and it makes conducting research misleading, like ethnography, in LGBTQ spots such as bars or soirées. Research should focus on LGBTQ people and not on the community, a subtle but substantial difference (Formby, 2017), which will guarantee that every group of the LGBTQ population(s) are taken into consideration, avoiding hyper-marginalization phenomena, also in the research context.

8. Mixed (& sequential) Methods Research

From a methodological point of view, we said that an important contribution to the LGBTQ field can be found in the application of mixed methods, mainly employed in Italy, with satisfactory results. This strategy uses and combines quantitative and qualitative approaches, and it seems to be more appropriate and flexible in dealing with the complexity characterizing the LGBTQ population(s).

For designing the present research project, a pragmatic approach that sequentially combines qualitative methods (semi-structured interviews) and a quantitative survey has been used. In this way, the depth of the information collected during interviews becomes the starting point for an efficient, functional and informative questionnaire, consequently answered by hundreds of people. If data has shown that dichotomies such as man-woman, male-female, homosexual-heterosexual, cisgender-transgender, etc. do not work and do not describe accurately the reality between two poles, from a methodological point of view, this research project tells us that dichotomy quantitative-qualitative, analogously, does not work. A questionnaire superficially drafted would not guarantee the information collected from an interview, but, at the same time, it would not be possible to interview hundreds of people so quickly, cheaply and through an engineered process. Social research in this field has to synergize the strengths of qualitative and quantitative methods, avoiding their weaknesses, in favour of the informative richness and deterring the non-responses.

9. Rethinking discrimination studies

The majority of the studies, both qualitative and quantitative, about LGBTQs concern discriminatory episodes experienced by interviewed people. This information, as well as highlighting the difficulties that an LGBTQ person suffers, should address the social policies of institutions so that they can directly contrast this kind of negative behaviour.

However, often discrimination is considered just as a set of extremely violent behaviours. The reality described by the survey points out that discriminatory behaviours can be very sneaky, both in and out of the household, both online and offline. We have also seen, specifically in the sporting context, that discriminatory events may not even occur, if a certain person accurately avoids coming out, due to fear of retaliation, but this could be considered as a form of discrimination as well.

A resultant suggestion is to rethink discrimination, and consequently the way to investigate it. A functional questionnaire has to ask not only about mockery, beating and screams but also about the silence, the indifference and being avoided. Investigating such sheer effects is obviously extremely challenging, mainly with quantitative methods, but the knowledge of discriminatory behaviours (mainly in scholastic, university and working contexts) cannot, and must not overlook the study of all the causes of social discomfort that an LGBTQ person can feel. The side effect of such an oversight would be the structural underestimation of all those negative behaviours towards LGBTQ people, which inevitably could lead to disinterest, inaccuracy and carelessness on LGBTQ civil rights claims.

Concluding, this project does not want to distance itself from the previous LGBTQ researches just from a methodological point of view: surveying LGBTQ people just with questions about discrimination could be considered as an heteronormativity side effect too, which considers the sexual and gender discrimination as obvious, when actually could be not. Many issues regarding LGBTQ people daily life, as important as discrimination, have been uninvestigated for years and social research should change tack, updating methods, data sources and the phenomena investigated, to fully understand the issues.

In this research project, the methodology became as important as the studied field and the attention paid to conduct, interviews before, and a survey after on a large number of individuals led to an acquaintance with the characteristics of the LGBTQ population(s), which are often difficult to define and to detect.

Leveraging on the methodological and empirical experience acquired, we chose to conclude this dissertation focusing on the most important issues faced throughout the survey's implementation. These nine issues above are intended to be guidelines for those who will further investigate the LGBTQ world, so multifaceted and often unexplored.

Social research has to provide policymakers with updated, accurate, accessible, and comparable data, which respects respondents' privacy, and, finally, which accurately and exhaustively describes the studied phenomena.

In this context, big data, coming from social media, can be a valid tool for studying the LGBTQ population(s). In fact, data characteristics, and the way they have been used, guarantee volume, in terms of sample size, the accuracy of the information and the privacy of respondents.

This social media data feature allows us to adopt a data-driven approach, which could encourage efficient and inclusive social policies, whose purpose is the respect of the civil rights of every person, independently of their sexual orientation and gender identity.

REFERENCES

Abdul Quader, Abu S., Douglas D. Heckathorn, Courtney McKnight, Heidi Bramson, Chris Nemeth, Keith Sabin, Kathleen, Gallagher, and Don C., Des Jarlais. 2006. *Effectiveness of Respondent-Driven Sampling for Recruiting Drug Users in New York City: Findings from a Pilot Study.* Journal of Urban Health?: Bulletin of the New York Academy of Medicine. 83 (3). 459-476.

Agger, Ben. 2015. *Oversharing: Presentations of self in the internet age*. New York and London: Routledge.

Ahmed, Sara. 2006. *Queer phenomenology: Orientations, objects, others.* Durham: Duke University Press.

Albury, K., et al. 2019. *Safety, risk and wellbeing on dating apps: Final report.* Retrieved 5th January, 2020 (https://apo.org.au/node/268156)

Allegro, Silvia. 2006. *Le sfide della maternità lesbica: uno studio sui percorsi di coming out madre-figli.* Omosapiens. Studi e ricerche sugli orientamenti sessuali. 43-54. Roma: Carocci.

Amaturo, Enrica, and Biagio Aragona. 2019. *Per un'epistemologia del digitale: note sull'uso di big data e computazione nella ricerca sociale.* Quaderni di Sociologia 81.81-LXIII: 71-90.

Amodeo, Anna Lisa, Nicola D. Casolare, and Simona Picariello. 2017. *La percezione dell'omofobia nello sport: uno studio esplorativo.* Mimesis.

Anderson, Chris. 2008. *The end of theory: The data deluge makes the scientific method obsolete.* Wired magazine. 44028.

Antonelli, Francesco, and Elisabetta Ruspini. 2016. *Il futuro della ricerca su generi e orientamenti sessuali. Sociologia del futuro Studiare la società del ventunesimo secolo*. 247-275 . Fabio, Corbisiero and Ruspini, Elisabetta. Padova: CEDAM.

Arcigay. 2005. MODIDI - Sesso *e salute di lesbiche gay e bisessuali oggi in Italia*. Summary document of main sociologic and health results. Retrieved October 1st, 2018 (https://www.arcigay.it/wp-content/uploads/modidi_opuscolo.pdf)

Badgett, M. V. Lee. 2009. *Best Practices for Asking Questions about Sexual Orientation on Surveys.* The William Institute.

Barad, Karen. 2003. *Posthumanist performativity: Toward an understanding of how matter comes to matter.* Journal of women in culture and society. 28 (3). 801-83.

Barad, Karen. 2007. *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning.* Duke University Press.

Barbagli, Marzio and Asher Colombo. 2007. *Omosessuali moderni. Gay e lesbiche in Italia.* Bologna: Il Mulino.

Benadusi, Lorenzo. 2007. *La storia dell'omosessualità maschile: linee di tendenza, spunti di riflessione e prospettive di ricerca*. Rivista di sessuologia. 31(1). 42005.

Benadusi, Lorenzo, Simonetta Piccone Stella, and Assunta Viteritti . 2009. *Dispari parità: genere tra educazione e lavoro.* Guerini e Associati.

Bernini, Lorenzo, Annalisa Zabonati and Massimo Prearo. 2015. *Chi ha paura del genere? Strumenti per una corretta informazione.* Retrieved August 07th, 2020 (http://www.cesp-pd.it/spip/IMG/pdf/-29.pdf)

Blumenfeld, Warren J., and R. M. Cooper. 2010. *LGBT and allied youth responses to cyberbullying: Policy implications.* The International Journal of Critical Pedagogy. 3.1. 112.

Bocchiaro, Piero, and Stefano Boca. 2002. *Differenze di genere nelle professioni tecnico-scientifiche: un'analisi sul ruolo dell'autoefficacia e della minaccia indotta dallo stereotipo. Giornale italiano di psicologia.* 29.3. 491-514.

Boyd, Danah Michele. 2008. *Taken out of context: American teen sociality in networked publics.* University of California, Berkeley.

Butler, Judith. 1990. *Gender trouble and the subversion of identity*. New York and London: Routledge.

Butler, Judith. 2015. *Senses of subject.* New York: Fordham University Press.

Cafasso Simone. 2017. Quanti *sono i gay in Italia? I conti non tornano*. Pagina99. Retrieved August 07th, 2017 (http://www.pagina99.it/2017/03/23/quanti-sono-i-gay-in-italia-i-conti-che-non-tornano/)

Campbell, Donald T., and Donald W. Fiske. 1959. *Convergent and discriminant validation by the multitrait-multimethod matrix.* Psychological bulletin. 56 (2). 81-99.

