Tops, Bottoms, and Versatiles: On the Need to Control for Intercourse Position Subgroups when Researching MSM

Young-Il Kim

STUDENT NUMBER: 6000630

Preface

I would like to thank my parents and grandparents, in particular my mother and grandmother, for always supporting me and for accepting me as I am unreservedly. I would like to thank Dr. Tali Spiegel and Prof. Dr. John de Wit for their guidance, comments and encouragement. Last, but certainly not least, I like to to thank Liyang Zhou and my partner Zhang Wei both for supporting me and helping me make sense of articles written in Chinese.

Abstract

It is common for gay men and MSM across the world to apply the terms 'top', 'bottom' and 'versatile' to themselves and to others. Some researchers take these terms to indicate preferences for insertive and/or receptive anal intercourse (IAI/RAI). Others take these terms to indicate actually enacted anal intercourse behaviour. Finally, some take these terms as indicating (sexual) identities related to anal intercourse position preferences and/or behaviour. However, what these terms mean /or what they conceptualise is not yet well-understood and so far this remains an understudied topic. The few studies that do focus on these subgroups (tops, bottoms and versatiles) find them to differ significantly with respect to many aspects. Given this, not controlling for these subgroups threatens to undermine the findings obtained in many studies involving gay men or MSM. In this study I conduct a meta-analysis of the distributions of these subgroups in the samples of the few studies that did record this information to fill a gap in the literature and to provide insights into what is potentially a serious methodological problem. It is argued that these subgroups ought to be controlled for in future research and, moreover, that research into more adequate measures ought to be conducted.

Key words: top, bottom, versatile, gay, homosexual, MSM.

Contents

Tops, Bottoms, and Versatiles: On the Need to Control for Intercourse Position Subgroups when Researching MSM

Young-Il Kim

1 Introduction

Gay men or MSM more generally have long since been a topic of considerable research interest as being a member of one of these sexual minority populations is associated with (amongst other things) less well-being, lower life satisfaction and poorer physical and mental health (Adelman, 1991; Meyer, 1995; Meyer & Dean, 1998; Frost & Meyer, 2009; Frost et al., 2015; Berggren et al., 2017).

Members of these minority groups in the West commonly use the terms 'top', 'bottom' and 'versatile' to refer to themselves and to others. A similar practice is observed in Latin America where in the same order the terms 'activo', 'pasivo' and 'moderno/internacional/versatil' are used instead, and in China, where these terms are replaced by the numbers '1', '0' and '0.5' respectively (Carrier, 1977; Wegesin & Meyer-Bahlburg, 2000; Hart et al., 2003; Gil, 2007; Moskowitz et al., 2008; Johns et al., 2012; Zheng & Zheng, 2009; Wei & Raymond, 2011). Usually, the term 'top' is regarded as equivalent to the terms 'activo' and '1', the term 'bottom' is regarded as equivalent to the terms 'pasivo' an '0', and the term 'versatile' is regarded as equivalent to the terms 'moderno', 'internacional', 'versatil' and '0.5' (Zheng et al., 2013; Zheng et al., 2015).

I purposefully and explicitly include the word 'term' here since it is not yet well-understood what the terms 'top', 'bottom' and 'versatile' conceptualise. Some researchers take these terms to indicate *preferences* for positions taken during anal intercourse. Consequently these researchers take or define a 'top' to be someone who *prefers* insertive anal intercourse (IAI), a 'bottom' to be someone who *prefers* receptive anal intercourse (RAI), and a 'versatile' to be someone who *prefers* IAI and RAI approximately equally (Wei & Raymond, 2011; Zhou et al., 2013; Pan et al., 2016). Rather than preferences, others take these terms as indicating *actually enacted anal course behaviour*, defining a 'top' as someone who *actually engages* mostly in IAI, a 'bottom' as someone who *actually engages* mostly in RAI, and a 'versatile' as someone who *actually engages* in both IAI and RAI approximately equally (Halkitis et al., 2005; Wei & Raymond, 2011; Lei et al., 2018). Finally, some take these terms as indicating (sexual) identities related to anal intercourse position preferences and/or behaviour (Wegesin & Meyer-Bahlburg, 2000; Hart et al., 2003; Zheng et al., 2012).

Researchers stipulating the terms 'top', 'bottom' and 'versatile' to indicate *(anal) intercourse position preference* (IPP) *usually* adopt an appropriate preference-based measure to divide the MSM constituting the sample used in their respective studies into subgroups of preference-tops, preference-bottoms and preference-versatiles (TBV_P-subgroups). Similarly, those stipulating these terms to indicate *(anal) intercourse position behaviour* (IPB) *usually* adopt an appropriate behavioural to distinguish between behavioural-tops, behavioural-bottoms and behavioural-versatiles (TBV_B-subgroups). Finally, those taking these terms to indicate *(anal) intercourse position identities* (IPI) *usually* adopt an appropriate identity-based measure to delineate subgroups of identity-tops, identity-bottoms or identity-versatiles (TBV_I-subgroups) from one another.

The fact that in quite a few studies a measure *in*appropriate to the stipulated meanings of the terms 'top', 'bottom' and 'versatile' is used and, furthermore, that this has so far gone completely unnoticed, indicates that it is still poorly understood what these terms mean or conceptualise. Moreover, it is a matter of serious methodological concern. At best it merely misrepresents research findings, e.g. participants were asked about anal intercourse *behaviour* but findings were presented as if applying to TBV_P–subgroups. At worst it undermines the research findings, e.g. when prior to answering which of the terms 'top', 'bottom' or 'versatile' applies to them, research participants were presented with *behavioural* definitions of the terms 'top' and 'bottom' and a *preference-based* definition of the term 'versatile'. As preference for a certain behaviour need not coincide with that behaviour being actually enacted, this entails the response options can no longer be regarded as being mutually exclusive, increasing (the risk of) measurement error (Weisberg, 2009; Rubin & Babbie, 2009; Toepoel, 2015).

Turning these matters of methodological concern into matters of great concern is the fact that regardless of the kind of measure used many studies find the three TBV_P-subgroups, the three TBV_B-subgroups and the three TBV_I-subgroups to differ significantly with respect to variables central to much of the research focussing on gay men or MSM more generally.

Studying TBV_B-subgroups, social dominance orientation (SDO) of tops was found to be significantly higher than that of bottoms and versatiles (Tan et al., 2013), bottoms and versatiles were found to be more gender-nonconforming than tops (Swift-Gallant et al., 2017) and to have earlier pubertal onset (Swift-Gallant et al., 2018). Using an Australian sample, Lyons et al. (2011) found versatiles were less likely to know their HIV status than tops or bottoms. Moreover, many epidemiological studies conducted across the world find men who mostly engage in RAI (behaviourally-defined bottoms) are at much higher risk of attracting STDs and/or HIV than men who mostly engage in IAI, i.e. behaviourally-defined tops (García et al., 2012; Lau et al., 2013; Feng et al., 2015; Meng et al., 2015).

Focusing on TBV_P-subgroups, Valentova et al. (2014) found men preferring "active" sex roles to also prefer having shorter partners whereas the opposite was true for men preferring "passive" sex roles. Moreover, bottoms and versatiles being more gender-nonconforming and having earlier pubertal onset was found to

hold too when using a preference-based measure (Swift-Gallant et al., 2017; Swift-Gallant et al., 2018). Other differences between the three different TBV_P-subgroups include, but are not limited to, differences in perception of HIV-risk and HIV-testing (Zhang et al., 2013b), carrying condoms (Wang et al., 2015), using condoms (Wang et al., 2015; Dangerfield et al., 2016) and being HIV positive (Zhou et al., 2013).