Card, Kiffer G. 2016. *The role of sex-seeking apps and websites in the social and sexual lives of gay, bisexual and other men who have sex with men: a cross-sectional study*. Sexual health. 14 (3). 229-237.

Carey-Mahoney, Ryan. 2016. *LGBT-who? Decoding the ever-changing acronym*. The Washington Post. Retrieved October 1st, 2018 (https://www.washingtonpost.com/news/soloish/wp/2016/06/10/lgbt-who-decoding-the-ever-changing-acronym/?noredirect=on&utm_term=.f1f1f598f73b).

Carnelli, Fabio, Romana De Micheli, and Maria Carolina Vesce. 2009. *Invisibilità manifeste*.

Castells, Manuel. 2002. *La nascita della società in rete.* Milano: Università Bocconi.

Ceron, Andrea, Luigi Curini, and Stefano M. Iacus. 2013. *Social Media e Sentiment Analysis: L'evoluzione dei fenomeni sociali attraverso la Rete.* vol.9. Milano: Springer Science & Business Media.

Chhetri, Gyanu. 2017. *Perceptions About The "Third Gender" In Nepal*. Dhaulagiri Journal of Sociology and Anthropology. 196-114.

Chung, Chia-Fang, Elena Agapie, Jessica Schroeder, Sonali Mishra, James Fogarty, and Sean A. Munson. 2017. *Personal Tracking Becomes Social: Examining the Use of Instagram for Healthy Eating. Proceedings of the SIGCHI conference on human factors in computing systems.* CHI 2017 Conference. Retrieved October 1st, 2018 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5432132/)

Coffman, Katherine B., Lucas C. Coffman, and Keith M. Marzilli Ericson. 2017. *The Size of the LGBT Population and the Magnitude of Anti-Gay Sentiment are Substantially Underestimated.* Management Science. 63.10. 3168-3186. Retrieved September 17th, 2017 (https://www.nber.org/papers/w19508)

Coletta, Claudio, and Rob Kitchin. 2017. *Algorithmic governance: Regulating the 'heartbeat of a city using the Internet of Things.* Big Data & Society. 4(2).

Collard, James. 1998. *Leaving the gay ghetto*. Newsweek. 17 August. 53.

Connell, Raewyn W. 2005. *Masculinities*. Berkeley University Press.

Cooper, Margaret, and Kristina Dzara. 2010. *The Facebook revolution: LGBT identity and activism. LGBT identity and online new media.* 114-126. New York and London: Routledge.

Corbisiero, Fabio, and Salvatore Monaco. 2017. *Città arcobaleno: Una mappa della vita omosessuale nell'Italia di oggi*. Donzelli ed.

Corbisiero, Fabio, Ed. 2013. *Comunità omosessuali. Le scienze sociali sulla popolazione LGBT.* Milano: FrancoAngeli.

Corbisiero, Fabio, ed. 2010. *Certe cose si fanno: identità, genere e sessualità nella popolazione LGBT.* Napoli: Gesco Ed.

Cozza, Michela. 2008. *Fare e disfare il genere. Studiare la tecnologia in un'ottica di genere*. II Convegno nazionale STS Italia: Catturare Proteo. Tecnoscienza e società della conoscenza in Europa, Università di Genova (2008): 19-21.

Creswell, John W., and Vicki L. Plano Clark. 2011. *Designing and conducting mixed methods research.* Los Angeles: Sage.

Cuccurullo, Alessia, and Nicola D. Casolare. 2013. *Omofobia e sport*. Retrieved October 2020, (http://www.sinapsi.unina.it/bullismoomofobico_omofobia_sport)

Daas, Piet, and Marco Puts. 2014. *Social Media Sentiment and Consumer Confidence.* Statistics Paper Series. 5. 46023. Retrieved February 14th, 2018 (https://www.ecb.europa.eu/pub/pdf/scpsps/ecbsp5.en.pdf)

Dara, Virginia. 2017. *Community LGTBQ sui social: iniziative, coinvolgimento e polemiche. Inside Marketing*. Retrieved April 14th, 2020 (https://www.insidemarketing.it/community-lgtbq-sui-social-iniziative/)

De Rosa, Eugenia, and Francesca Inglese. 2018. *Diseguaglianze e discriminazioni nei confronti delle persone LGBT: quale contributo dalla statistica ufficiale?*. Rivista Italiana di Economia Demografia e Statistica . 72 (4). 77-88.

De Rose, Alessandra, and Catherine Marquette. 2011. *Same-sex families in Italy compared to Spain.* Journal of Family Research. 53-69.

De Vita, Luisa, and Orazio Giancola. 2017. *Between Education and Employment: Women's Trajectories in STEM Fields.* Polis. 31(1). 45-72.

Denny, Matt. 2017. *A bitesize intro to... Web Scraping.* Retrieved April 24th, 2020 (https://campus.sagepub.com/blog/a-bitesize-intro-to-web-scraping)

D' Ippoliti, Carlo, and Alexander Schuster. 2011. *DisOrientamenti. Disciminazione ed esclusione sociale delle persone LGBT in Italia.* Vol. 4. Roma: Armando Editore.

Duffy, B., et al. 2018. *Beyond Binary. The lives and choices of Generation Z.* Retrieved April 29th, 2020 (https://www.ipsos.com/sites/default/files/2018-08/ipsos_-_beyond_binary_-_the_lives_and_choices_of_gen_z.pdf)

Ellis, Renee et al. 2017. *Assessing the Feasibility of Asking Sexual Orientation and Gender Identity in the Current Population Survey: Results from Cognitive Interviews*. Washington, DC: Center for Survey Measurement, US Census Bureau and Office of Survey Methods Research, Bureau of Labor Statistics.

European Union. 2016. *List of actions by the Commission to advance LGBTI equality*. Luxembourg: Publications Office of the European Union.

Fabris, Giampaolo, and Rowena Davis. 1978. Il *mito del sesso rapporto sul comportamento sessuale degli italiani*. Mondadori.

FADA. 2017. *Survey of attitudes towards lesbians, gay and bisexual people.* retrieved 1st April, 2020 (https://www.antidiskriminierungsstelle.de/SharedDocs/Pressemitteilungen/EN/2017/20170112_Umfrage_TJ_2017.html?cms_gtp=12797538_Dokumente%253D3)

Fields, Jessica. 2001. *Normal queers: Straight parents respond to their children's "coming out"*. Symbolic interaction. 24.2. 165-167.

Fiore, Crescenzo, ed. . 1991. *Sorriso di Afrodite. Rapporto sulla condizione omosessuale in Italia*. Firenze: Vallecchi.

Formby, Eleanor. 2017. Exploring *LGBT spaces and communities: Contrasting identities, belongings and wellbeing*. Taylor & Francis.

FRA. 2012. *EU LGBT survey Technical report Methodology, online survey, questionnaire and sample.* Retrieved August 18th, 2018, (https://fra.europa.eu/sites/default/files/eu-lgbt-survey-technical-report_en.pdf)

FRA. 2020. *EU LGBT survey II - A long way to go for LGBTI equality*. Retrieved May 14th, 2020, (https://fra.europa.eu/en/publication/2020/eu-lgbti-survey-results)

Franchi, Marina, and Giulia Selmi. 2020. *Same-Sex Parents Negotiating the Law in Italy: Between Claims of Recognition and Practices of Exclusion*. Same-Sex Families and Legal Recognition in Europe. 73-93. Springer.

Franchi, Marina, and Giulia Selmi. 2018. *Challenging the unthinkable: gay and lesbian parents between redefinition and exclusion in Italy.* About Gender. International Journal of Gender Studies. 7(14). 44197.

Fryrear, Andrea. 2016. *How to write gender questions for a survey.* Retrieved March 18th, 2020, (https://www.surveygizmo.com/resources/blog/how-to-write-survey-gender-questions/)

Gable, Guy G. 1994. *Integrating case study and survey research methods: an example in information systems*. European journal of information systems. 3 (2). 112-126.

Gasperoni, Giancarlo. 1996. *Rendimento Scolastico e Istruzione Secondaria Superiore*. Bologna: Il Mulino

Gates, Gary J. 2011. *How many people are LGBT?* Los Angeles, California, USA: University of California: The Williams Institute.

Gates, Gary J. 2017. *LGBT data collection amid social and demographic shifts of the US LGBT community*. 1220-1222.

Gates, Gary J., and Frank Newport. 2012. *Special report: 3.4% of US adults identify as LGBT*. Washington, DC: Gallup.

Ghaziani, Amin. 2011. *Post-gay collective identity construction.* Social Problems. 58. 99-125.

Giddens, Anthony. 1992. *The transformation of intimacy: Sexuality, love and eroticism in modern societies.* John Wiley & Sons.