Looking at TBV_I-subgroups, significant differences between the TBV_I-subgroups were found with respect to sexual behaviour (Zheng et al., 2012; Lei et al., 2018), gendered traits and interests (Zheng et al., 2012), femininity (Wegesin & Meyer-Bahlburg, 2000), self-ascribed masculinity, femininity and cognitive styles (Zheng et al., 2015), condom use (Dangerfield et al., 2016), risk of getting infected with HIV (Wei & Raymond, 2011) and sexual satisfaction (Zheng & Zheng, 2017). Moreover, as both Hart et al. (2003) and Gil (2007) found significant differences with respect to internalised homophobia and internalised homophobia itself linked to, for instance, relationship quality (Frost & Meyer, 2009), depressive symptoms (McLaren, 2016), and intimacy (Meyer & Dean, 1998), the TBV_I-subgroups plausibly differ with respect to these and other outcomes correlated with internalised homophobia too.

Given that the three TBV_P-subgroups, the three TBV_B-subgroups and the three TBV_I-subgroups differ significantly with regards to so many aspects it is possible and plausible that both the effect size and the statistical significance of the findings obtained from analysing a sample of gay men/MSM depend to some degree on the particular distribution of the TBV_P-subgroups, the distribution of the TBV_B-subgroups and the distribution of the TBV_I-subgroups (these are plausibly correlated) in that sample. This is problematic since in most studies the TBV_P-distribution, the TBV_B-distribution, and/or the TBV_I-distribution is unknown because participants were not asked about their IPP, IPB or IPI. Not only may this affect the findings obtained in any one particular study, it also makes it problematic to compare the results of different studies, as differences in the results obtained may be due to differences in the TBV_P-distribution, the TBV_B-distribution, and/or the TBV_I-distribution of the samples used. Whether these distributions differ significantly between samples is currently not known, because (a) most studies do not ask participants about IPP, IPB or IPI; (b) the studies that do ask participants about IPP, IPB or IPI do not compare the TBV_P, TBV_B or TBV_I-distribution in the obtained sample to the corresponding distribution in samples obtained in other studies; (c) if the TBV_P, TBV_B or TBV_I-distribution in the sample obtained is compared to the corresponding distribution in samples obtained in other studies, these comparisons are not systematic but involve one or at most two comparisons.

In this paper I aim to fill this gap in the literature by systematically collecting data from studies that did ask participants about IPP, IPB or IPI, and by then answering the following research questions:

R1: What is the average TBV_I-distribution of the samples from studies where participants were asked about IPI, the average TBV_P-distribution of the samples from studies where participants were asked about IPP, and the average TBV_B-distribution of the samples where participants were asked about IPB?

R2: Do the average TBV_I-distribution, the average TBV_P-distribution, and the average TBV_B-distribution differ from one another?

I am purposefully avoiding any mention of TBV_P-, TBV_B- or TBV_I-distributions *in the population* in formulating R1. This is because the samples obtained in most studies involving populations of gay men or MSM are obtained using non-probability sampling-methods. Similarly to how the results obtained in such studies cannot be generalised to the underlying population from which the non-probability samples were taken, the average TBV_P-, TBV_B- or TBV_I-distributions calculated from a (large) number of samples must not be taken as an approximation of the TBV_P-, TBV_B- or TBV_I-distributions in some underlying population of gay men or MSM from which these samples were taken. This holds true especially since the different samples used for calculating the average TBV_P-, TBV_B- and TBV_I-distributions were often obtained using different (non-probability) sampling methods, but also because the samples included in this study were obtained from different regions whose gay and MSM populations may differ from one another in various regards. In particular, the distribution of the TBV_P-, TBV_B- and TBV_I-subgroups may be different in gay and MSM populations in different regions. Although a systematically collected sample consisting of non-probability samples obtained across the world does not allow for answering questions regarding the underlying populations, the following research questions can be asked (and answered) legitimately:

R3: Do the average TBV_I-distribution, the average TBV_P-distribution, and the average TBV_B-distribution differ between the regions Asia, Latin America and the West?