Gluckman, Peter. 2017. *How big data can contribute to more efficiency and impact in social policies.* Medium. Retrieved 4th February 2020, (https://medium.com/reaffirming-social-values-in-uncertain-times/how-big-data-can-contribute-to-more-efficiency-and-impact-in-social-policies-1cfee2d0728a)

Gobo, Giampietro. 2015. *La nuova survey: sondaggio discorsivo e approccio internazionale.* Roma: Carocci.

González-Bailón, Sandra. 2013. *Social science in the era of big data*. Policy & Internet. 5 (2). 147-160.

Government Equalities Office. 2018. *National LGBT survey. summary report*. Retrieved 16th February 2020, (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/722314/GEO-LGBT-Survey-Report.pdf)

Graglia, Margherita. 2012. *Omofobia: strumenti di analisi e di intervento*. Roma: Carocci.

Green, Melissa S. 2012. *Anchorage LGBT discrimination survey.* Retrieved April 19th, 2019, (https://alaskacommunity.org/news/anchorage-survey/)

Grov, Christian. 2014. *Gay and bisexual men's use of the Internet: research from the 1990s through 2013*. The Journal of Sex Research. 51 (4). 390-406.

Hanft, Ruth S. 1981. *Use of social science data for policy analysis and policymaking*. The Milbank Memorial Fund Quarterly. Health and Society. 59 (4). 596–613.

Haraway, Donna J. 1991. *Simians, cyborgs and women: The reinvention of nature*. New York: Routledge.

Haraway, Donna J. 1997. *MODEST WITNESS@ SECOND MILLENNIUM. FEMALEMAN MEETS ONCOMOUSE: Feminism and Technoscience*. New York: Routledge.

Heckathorn, Douglas D. 1997. *Respondent-driven sampling: a new approach to the study of hidden populations. Social Problems*. 44.2. 174-199.

Hughes, Mark. 2006. *Queer ageing. Gay and lesbian issues and psychology review*. 2 (2). 54-7.

IFOP. 2017. *To bi or not to bi ? Enquête sur l'attirance sexuelle entre femmes*. Retrieved 2nd April, 2020 (https://www.ifop.com/publication/to-bi-or-not-to-bi-enquete-sur-lattirance-sexuelle-entre-femmes/)

ILGA. 2018. *Annual report*. Retrieved 3rd March, 2020 (https://ilga.org/downloads/ILGA_Annual_Report_2018_eng_web.pdf)

Inghilleri, Marco, and Elisabetta Ruspini. 2011. *Sessualità narrate. Esperienze di intimità a confronto: Esperienze di intimità a confronto*. Milano: Franco Angeli.

Irlam, C. B. 2012. *LGBTI Data: developing an evidence-informed environment for LGBTI health policy*. Sydney: National LGBTI Health Alliance.

Istat. 2011. *La popolazione omosessuale nella società italiana*. Statistiche report. Retrieved August 13th, 2018, (https://www.istat.it/it/archivio/62168)

Istat. 2018. *Permanent census of the Population and Housing*. Retrieved August 26th, 2019, (https://www.istat.it/en/permanent-censuses/population-and-housing)

Jenzen, Olu, and Irmi Karl. 2014. *Make, share, care: Social media and LGBTQ youth engagement*. Journal of Gender, New Media, and Technology. 5.

Jick, Todd D. 1979. *Mixing qualitative and quantitative methods: Triangulation in action. Administrative science quarterly.* 24.4. 602-61.

Johnson, Patrick M. 2020. *Coming Out Queer Online: Identity, Affect, and the Digital Closet*. Lexington Books.

Johnson, R. Burke, Anthony J. Onwuegbuzie, and Lisa A. Turner. 2016. *Toward a definition of mixed methods research*. Journal of mixed methods research. 1 (2). 112-133.

Kinsey, Alfred C., Wardell B. Pomeroy, and Clyde E. Martin. 1948. *Sexual behavior in the human male*. The Journal of Nervous and Mental Disease. 109.3. 283.

Kinsey, Alfred C., Wardell B. Pomeroy, and Clyde E. Martin. 1953. *Sexual behavior in the human female*. Indiana University Press.

Kitchin, Rob. 2014. *Big Data, new epistemologies and paradigm shifts*. Big data & society. 1(1).

Klein, Fritz. 1993. *The bisexual option*. New York: The Haworth Press.

Knight, Kyle. 2014. *Surveying Nepal's Sexual and Gender Minorities: An Inclusive Approach*.

Knight, Kyle. 2011. *What we can learn from Nepal's inclusion of 'third gender' on its 2011 census.* The New Republic.

Lam, Athena. 2016. *Counting the LGBT population: 6% of Europeans identify as LGBT*. Dalia Research. Retrieved 3rd April, 2020 (https://daliaresearch.com/blog/counting-the-lgbt-population-6-of-europeans-identify-as-lgbt/)

Laney, Douglas. 2003. *3D data management: Controlling data volume, velocity and variety*. META Group research note . 6(70).

Lelleri, Raffaele (Ed.). 2006. *Survey nazionale su stato di salute, comportamenti protettivi e percezione del rischio HIV nella popolazione omo-bisessuale*. Report finale. Bologna: Arcigay.

Lombi, Linda. 2015. *La ricerca sociale al tempo dei Big Data: sfide e prospettive*. Studi di sociologia. 215-227.

Lupton, Daniel. 2013. *Introducing digital sociology*. Sydney: University of Sydney.

Lupton, Daniel. 2015. *Digital sociology*. New York and London: Routledge.

Lykke, Nina. 2011. *Intersectional analysis: Black box or useful critical feminist thinking technology*. 207-22.

Lyon, David. 2014. *Surveillance, Snowden, and big data: Capacities, consequences, critique*. Big data & society 1(2).

Macintosh, Lori, and Mary Bryson. 2007. *Queer (re) presentations*. Journal of LGBT Youth. 5(1).

MacLeod, Caitlin and Victoria McArthur. 2018. *The construction of gender in dating apps: an interface analysis of Tinder and Bumble*. Feminist Media Studies. 43466.

Mäkitalo, Åsa, Todd E. Nicewonger, and Mark Elam, Eds. 2020. *Designs for experimentation and inquiry. Approaching Learning and Knowing in Digital Transformation*. London and New York: Routledge, Taylor & Francis Group.

Marcus, Sharon. 2005. *Queer Theory for Everyone: A Review Essay*. Journal of Women in Culture and Society. 31(1). 191-218.

Marres, Noortje. 2012. *The redistribution of methods: on intervention in digital social research, broadly conceived.* The sociological review. 60. 139-165.

Matsick, Jes L., and Jennifer D. Rubin. 2018. *Bisexual prejudice among lesbian and gay people: Examining the roles of gender and perceived sexual orientation*. Psychology of Sexual Orientation and Gender Diversity. 5.2. 143-155.

Matthews, John, and Elizabeth P. Cramer. 2008. *Using technology to enhance qualitative research with hidden populations.* Qualitative Report. 13(2). 301-315.

McDonnell, Diarmuis. 2020. *Web-scraping for Social Science Research: Websites as a Source of Data*. Retrieved April 23th, 2020 (https://ukdataservice.ac.uk/media/622652/web_webscrapingwebsites23apr20.pdf)

McMillan, David W., and David M. Chavis. 1986. *Sense of community: A definition and theory*. Journal of community psychology. 14.16-23.

Michael, Robert T., et al. 1994. *Sex in America: A definitive survey*.

Mieli, Mario. 1977. *Elementi di critica omosessuale*. Milano: Feltrinelli Editore.

Monaco, Salvatore. 2019. *Mixed methods e e-research: frontiere possibili per lo studio delle hidden population*. Sociologia Italiana. 14. 95-105.

Mowlabocus, Sharif. 2010. *Gaydar Culture: Gay men, technology and embodiment in the Digital Age.* Routledge.

Munk, Kristian, Mette Simonsen Abildgaard, Andreas Birkbak, and Morten Krogh Petersen. 2016. *Re-Appropriating Instagram for Social Research: Three Methods for Studying Obesogenic Environments.* Retrieved August 27th, 2018, (https://www.researchgate.net/publication/304918866_Re-Appropriating_Instagram_for_Social_Research_Three_Methods_for_Studying_Obesogenic_Environments)

Nagarajan, Rema. 2014. *First count of third gender in census: 4.9 lakh*. The Times of India.

Nelson, Kenneth. 2011. *Improving the efficiency of social protection: Synthesis Report*.

Norton, Rictor. 2016. *Myth of the modern homosexual: Queer history and the search for cultural unity*. Bloomsbury Publishing.