R4: Do the regional average TBV_P-, TBV_B- and TBV_I-distributions differ from the global TBV_P-, TBV_B- and TBV_I-distributions?

In order to assess the usefulness of the practice adhered to by some researchers whereby the TBV_X-distribution (X = P, B or I) in the sample obtained for their study is compared to the corresponding distribution in one or two other samples, I aim to answer the following research question:

R5: What is the probability that the TBV_X-distribution (X = P, B or I) in a given sample significantly differs from the TBV_X-distribution (X = P, B or I) in another sample?

In this study I do not aim to merely describe and compare distributions but also to explain and understand what factors (if any) affect or may predict the proportion of tops, bottoms and versatiles in a given sample.

Similarly to how the terms 'top', 'bottom' and 'versatile' may be used to indicate either preferences, actually enacted behaviour, or a sexual identity, the term 'gay' or 'homosexual' can be used to indicate either preferences (Laner & Laner, 1979; Plug & Berkhout, 2004), actually enacted behaviour (Kinsey, 1941; Cochran & Mays, 2000; Bauer & Brennan, 2013), or (sexual) identity (Harada, 2002; Rosario et al., 2006; Hammack &

4

Cohler, 2009). More common, however, is to see these terms in relation to the concept of sexual orientation, which is usually considered as *not* being defined along just one of the aforementioned dimensions but as including multiple dimensions that each capture part of the concept (Kinsey et al., 1948; Joseph et al., 1991; Dube, 2000; Gates, 2011).

Not surprisingly, not all MSM identify as gay as category of MSM also includes, amongst others, men identifying as bisexual (Hodge, 1995; Boellstorff, 2011). Similarly to many studies finding significant differences between the TBV_X-subgroups, there is a large body of literature showing gay-identified MSM (GI-MSM) to differ significantly from non-gay-identified MSM (NGI-MSM) in many aspects, e.g. with respect to sexual behaviour, sexual risk taking and HIV infection rates (Weatherburn et al., 1998; Heckman et al., 1999; Rietmeijer et al., 1998; Washington et al., 2006; Wohl et al., 2002). As the earlier mentioned studies also found significant differences between tops and bottoms, regardless of whether these two subgroups were defined behaviourally, on the basis of preference, or on the basis of identity, it is natural to ask if there is a relation between being GI or NGI and being top or bottom. In this paper, I examine this by answering the following two research questions:

R6: What is the relation between the proportion of GI men and the proportion of *behaviourally defined tops* in samples consisting of MSM, when sampling period and mean age of the sample are controlled for?

R7: What is the relation between the proportion of GI men and the proportion of *behaviourally defined tops* in samples consisting of MSM, when sampling period and mean age of the sample are controlled for?

Some remarks are in order: (1) research questions R6 and R7 only ask whether a relation holds in samples, rather than in some underlying population from which these samples were taken. This formulation is deliberate, since answering questions about some underlying population would at minimum require assuming that these samples were all taken from the same population and that they were all representative of this population. Both these assumptions are implausible. Moreover, in this paper I examine if there are significant differences between MSM populations in different geographical regions, so it would be premature to assume any findings may be generalised to some global population of MSM; (2) versatiles are purposefully excluded from the present study. Whilst studying versatiles would certainly be interesting, problems with inadequate measures or inconsistent use of measures mostly concern versatiles, and given that this area is understudied, there simply is not enough data available for carrying out the required analysis; (3) limitations on the available data also allow for a focus on behavioural-defined tops and behavioural-defined bottoms only. It also determines the choice of control variables, although there are also good theoretical reasons for controlling for these factors.