OECD. 2020. *Social policies and data*. Retrieved February 20th 2020, (http://www.oecd.org/els/soc/)

Omnicore. 2020. *Instagram by the Numbers: Stats, Demographics & Fun Facts*. Retrieved April 20th 2020, (https://www.omnicoreagency.com/instagram-statistics/)

Papuli, Celeste. 2019. *Io volevo andare nella foresta. Storie di vita per una sociologia dell'esperienza trans*. Mimesis / Eterotopie.

Pascoe, Cheri J. 2011. *Resource and risk: Youth sexuality and new media use*. Sexuality Research and Social Policy. 8(1). 5-17.

Paterlini, Piergiorgio. 2006. *Matrimoni gay. Dieci storie di famiglie omosessuali*. Torino: Einaudi.

Pew Research Center. 2020. *The Virtues and Downsides of Online Dating*. Retrieved August 19th, 2020, (https://www.pewresearch.org/internet/2020/02/06/the-virtues-and-downsides-of-online-dating/)

Pew Research Center. 2013. *A survey of LGBT Americans: attitudes, experiences and values in changing times.* Retrieved August 19th, 2018, (https://www.pewsocialtrends.org/2013/06/13/a-survey-of-lgbt-americans/)

Pinch, Trevor J. and Wiebe E. Bijker. 1984. *The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other*. Social Studies of Science. 14(3). 399–44.

Poggio, Barbara, and Giulia Selmi. 2012. *Sfidare i confini del genere*. AG About Gender-Rivista internazionale di studi di genere . 1(2). 1-8.

Porrovecchio, Alessandro. 2012. *Sessualità in divenire: adolescenti, corpo e immaginario.* Milano: Franco Angeli.

Prensky, Marc. 2009. *Homo sapiens digital: From digital immigrants and digital natives to digital wisdom.* Innovate: Journal of online education. 5(3).

Pullen, Christopher, and Margaret Cooper, eds. 2010. *LGBT identity and online new media*. Routledge.

Randall, Sell. 2020. *LGBTDATA.COM.* Retrieved March 1st, 2020 (http://www.lgbtdata.com/data.html)

Redattore Sociale. 2013. *Parlare civile. Comunicare senza discriminare*. Milano: Mondadori Editore.

Ricciardelli, Rosemary, Kimberley A. Clow, and Philip White. 2010. *Investigating hegemonic masculinity: Portrayals of masculinity in men's lifestyle magazines*. Sex Roles. 63.1-2. 64-78.

Richardson, Diane, and Steven Seidman. 2002. *Handbook of Lesbian and Gay Studies*. SAGE.

Rinaldi, Cristina, and C. Cappotto. 2014. *Figli di genitori "normali". Omosessualità e dinamiche familiari tra tradizione, resistenza e sperimentazione*. 96-118.

Rizzo, Amelia. 2018. *Management dell' identità sui Social Network*. PSYCOMEDIA Telematic Review. Retrieved October 15th , 2018, (http://www.psychomedia.it/pm/telecomm/massmedia/rizzo.htm)

Romano, Monica. 2011. *Diversity & Inclusion: la condizione transgender nel mondo del lavoro*.

Rosa, Hartmut. 2010. *High-speed society: Social acceleration, power, and modernity*. Penn State Press.

Ross, Micheal W., Ronny Tikkanen, and Rigmor C. Berg. 2014. *Gay community involvement: its interrelationships and associations with Internet use and HIV risk behaviors in Swedish men who have sex with men. Journal of homosexuality*. 61 (2). 323-333.

Rossman, Gretchen B., and Bruce L. Wilson. 2013. *Reassembling social science methods: the challenge of digital devices*. Theory, Culture & Society. 30 (4). 22-46.

Ruspini, Elisabetta. 2014. *Le differenze di genere. Sociologia. Problemi, teorie e intrecci storici*. Roma: Carocci.

Russell, Stephen T., et al. 2010. *Safe schools policy for LGBTQ students and commentaries*. Social Policy Report. 24.4. 1-25.

Saewyc, E., et al. 2007. *Not yet equal: The health of lesbian, gay, & bisexual youth in BC*. Vancouver, BC: McCreary Centre Society.

Saraceno, Chiara, Alessandro Casiccia, Chiara Bertone and Paola Torrioni. 2003. *Diversi da chi? Gay, lesbiche e transessuali in un'area metropolitana*. Milano: Guerini e Associati.

Savage, Mike, and Roger Burrows. 2007. *The coming crisis of empirical sociology.* Sociology. 41.5. 885-899.

Schröder, Jonas. 2018. *Social Network Analysis of Related Hashtags on Instagram (using InstaCrawlR).* Retrieved September 10th, 2018 (https://towardsdatascience.com/social-network-analysis-of-related-hashtags-on-instagram-using-instacrawlr-46c397cb3dbe)

Simon Rosser, Brandon R., William, and Richard Weinmeyer. 2008. *Are gay communities dying or just in transition? Results from an international consultation examining possible structural change in gay communities.* AIDS Care. 20 (5). 588-595.

Singh, Anneliese A., and Kimber Shelton. 2011. *A Content Analysis of LGBTQ Qualitative Research in Counseling: A Ten-Year Review.* Journal of Counseling & Development. 89(2). 217-226.

Snee, Helene, Christine Hine, Yvette Morey, Steven Roberts, and Hayley Watson. 2016. *Digital Methods for Social Science - An Interdisciplinary Guide to Research Innovation.* New York and London: Palgrave Macmillan ed.

Spitale, Samuele C. 2015. *Gay Ghosting Is an Epidemic*. Advocate. Retrieved 15th April, 2019 (https://www.advocate.com/commentary/2015/10/13/gay-ghosting-epidemic)

Sumter, Sindy R., and Laura Vandenbosch. 2019. *Dating gone mobile: Demographics and personality-based correlates of using smartphone-based dating applications among emerging adults*. New media & society. 21.3. 655-673.

Trappolin, Luca. 2008. *Omosapiens 3, Per una sociologia dell'omosessualità*. Roma: Carocci.

UN. 2017. *The price of exclusion*. Retrieved September 2nd, 2018, (https://www.unfe.org/the-price-of-exclusion/)

UNAR-Istat. 2020. *Il diversity management per le diversità lgbt+ e le azioni per rendere gli ambienti di lavoro più inclusivi.* Retrieved November 13rd, 2020 (http://www.unar.it/wp-content/uploads/2020/11/report-Diversity-UNAR-ISTAT.pdf)

UNECE. 2019. *Measuring gender identity*. Unece working paper. 5.

Van Gelder, Linda, Dominik Koehler, and Nicholas Menzies. 2018. *Life on the Margins: Survey Results of the Experiences of LGBTI People in Southeastern Europe.* Retrieved April 5th, 2020 (https://blogs.worldbank.org/europeandcentralasia/life-margins-experiences-lgbti-people-southeastern-europe)

Wallace, Frankie. 2019. *The Pros and Cons of Social Media. The gay & lesbian review*. Retrieved April 5th, 2020 (https://glreview.org/the-pros-and-cons-of-social-media/)

Waskul, Dennis D. . 2003. *Self-games and body-play: Personhood in online chat and cybersex*. New York: Peter Lang.

Webb, Eugene J., Daniel T. Campbell, Richard D. Schwartz, and Lee Sechrest. 1999. *Unobtrusive measures*. Vol. 2. Newbury Park: Sage Publications.

Wesling, Meg. 2008. *Why queer diaspora?*. Feminist Review. 90.30-47.

Westbrook, Laurel, and Aliya Saperstein. 2015. *New categories are not enough: Rethinking the measurement of sex and gender in social surveys*. Gender & Society. 29 (4). 534-560.

Whittle, Stephen, Turner Lewis, and Al-Alami Maryam. 2007. *Engendered Penalties: Transgender and Transsexual People's Experiences of Inequality and Discrimination*. Press For Change.

Wiederhold, Brenda K. 2014. *Cyberbullying and LGBTQ youth: A deadly combination*. 569-570.

Wilson, Tom and Shalley, Fiona. 2018. *Estimates of Australia's non-heterosexual population*. Charles Darwin University working paper. 2.

World Bank. 2018. *Life on the Margins: Survey Results of the Experiences of LGBTI People in Southeastern Europe*. Retrieved April 4th, 2020 (http://documents.worldbank.org/curated/en/123651538514203449/pdf/130420-

REPLACEMENT-PUBLIC-FINAL-WEB-Life-on-the-Margins-Survey-Results-of-the-Experiences-of-LGBTI-People-in-Southeastern-Europe.pdf)

Wronski, Laura. 2019. *Why (and how!) to ask survey questions on sexual orientation and gender identity*. Retrieved March 4th, 2019 (https://www.surveymonkey.com/curiosity/ask-survey-questions-sexual-orientation-gender-identity/)

Zanola, Elisa. 2014. *The Sociological Research on LGBT population in Italy*. Italian Sociological Review. 4(3). 383-398.