Various authors remark that concept of homosexual identity may differ between the regions under consideration in this study (Lancaster, 1988; Altman et al., 2002; Decena, 2004; Wong, 2011). Furthermore, attitudes towards homosexuality differ between these three regions, with attitudes in the West generally being more

positive than attitudes in Asia and Latin America (Crawford & Solliday, 1996; Yang, 1997; Banse et al., 2001; Kelley et al., 2001; Liu & Choi, 2006; Besen et al., 2007; Gerhards, 2010; Smith, 2011; Ayoub & Garretson, 2017; Xie & Peng, 2018). As such, it is reasonable to expect there to be regional differences in the relations enquired about in R6 and R7 and hence I also aim to answer the following two final research questions:

R8: Is the main effect enquired about in R6 moderated by sampling region?

R9: Is the main effect enquired about in R7 moderated by sampling region?

Scientific and societal relevance

Scientifically, answering R1–R5 is important for a few reasons. Firstly, it addresses an *epistemological* gap in the literature that no one so far has touched upon. Whilst it is known that the labels 'top', 'bottom' and 'versatile' are used ubiquitously in populations of gay men/MSM, not much is known about what these terms mean to them and thus what type of measure would be most appropriate. This paper is not concerned with this issue, however, for this would require a qualitative approach. However, the findings of this paper do bear on the need of conducting such qualitative research, for it may very well be that what type of measure is used makes little difference. Whilst there are some studies showing that there is some incongruence between preferences and enacted behaviour, the two are nevertheless strongly related. However, these studies examine this only within the confines of a single study whilst the current study looks at many studies. As such, it provides the much needed overview of whether TBV_X-distributions differ significantly depending on what measure is used, and hence, whether there is a *practical* need to sharply distinguish the type of measure used and consequently, for research into developing more accurate measures. Research questions R6–R9 build upon this and look for factors that explain the occurrence of distributions in samples. Since research involving populations of gay men/MSM use non-probability samples, and the various studies included in the "sample of samples" used for this study also use different measures (even if the same type of measure is used, there are also variations within a particular type of measure) and sampling methods, I am reluctant to generalise the findings from being about a collection of samples to some underlying population, anyone willing to interpret the results as such is free to do so.

Secondly, this study addresses a *methodological* concern that has been touched upon a few times. Though TBV_X-subgroups are an understudied topic, the research that has been carried out shows there are significant differences between these subgroups, regardless of what measure one uses, with respect to variables that are of key interest in conducting research into gay men/MSM. Moreover, research findings are potentially sensitive to the particular distribution of TBV_X-subgroups used in the sample. This risk always persists, unless one controls for TBV_X-subgroups. However, whether the problem is large or small in scope depends in part on the answer that is obtained for R1–R5: if comparing various (average) distributions with one another and these do not statistically differ from one another, we have less reason to call into question the results obtained in various studies in which TBV_X-subgroups were not controlled for, because although results may be sensitive to the

distribution of TBV_X-subgroups in the sample used, if these distributions in different samples are comparable most of the time, so are findings based on such samples and the risk that particular findings are merely the result of some extreme TBV_X-distribution is lower. Note, however, that the doubt always remains whether particular findings were due to an extreme or unusual distribution of the TBV_X-subgroups unless TBV_X-subgroups are controlled for. If this is done, a new concern arises, in particular if these subgroups turn out to differ significantly with regards to even more aspects than what has so far been uncovered, namely whether it is appropriate to ever try and generalise results to a population of gay men or MSM.

From a societal point of view, the study is important for several reasons too. Firstly, as men who engage in RAI are more at risk of STDs and HIV, knowing the TBV_X-distributions allows us to better estimate how many men this concerns, and, in turn, to more appropriately align policy measures aimed at addressing the specific challenges faced by these already vulnerable subgroups. Secondly, if current research into populations of gay men/MSM suffers from a methodological problem, than addressing this issue is also of societal concern, because it is through science that we learn about and chart existing problems, and explore possible solutions. Thus, addressing the methodological concern has direct bearing upon the societal value of this study too.