Zervoulis, Karyofyllis, David S. Smith, Rhiannon Reed, Sokratis Dinos. 2020. *Use of 'gay dating apps' and its relationship with individual well-being and sense of community in men who have sex with men*. Psychology & Sexuality. 11:1-2. 88-102.

Zevallos, Zuleyka. 2014. *Sociology of Gender. The Other Sociologist*. Retrieved July 10th, 2018 (https://othersociologist.com/sociology-of-gender/)

Zhang, Weiwu, Thomas J. Johnson, Trent Seltzer, and Shannon L. Bichard. 2010. *The revolution will be networked: The influence of social networking sites on political attitudes and behavior*. Social Science Computer Review. 28 (1). 75-92.

APPENDIX

Q) Questionnaire dimensions structure and questions

Introductory letter to ask people to participate in the survey, explaining why they have been selected and the aim of the research project.		
Dimension	**Questions**	**Modalities**
Sociodemographic	1) How old are you?	(years old)
	2) In which region were you born?	20 Italian regions (NUTS 2)
	3) What is the population of your hometown size?	• ≤5000 • 5'001-20'000 • 20'000-50'000 • 50'000-250'000 • >250'000 inhabitants
	4) In which region do you live?	20 Italian regions (NUTS 2)
	5) What is the population of your city of residence?	• ≤5000 • 5'001-20'000 • 20'000-50'000 • 50'000-250'000 • >250'000 inhabitants
	6) What is your educational level?	• Elementary education or none • Junior school license • Professional qualification (3 years high school) • High school diploma (5 years) • Bachelor's degree • Master's degree • Post-graduate master • PhD
	7) Are you at the present in a relationship?	• No, single • Yes, with a same-gender person • Yes, with another-gender person • Other:_____
	8) What is your marital status?	• Unmarried • In a civil union • Married • Divorced • Other:_____
	9) Who do you live with?	• On my own, • With my sons/daughters • With my partner

		• With my partner and (my/his/her/our) sons/daughters • With friends/flatmates • With my parents • With other relatives (uncles/aunts, grandparents, etc.)
	10) What do you do?	• I Work (to question 11) • I am looking for a job • I study (to question 13) • I am retired • Other:_____
Work	11) What's your job?	• Businessmen, manager • High specialization job (researcher, engineer, etc.) • Technical job • Executive office job (secretary, accountant, etc.) • Commercial job (seller, shop assistant, etc.) • Artisan, specialized worker, farmer • Driver, factory worker • Non-qualified job • Armed force
	12) What is your net monthly wage?	• 0 – 500 € • 501 - 1000 € • 1000 - 1500 € • 1500 - 2000 € • 2000 - 3000 € • > 3000 €
Education	13) What are you studying for?	• Junior school license • High school diploma (to question 14) • Bachelor's degree (to question 15) • Master's degree (to question 15) • Post-graduate master (to question 15) • PhD (to question 15)
	14) Which kind of diploma are you studying for?	• Professional institute • Technical institute • Teaching institute • Scientific high school • Classical high school • Artistic high school
	15) Which academic area are you studying for?	• Agricultural sciences • Architecture

		• Chemical-Pharmaceutical • Defence and security • Economic-statistical • Sport sciences • Geo-biological • Law • Engineering • Teaching • Literature • Language • Medicine • Social and political sciences • Psychological • Scientific • Other:_____
Sex, gender, sexual orientation	16) Birth sex?	• Male • Female • Intersex
	17) How would you define your gender identity?	• Male • Female • Transgender • Queer, gender-fluid, non-binary • Other:_____
	18) How would you define your sexual orientation?	• Asexual • Bisexual • Heterosexual • Homosexual • Other:_____
	19) Agreement with sentences about gender identity, sex and sexual orientation: 5-levels Likert scale (strongly disagree, disagree, undecided, agree, strongly agree)	• Talking about gender identity, and not of sex, comforts me, because being classified as male or female is not enough • Proliferation of gender identities is getting ridiculous and it increases the distance between LGBTQ and heterosexual communities • My gender identity cannot be labelled. I perform it daily • "Queer" suits my identity • The proliferation of sexual orientations is getting ridiculous and it increases the distance between LGBTQ and heterosexual communities

		• The multitude of gender identities and sexual orientations highlights the LGBTQ community's thousand facets • The multitude of gender identities and sexual orientations makes the LGBTQ civil rights claims stronger • Some sexual orientations are more discriminated than others
Coming-out	20) Who was the first person you came-out to?	• None, I have never come-out (to question 22) • My mother • My father • My sister • My brother • My partner • Friends • Colleagues • Superiors at work (teachers at school) • Team-mates, sport-mates • Relatives (uncles/aunts, grandparents, etc.)
	21) Who knows your sexual orientation today?	• My mother • My father • My sister • My brother • My partner • Friends • Colleagues • Superiors at work (teachers at school) • Team-mates, sport-mates • Relatives (uncles/aunts, grandparents, etc.)
Social Network (SN)	22) How often do you use these SNs? (SNs: Instagram, Facebook, Twitter, Others)	• I don't have any account for this SN • I've got an account but I don't use it anymore • Once a week • Once a month • Once a day • More than once a day
	23) Which device do you use for browsing SNs?	• PC • Tablet • Smartphone

		• Other:_____
	24) Which SN do you use for these purposes? (SNs: Instagram, Facebook, Twitter, Others)	• To post selfies • To post my photo/videos, made by other ones • To post group photos of me and friends/colleagues of mine • To post photo/videos of me and my partner • To post photo/videos of activities that I carry out and which characterize me (sport, music that I listen to, books that I read, pics of my dog/cat) • To post contents that I'd like my friends to repost (pics taken from the web, events, meme, magazine/newspaper articles) • To post 24h-lasting stories • To keep me informed, reading posts • To comment on posts that I agree with (also just liking them) • To comment posts that I don't like, giving my reasons to a comment • Using hashtag to index my post and make it traceable by other users • To follow VIPs and celebrities • To chat with people that I know in real life • To chat with people that I do not know in real life
	25) Agreement with sentences about social media: 5-levels Likert scale (strongly disagree, disagree, undecided, agree, strongly agree)	• SNs are useful to meet new people • SNs are useful to show ourselves • SNs are useful to feel less alone • SNs are dangerous, because we only communicate through a keyboard • SNs make relations distant and superficial • SNs, carefully used, represent a good tool to be informed • SNs take too much time • SNs allow me to share my opinions

		• I use SNs to find out all the details of a person I have just met, looking at what he/she posts • SNs are a double-edged sword: they give freedom of expression to many people, but can be used unhelpfully
	26) Have you ever used this SN for these purposes? (SN: Instagram, Facebook, Twitter, Others)	• To meet people with whom you later hanged out with • To meet people, without hanging out with them • To participate in an LGBTQ friendly event, of which you heard on SNs (Facebook events, Instagram event page, etc.) • To keep informed, reading articles, listening to podcasts or watching videos about LGBTQ claims • To keep informed, reading articles, listening to podcast or watching videos about coming-out • To express solidarity in cases of homophobia events (e.g Orlando massacre) • To express your satisfaction for the recognition of LGBTQ civil rights (e.g Civil Unions for same-sex couples) • To express your annoyance for the lack of LGBTQ civil rights (e.g Civil Unions for same-sex couples)
	27) Is your sexual orientation deducible by your SNs profiles?	• Yes, it is. • There are pictures with my partner/s but I have never come-out on SNs. • No, it is not.
Dating apps (DAs)	28) Have you ever used DAs, such as Grindr, Tinder, Her, Wapa,Happn, etc.?	• Yes, I have. I still use DAs. *(to 30)* • Yes, I have used them, but I don't use DAs anymore. *(to 29)* • No, I have never used DAs. *(to 32)*
	29) Why did you stop using Das?	• Because I got engaged with a person met through a DA • Because chatting and swiping people pictures, takes too much time

		• Because I did not meet very interesting people • Because people I met, in the real-life, were not as interesting as on DAs • Because I prefer to meet people in a more "traditional" way • Because DAs are just another way of avoiding out sexual orientation and I met too many people engaged in a heterosexual relationship • Other:_____
	30) Have you ever hang out with someone known on a DA?	• Yes, I have, more than once • yes, I have, just once • No, I have never hanged out with someone met on a DA
	31) Do/did you have a public profile on the Das you used?	• Yes, I used my profile picture • No, I used a fake/anonymous profile
Socialization	32) Do you frequent these kinds of places to meet people?	• LGBTQ places, or gay bars • LGBTQ discos • LGBTQ associations • No, I don't • Other:_____
	33) Do you feel part of the LGBTQ community?	• Yes, I do. • Yes, I do, but partially. • No, I do not, but I support it. • Other:_____
	34) Agreement with LGBTQ community claims: 5-levels Likert scale (strongly disagree, disagree, undecided, agree, strongly agree)	• Civil unions between same-sex couples • Marriage between same-sex couples • Child adoption for same-sex couples • Homophobia crime introduction in the Italian legislation • Greater expressive freedom
	35) Have you ever endorsed an LGBTQ claim on SNs?	• Yes, I did • No, I did not • I do not remember
	36) Have you ever been on a Pride?	• Yes, I have • No, I have not • No, but I wish I had

	37) Agreement with sentences about internet and LGBTQ associations : 5-levels Likert scale (strongly disagree, disagree, undecided, agree, strongly agree)	• Writing a post on SN allows me to easily reach many people • Instagram can be used to meet people who have my interests and hobbies • SNs are substituting LGBTQ associations as places for sharing sexual orientation and gender identity experiences • Associations are important for LGBTQ rights claims and to support those who need help, but to socialize nowadays there are more efficient tools available
Discrimination	38) Have you ever been a victim of discriminatory events into your household, and by whom? (My mother, my father, my sister, my brother, other relatives)	• I have been mocked • I have been insulted, e.g. swear words and/or aggressively treated • I have been shoved • I have been hit or beaten up • I was told to hide my sexual orientation • I was accused of causing a great sorrow, hurting my parents • People started to avoid me after I came out • I have been insulted or mocked, with homophobic comments on SNs • I was accused of spreading sexual-transmitted diseases (e.g AIDS)
	39) Have you ever been a victim of discriminatory events out of your household, and by whom? (My partner, my friends, my colleagues/classmates, my superiors/teachers, my sport mates, unknown people)	• I have been mocked • I have been insulted, e.g. swear words and/or aggressively treated • I have been shoved • I have been hit or beaten up • I was told to hide my sexual orientation • I was accused of causing a great sorrow, hurting my parents • People started to avoid me after I came out • I have been insulted or mocked, with homophobic comments on SNs

		• I was accused of spreading sexual-transmitted diseases (e.g AIDS)
	40) Have you ever been mobbed for sexual reasons, and by whom?	• Yes, I have, by colleagues • Yes, I have, by superiors • Yes, I have, by other people • No, I have never been mobbed • Other:_____
	41) Have you ever been bullied for sexual reasons at school/university, and by whom?	• Yes, I have, by classmates • Yes, I have, by other students • No, I have never been bullied • Other:_____
	42) Have you ever been discriminated in sport, and by whom?	• Yes, I have, by teammates • Yes, I have, by opposing team members • Yes, I have, by the coach • Yes, I have, by other sport club members • Yes, I have, by supporters • No, I have never been discriminated in sport • Other:_____
	43) Have you ever been teased on a SN?	• Yes, I have, by people whom I know in real life • Yes, I have, by people whom I do not know in real life • No, I have never been teased on SNs • Other:_____
Snowballing	Usernames to interview in turn	(3 possible usernames)
Privacy	Authorization to handle data, according to the law 196/2003	• Yes, I authorize the use of my data • No, I do not authorize the use of my data