2 Theory & Hypotheses

Hegemonic Masculinity Theory

In Connell's (1995)'s theory of hegemonic masculinity, hegemonic masculinity (HM) is defined as set of values that serves to legitimise the dominant position of men in society and functions to organise society in gender unequal ways (Jewkes et al., 2015). It is assumed that there is a most-valued way of manhood, a gold-standard upon what it means to be a "real man" is defined. All men position themselves in relation to this standard and occupy a position relative to it. This gives rise to a hierarchy of masculinities in which men who do not (or who are not able or willing to) embody hegemonic masculinity occupy sub-ordinated masculinities (Connell, 1995; Connell & Messerschmidt, 2005)

HM culture dichotomises sexual roles into an insertive and a receptive role. As the insertive role is seen as the "man's role" and the receptive role as the "woman's role", this aligns gender roles, gender identity and sexuality with sexual acts (Lovaas & Jenkins, 2007; Acaba, 2013; Lick & Johnson, 2015). Moreover, this embeds the meanings of the terms 'top' and 'bottom' (the literature mostly remains silent about the term 'versatile') within gender discourses as the same framework for describing heterosexual intercourse is applied to sexual acts between to men, which are framed too in terms of an insertor and an insertee. This is irrespective of whether the terms 'top', 'bottom' and 'versatile' purvey preferences regarding anal intercourse, actually enacted anal intercourse behaviour, or a sexual identity, as in all three cases the term 'top' is associated with being the insertor, and the term 'bottom' is associated with being the insertee. As such, the term 'top' becomes associated with masculinity, and the term 'bottom' with femininity (Pathela et al., 2006; Yarhouse et al., 2011).

As HM is associated with displaying certain characteristics associated with and befitting to the gold-standard of masculinity, tops are attributed characteristics such as physical strength, dominance, and power. Bottoms, on the other hand, become associated with subordinated masculinities or femininity and are attributed characteristics such as passivity and being slender (Kippax & Smith, 2001; Carballo-Diéguez et al., 2004; Johns et al., 2012; Ravenhill & de Visser, 2018).

According to Lick & Johnson (2015), the scientific evidence for an overlap between gendered attributes and the labels 'top' and 'bottom' is mixed. As evidence against such view, they point towards the findings of Kippax & Smith (2001) that some bottoms actively oppose ascriptions of femininity and being powerless, viewing the receptive role as giving one power over one's sexual partner (Kippax & Smith, 2001), and towards the findings of Wegesin & Meyer-Bahlburg (2000) that tops and bottoms do not differ significantly with regards to gender-role identity at any point during the life-course.

Before considering evidence pointing in the opposite direction, it useful to distinguish between self-ascribed masculinity (SM) and ascribed masculinity (AM). SM concern ascriptions of masculinity relating to the self, and depends on the extent to which someone accepts the socially constructed image of masculinity; AM

concerns ascriptions of masculinity to others, and depends on the extent to which someone embodies the socially constructed image of masculinity (Thompson et al., 1992).

Making this distinction, there is strong evidence suggesting the terms 'top' and 'bottom' are gendered. Considering ascriptions of masculinity/femininity first, Han (2008) found American Asian Pacific Islander men were seen as feminine and expected to be bottoms, whereas black men were seen as hyper-masculine and expected to be tops. Similar findings regarding ascriptions of masculinity to black men and the expectation for them to be tops were obtained by Whitehead (1997) and Bowleg (2004). Further evidence comes from Tskhay & Rule (2013), who found that heterosexual research participants were able to accurately identify gay men's sexual role preferences on the basis of gendered facial features. Similar perceptions of tops being masculine and bottoms being feminine were found in studies conducted in Latin America (Carballo-Diéguez et al., 2004). Finally, looking at what is to my knowledge the only study into self-ascriptions of masculinity, Zheng et al. (2012) finds that tops scored higher on SM than bottoms which shows that it is not just others who think of tops as more masculine than bottoms, but that tops considers themselves to be more masculine than bottoms too, and, vice versa, that bottoms consider themselves to be less masculine than tops.