R) Web-scraping R script

```
#-----------------------------------------------------
# Part of InstaCrawlR
# Git Hub: https://github.com/JonasSchroeder/InstaCrawlR
# Code by Jonas Schröder
# See ReadME for instructions and examples
#-----------------------------------------------------

#install.packages(c("readr","rlist", "stringr", "rvest", "readr", "tidyverse",
"ggplot2"))

library(readr)
library(rlist)
library(stringr)
library(rvest)
library(readr)
library(tidyverse)
library(ggplot2)


#---------------------------------------
# Functions to extract POST meta data
# @handle of author
# number of likes and comments
#---------------------------------------

extractMetaDataPost <- function(index){
    for(i in index:length(url_post_list)){
        print(index)
        url <- url_post_list[[i]]
        error <- tryCatch(
            source_temp <- read_lines(url),
            error=function(e) e
        )
        #no error
        if(!inherits(error, "error")){
            #source_temp <- read_lines(url_post_list[53])
            for(i in 1:length(source_temp)){

                # Extract date and time of post
                timestamp <- str_extract(source_temp[i],
"taken_at_timestamp\":[:digit:]*") %>%
                    str_remove("taken_at_timestamp") %>%
                    str_remove("\":")
                if(is.na(timestamp) == F){
                    datetime_temp <- as.POSIXct(as.numeric(timestamp),
origin="1970-01-01")
                    #print(datetime_temp)
                    post_datetime[index] <- datetime_temp
                }

                if(str_sub(source_temp[i], 14, 61) == "script
type=\"text/javascript\">window._sharedData"){
                    # Sponsored Post
                    sponsor_temp <- unlist(str_split(source_temp[i],
"sponsor")[[1]][3])

                    if(is.na(sponsor_temp)){
                        # Post not sponsored
```

```
                        post_sponsor[index] <- "not sponsored"
                    } else {
                        #Get Profile from ID using
https://www.instagram.com/web/friendships/USER-ID/follow/
                        sponsor_temp <- str_split(sponsor_temp,
"location")[[1]][1]
                        post_sponsor[index] <- str_split(sponsor_temp,
"\"")[[1]][5]
                    }


                    # Text
                    text_temp <- unlist(str_split(source_temp[i],
"edge_media_to_caption")[[1]][2] %>%
                                                str_split("caption_is_edited"))[1]
%>%
                        str_sub(30, -8)
                    #print(text_temp)
                    post_text[index] <- text_temp


                    # Hashtags in Text
                    post_hashtags_temp <-
paste(unlist(str_extract_all(text_temp,
                                                                        "#([A-
Za-z0-9_](?:(?:[A-Za-z0-9_]|(?:\\.(?!\\.))){0,28}(?:[A-Za-z0-9_]))?)")),
collapse = ' ')

                    if(post_hashtags_temp == "character(0)"){
                        post_hashtags[index] <- "No Hashtags"
                    } else {
                        post_hashtags[index] <- post_hashtags_temp
                    }


                    # Mentions in Text
                    post_mentions_temp <-
paste(unlist(str_extract_all(text_temp,
                                                                    "@([A-Za-z0-
9_](?:(?:[A-Za-z0-9_]|(?:\\.(?!\\.))){0,28}(?:[A-Za-z0-9_]))?)")), collapse =
' ')

                    if(length(post_mentions_temp) == 0){
                        post_mentions[index] <- "No Mentions"
                    } else {
                        post_mentions[index] <- post_mentions_temp
                    }
                }


                if(str_sub(source_temp[i], 6, 17) == "meta content"){
                    #author is always the first @handle
                    profile_temp <- str_extract(source_temp[i],
                                                "@([A-Za-z0-9_](?:(?:[A-Za-z0-
9_]|(?:\\.(?!\\.))){0,28}(?:[A-Za-z0-9_]))?)")
                    # post content preview
                    #text_temp <- str_extract(source_temp[i], "“(.*?)”") %>%
                    #    str_remove("“") %>%
                    #    str_remove("”")


                    #number of likes (nol), number of comments (noc)
                    nol_temp <- cleanNum(str_extract(source_temp[i],
"[:digit:]*[:punct:]?[:digit:]*[:alpha:]?[:space:]Likes") %>%
                                            str_remove("\"") %>%
```

```
                                            str_sub(1, -7))
                    noc_temp <- cleanNum(str_extract(source_temp[i],
"[:digit:]*[:punct:]?[:digit:]*[:alpha:]?[:space:]Comments") %>%
                                            str_remove("\"") %>%
                                            str_sub(1, -10))

                    #print(profile_temp)
                    #print(text_temp)
                    #print(nol_temp)
                    #print(noc_temp)

                    if(length(profile_temp) < 1){
                        print("some strange error")
                        insta_profiles[index] <- "strange error"
                    } else {
                        #no errors, save data
                        insta_profiles[index] <- profile_temp
                        #post_text[index] <- text_temp
                        post_likes[index] <- nol_temp
                        post_comments[index] <- noc_temp
                    }
                }
            }

        } else {
            print("Post is not available anymore")
            print(error)
            insta_profiles[index] <- "post deleted"
        }

        index <- index + 1
        assign("index", index, envir = .GlobalEnv)
        assign("insta_profiles", insta_profiles, envir = .GlobalEnv)
        assign("post_text", post_text, envir = .GlobalEnv)
        assign("post_sponsor", post_sponsor, envir = .GlobalEnv)
        assign("post_hashtags", post_hashtags, envir = .GlobalEnv)
        assign("post_mentions", post_mentions, envir = .GlobalEnv)
        assign("post_likes", post_likes, envir = .GlobalEnv)
        assign("post_comments", post_comments, envir = .GlobalEnv)
        assign("post_datetime", post_datetime, envir = .GlobalEnv)
        #Sys.sleep(2)
    }
}


# Extract Profile Infos
getProfileURL <- function(){
    for(i in 1:length(insta_profiles)){
        print(i)
        if(is.na(insta_profiles[i])){
            print("NA; not a post link")
            profile_url <- "not a post link"
        } else {
            if(insta_profiles[i] == "post deleted"){
                print("post deleted")
                profile_url <- "post deleted"
            } else {
                profile_name <- str_remove(insta_profiles[i], "@")
                profile_url <-
str_glue("https://www.instagram.com/{profile_name}")
                print(profile_url)
            }
        }
        url_profile_list <- list.append(url_profile_list, profile_url)
    }
```

```
    return(url_profile_list)
}

#------------------------------------------------------------------------------
-
# Functions to extract PROFILE meta data: Number of followers, following,
posts
# Clean Data
# Later match data for unique list with larger list that contains duplicates
#------------------------------------------------------------------------------
-

extractMetaDataProfile <- function(index){
    for(i in index:length(url_profile_list_unique)){
        print(index)
        # check for appropriate profile links
        if(url_profile_list_unique[i] != "post deleted" &
url_profile_list_unique[i] != "not a post link"){
            url <- url_profile_list_unique[index]
            error <- tryCatch(
                profile <- read_lines(url),
                error=function(e) e
            )
            #no error
            if(!inherits(error, "error")){
                #profile <- read_lines(url)
                for(j in 1:length(profile)){
                    if(str_sub(profile[j], 14, 43) == "meta
property=\"og:description\""){
                        #print("found line")
                        source_temp <- profile[j]
                        #print(source_temp)
                        meta <- strsplit(source_temp, "\"")[[1]][4] %>%
                            strsplit("[: :]") %>% unlist()
                        #print(meta)

                        follower_unique[index] <- getNoFollower(meta)
                        following_unique[index] <- getNoFollowing(meta)
                        posts_unique[index] <- getNoPosts(meta)

                        assign("follower_unique", follower_unique, envir =
.GlobalEnv)
                        assign("following_unique", following_unique, envir =
.GlobalEnv)
                        assign("posts_unique", posts_unique, envir =
.GlobalEnv)
                        index <- index + 1
                    }
                }
            } else {
                # probably HTTP 429 return
                print("something's wrong with the url")
                print(error)

                # HTTP 404 error -> post not available anymore
                is404 <- str_extract(error, "404")
                is429 <- str_extract(error, "429")
                if(is.na(is404) == F){
                    # next
                    index <- index + 1
                }
                if(is.na(is429) == F){
                    Sys.sleep(30)
                }
            }
```

```
        } else {
            # No profile link extracted: either post not available or wrong
link format (e.g. profile instead of post)
            print("Ignoring this element")
            index <- index + 1
        }

        assign("index", index, envir = .GlobalEnv)

        # optional: sleep to reduce risk of HTTP 429 returns
        Sys.sleep(1)
    }
}

getNoFollower <- function(meta){
    follower <- meta[1]
    #print(follower)
    if(length(cleanNum(follower) != 0)){
        follower <- cleanNum(follower)
        #print(follower)
        return(follower)
    } else {
        print("something's wrong: follower")
    }
}

getNoFollowing <- function(meta){
    following <- meta[3]
    if(length(cleanNum(following) != 0)){
        following <- cleanNum(following)
        #print(following)
        return(following)
    } else {
        print("something's wrong: following")
    }
}

getNoPosts <- function(meta){
    posts <- meta[5]
    if(length(cleanNum(posts) != 0)){
        posts <- cleanNum(posts)
        #print(posts)
        return(posts)
    } else {
        print("something's wrong: posts")
    }
}

cleanNum <- function(to_test){
    #clean <- as.numeric(to_test)
    if(str_detect(to_test, "\\.") & str_detect(to_test, "k")){
        clean <- str_remove(to_test, "\\.") %>%
            str_remove("k") %>%
            as.numeric() * 100
    } else if(str_detect(to_test, "\\.") & str_detect(to_test, "m")){
        clean <- str_remove(to_test, "\\.") %>%
            str_remove("m") %>%
            as.numeric() * 100000
    } else if(str_detect(to_test, "\\,")){
        clean <- str_remove(to_test, ",") %>%
            as.numeric()
    } else if(str_detect(to_test, "k")){
        clean <- str_remove(to_test, "k") %>%
            as.numeric() * 1000
    } else if(str_detect(to_test, "m")){
```

```
        clean <- str_remove(to_test, "m") %>%
            as.numeric() * 1000000
    } else {
        clean <- as.numeric(to_test)
    }
    return(clean)
}

# Match Data from url_profile_list_unique with larger url_profile_list
matchProfileData <- function(index){
    for(i in 1:length(url_profile_list)){
        for(j in 1:length(url_profile_list_unique)){
            if(url_profile_list[i] == url_profile_list_unique[j]){
                follower[i] <- follower_unique[j]
                following[i] <- following_unique[j]
                posts[i] <- posts_unique[j]

                assign("follower", follower, envir = .GlobalEnv)
                assign("following", following, envir = .GlobalEnv)
                assign("posts", posts, envir = .GlobalEnv)

                break
            }
        }
    }
}

#------------------------------------------------------------------
# START SCRIPT HERE
#------------------------------------------------------------------

setwd("C:\\Users\\M\\Desktop\\MARCO\\Dottorato_casa\\InstaCrawlR-
master\\out\\urls")

# Load Log Files and Import Your List of URLs
# German Excel uses ";" as separator -> read_csv2().
# If your Locale is EN, use read_csv() instead of read_csv2()
url_post_list <- unlist(read_csv("url_lesbicheitalia.csv"))
############################################################

# Optional: Take a Subset for Sampling / Testing (e.g., last 50 entries)
# url_post_list <- tail(url_post_list, 50)

# Extract Meta Data from Post: Author's @handle, Text, Hashtags, Mentions,
Number of Likes and Comments,
# Datetime, whether the Post is Sponsored or not (incl. user ID of the
sponsoring company)
insta_profiles <- c()
post_text <- c()
post_mentions <- c()
post_hashtags <- c()
post_sponsor <- c()
post_likes <- c()
post_comments <- c()
post_datetime <- c()
index <- 1
extractMetaDataPost(index)


# Save List of Profiles and Post Links
# Note: NA created in extractMetaDataPost() when link in list is not a post
link
# Optional: Deleting Entries with "post deleted"
profile_save <- insta_profiles
post_url_save <- url_post_list
```

```
if(any(insta_profiles == "post deleted") == F){
    print("no deleted posts in url_post_list")
} else {
    #deleting posts is optional
    #url_post_list <- url_post_list[-which(insta_profiles == "post deleted")]
    #insta_profiles <- insta_profiles[-which(insta_profiles == "post
deleted")]
}

# Get profile URL fro

m @handle
url_profile_list <- list()
url_profile_list <- unlist(getProfileURL())

# Get Rid of Duplicates
# Extract Profile Meta Data: Number of Followers, Following, Posts
url_profile_list_unique <- unique(url_profile_list)
follower_unique <- c()
following_unique <- c()
posts_unique <- c()
index <- 1
extractMetaDataProfile(index)

# Match Data for Export
follower <- c()
following <- c()
posts <- c()
index <- 1
matchProfileData(index)

# Combine and Export Data
export3 <- as.data.frame(cbind(insta_profiles, url_profile_list, follower,
following, posts,
                               unlist(url_post_list), post_text,
post_mentions, post_hashtags,
                               post_sponsor, post_datetime, post_likes,
post_comments), row.names = F)
write.csv(export3, "db_url_lesbicheitalia.csv", fileEncoding = "UTF-8", quote
= T)
########################################################################
```

```
#-------------------------------------------------------
# Part of InstaCrawlR
# GitHub: https://github.com/JonasSchroeder/InstaCrawlR
# Code by Jonas Schröder
# See ReadME for instructions and examples
#-------------------------------------------------------

library(stringr)

#Import Table and Extract Hashtags
text <- list()
htemp <- list()
htags <- data.frame()
data <- read.csv("table-HASHTAG-cleared.csv", sep = ";")
data <- as.matrix(data[-1])

maxrows <- nrow(data)
for(i in 1:maxrows){
    text[i] <- as.character(data[i,5])
    htemp <- str_extract_all(text[i], "#\\S+", TRUE)

    if(ncol(htemp) != 0){
        for(j in 1:ncol(htemp)){
            htags[i,j] <- htemp[1,j]
        }
    }
}

#Save Hashtags as csv for Excel
write.csv(htags, "ht_unsort_HASHTAG.csv", fileEncoding = "UTF-8")
df_htags <- as.data.frame(table(unlist(htags)))
write.csv(df_htags, "ht_sort_HASHTAG.csv", fileEncoding = "UTF-8")
```

```
#-------------------------------------------------------
# Part of InstaCrawlR
# Git Hub: https://github.com/JonasSchroeder/InstaCrawlR
# Code by Jonas Schröder
# See ReadME for instructions and examples
# Last Updated March 2019
#-------------------------------------------------------

library(jsonlite)
library(stringr)
library("jpeg")
library(tidyr)
library(utf8)

#-----------------------------------------------------------
#Download JSON File from Instagram for a specific Hashtag
#-----------------------------------------------------------
hashtag <- "lgbtitaly"
url_start <- str_glue("http://instagram.com/explore/tags/{hashtag}/?__a=1")
json <- fromJSON(url_start)
edge_hashtag_to_media <- json$graphql$hashtag$edge_hashtag_to_media
end_cursor <- edge_hashtag_to_media$page_info$end_cursor
posts <- edge_hashtag_to_media$edges$node

#-----------------------------
#Extract Information per Post
#-----------------------------
index <- 1
post_id <- list()
post_text <- list()
post_time <- list()
post_likes <- list()
post_owner <- list()
post_img_url <- list()

extractInfo <- function(index){
    print("extractInfo function called")
    maxrows <- nrow(posts)
    for(i in 1:maxrows){
        if(i == maxrows){
            assign("index", index, envir = .GlobalEnv)
            assign("post_id", post_id, envir = .GlobalEnv)
            assign("post_text", post_text, envir = .GlobalEnv)
            assign("post_time", post_time, envir = .GlobalEnv)
            assign("post_img_url", post_img_url, envir = .GlobalEnv)
            assign("post_likes", post_likes, envir = .GlobalEnv)
            assign("post_owner", post_owner, envir = .GlobalEnv)
            getNewPosts(index)
        } else {
            post_id[index] <- posts[i,5]

if(length(posts$edge_media_to_caption$edges[[i]][["node"]][["text"]])==0){
                post_text[index] <- "no-text"
                print("no text in post")
            } else {
                temp <-
posts$edge_media_to_caption$edges[[i]][["node"]][["text"]]
                post_text[index] <- gsub("\n", " ", temp)
            }

            post_time[index] <- toString(as.POSIXct(posts[i,7], origin="1970-
01-01"))
            post_img_url[index] <- posts[i,9]
            post_likes[index] <- posts[i,11]
            post_owner[index] <- posts[i,12]
```

```
            #optional: download image
            #img_dir <- str_glue("images/{index}_{hashtag}_post_img.jpg")
            #download.file(posts[i,8], img_dir, mode = 'wb')

            index <- index + 1
        }
    }
}

#------------------------------
#Get New Posts from Instagram
#------------------------------
getNewPosts <- function(index){
    print("getNewPosts function called")
    url_next <- str_glue("{url_start}&max_id={end_cursor}")
    json <- fromJSON(url_next)
    edge_hashtag_to_media <- json$graphql$hashtag$edge_hashtag_to_media
    end_cursor <- edge_hashtag_to_media$page_info$end_cursor
    posts <- edge_hashtag_to_media$edges$node
    assign("end_cursor", end_cursor, envir = .GlobalEnv)
    assign("posts", posts, envir = .GlobalEnv)
    print(index)
    Sys.sleep(5)
    extractInfo(index)
}

#Start the Madness
extractInfo(index)

#------------------------------
#Export Dataframe to CSV()
#------------------------------
table <- do.call(rbind.data.frame, Map('c', post_id, post_img_url, post_likes,
post_owner, post_text, post_time))
head(table)
dim(table)
colnames(table) <- c("ID", "URL", "Likes", "Owner", "Text", "Date")
#time <- Sys.time()
#filename <- str_glue("table-{hashtag}-{time}.csv")
write.csv(table, "out_lgbtitaly.csv")


#May run first to set TZ
Sys.setenv(TZ="Europe/Berlin")
Sys.getenv("TZ")
```

Update October 2018:
I added a new script (databaseCreator.R) which enables you to build your own Instagram database that you can use for Social Media Monitoring, comparing and selecting Influencers, or Competitive Analyses. databaseCreator scrapes Instagram based on a list of post URLs for Post Meta Data (text, hashtags, mentions, number of likes and comments) and Profile Meta Data (Author's @handle, number of followers, following, and posts).

More about databaseCreator in this Medium article.https://medium.com/@jonas.schroeder1991/build-your-own-instagram-database-134281e8ee92

InstaCrawlR
Crawl public Instagram data using R scripts without API access token.

Here's an example:
https://medium.com/@jonas.schroeder1991/social-network-analysis-of-related-hashtags-on-instagram-using-instacrawlr-46c397cb3dbe

Please consult "InstaCrawlR Instructions.pdf" for more information on what InstaCrawlR can and can't do and how to use it.

Jonas

Instagram is constantly changing their API's functionality (platform changelog). Following Facebook's Cambridge Analytica incident and the resulting public pressure, the API use got restricted even more severely in April 2018. The new limit is now 200 calls per user per hour instead of 5,000. More restrictions are announced to become active in July and December 2018.

The company's rational for restricting access to data is probably to prevent spamming behavior and data exploitation. However, since Social Media Platforms is now an integral part of everyday life, data gathered from these services have become more and more interesting for academic researchers.

In 2016, Instagram totally changed their API system. Developers have to submit their app to a rigorous permission review process in order to get an access token. Since academic researchers are not programming applications that are suitable for this review process (e.g., video-screen casting the app's functionality from an end user's point of view), they are basically unable to officially access valuable data for their research.

InstaCrawlR is a collection of R scripts that can be used to crawl public Instagram data without the need to have access to the official API. Its functionality is limited compared to what is possible using the official API. However, it seems to be the only option for non-developers to gather and analyze Instagram data.

Please note two things: As of July 2018, the scripts run as intended. This can change any time soon since Instagram is constantly limiting their API's functionality. Also keep in mind that using these scripts can have legal consequences since Instagram does not allow automated scripts. I am not responsible for consequences of any kind.

USE AT YOUR OWN RISK. BE ETHICAL WITH USER DATA.

What it can do

InstaCrawlR consist of four scripts – jsonReader, hashtagExtractor, graphCreator, and g2gephi – which are described in the instruction PDF. InstaCrawlR can be used to download and analyze the most recent posts for any specific hashtag that can be found on Instagram's Explore page (instagram.com/explore/tags/HASHTAG/). More specifically it can:

• Download the most recent posts for any hashtag

• Export a csv file that shows post ID, URL, number of likes, post owner ID, post text,
and post date

• Automatically extract related hashtags from post text

• Images can be automatically downloaded, too

• Export related hashtags and frequency

• Create a graph showing the relationship of related hashtags (social network analysis)

• Export graph for further analysis in Gephi

What it can't

• No specification of a certain timeframe (only most recent)

• No information on who liked the posts (only counter)

• Only post owner ID, not profile name

• Suspicious posts must be filtered out by hand using Excel

• No location information available

Please consult the instructions PDF for details.

Closing Words

You can use the script or parts in your own code. Please note that I am not a professional developer or trained programmer. I am sure InstaCrawlR's code can be simplified and improved a lot. Feel free to clean up my code or change it to increase its capabilities.
Again, use the scripts at your own risk. I am not reliable for any consequences. InstaCrawlR may only function for a limited time since Instagram is constantly changing their system. I will not necessarily support InstaCrawlR in the future.
If you have any comments or suggestions you can reach me on LinkedIn. I am always looking forward to a nice conversation about the future of digital marketing, entrepreneurship, and data science.

Best regards,
Jonas Schröder
University of Mannheim, July 2018

ACKNOWLEDGEMENTS

I would like to thank the following people, without whom I would not have been able to complete this research, and without whom I would not have made it through my PhD.

In the first place, my Mum who spent a lot of hours sending the questionnaires and answering questions about the survey.

The other PhD students for being a great team; especially Letizia Zampino, who supported me when I wanted to give up.

My supervisor, Assunta Viteritti, for leaving me free to express myself.

The Istat, in the person of my superiors and colleagues, which gave me the leave of absence to deepen such an innovative and unexplored issue, proving to be a cutting-edge research institute.

All the LGBTQ people I interviewed, and the professors I asked for advice (Fabio Corbisiero, Luca Massimiliano Visconti, Salvatore Monaco), whose insight and knowledge into the subject matter steered me through this research.

The GIFTS, Italian Network of Gender, Intersex, Feminist, Transfeminist and Sexuality Studies that welcomed me.

My family, proud of me.

Last, but not least, I would like to thank the person who shared with me the quarantine, making the seclusion lighter, tender and even amusing.

Thanks again to all of them for getting me here.