The previous paragraph provided evidence that it seems to be the common view that tops and bottoms differ with respect to masculinity. In accordance with the hegemonic masculinity theory, this may be because tops possess characteristics associated with masculinity, whilst bottoms possess characteristics associated with femininity. There is evidence that this is indeed the case e.g. Moskowitz & Hart (2011) found masculinity and self-reported penis size to differ significantly between tops, bottoms and versatiles regardless of whether a behavioural or preference-based measure was used. Moreover, using Chinese samples, the findings of Zheng & Zheng (2016) indicate tops and bottoms differ with respect to partner choice, with bottoms prefer partners with more masculine facial, body, and personality traits and the opposite being the case for tops. In a different study, Zheng & Zheng (2017) found tops to be more ego-focuses than bottoms with regards to sexual satisfaction, suggesting there may be difference with regards to dominance. Indeed, Tan et al. (2013) found lower social dominance orientation (SDO) to be an important moderator for engaging in receptive anal intercourse. Finally, bottoms were found to have displayed higher degrees of childhood gender-nonconformity by Weinrich et al. (1992), and to have enjoyed gender-atypical interests during childhood more Carrier (1977).

Taken together, and *pace* Lick & Johnson's (2015) reservations, it seems that hegemonic masculinity theory is appropriate for analysing differences between tops and bottoms. Tops have more masculine characteristics than bottoms and are also viewed as more masculine by others; bottoms have more feminine characteristics than tops and are also viewed as more feminine by others. In a hegemonic masculine society, this places tops above bottoms on the masculinities ladder. However, by itself this does not allow for deriving hypothesis regarding the relationship between being top/bottom and being gay-identified/non-gay-identified. Such a relation can be explained by considering social identity theory (SIT).

Social Identity Theory

In order to make sense of the complex world surrounding us, people simplify reality by categorising things into a number of differentiated groups or classes. Similarly, they simplify social reality by categorising other people into a number of differentiated groups or social categories on the basis of shared values, prototypical behaviours, and characteristics, e.g. sexual orientation (Leaper, 2011), a process known as *social categorisation* (Hogg & Reid, 2006; Ellemers & Haslam, 2011; Zepp et al., 2013; Ohlert & Zepp, 2016). When a person categorises himself into one or more of social categories or groups, this is known as social self-categorisation (Cox & Gallois, 1996) and necessarily involves cognitively awareness of belonging to these particular social categories or these particular groups (Tajfel, 1974; Tajfel, 1978a; Tajfel, 1978b).

Social identity theory (SIT) posits people to have both a personal identity (PI) and a social identity (SI). The former consists of values, characteristics and behaviours unique to the individual which distinguish him or her from others (Cox & Gallois, 1996), the latter derives from the membership of particular social categories/groups the individual belongs to through the process of social self-categorisation (Scheepers & Ellemers, 2019). Together SI and PI comprise a person's self-concept (Turner, 1987; Hogg & Abrams, 1988).

SIT assumes people strive to evaluate their own worth positively, i.e. to have a positive self-concept and positive self-esteem. For this it is necessary that they perceive both their SI and SI as positive (Abrams, 2001; Oldmeadow & Fiske, 2010). Positive (negative) personal self-esteem derives from comparing oneself favourably (unfavourably) to other individuals, whereas positive (negative) social self-esteem derives from comparing the social categories and/or groups one is a member of favourably (unfavourably) to social categories and/or groups one is not a member of (Luhtanen & Crocker, 1992). People may achieve positive personal self-esteem by comparing themselves with others along dimensions that allow them to make downward comparisons; similarly members of a group can achieve positive social self-esteem by comparing the in-group to out-groups along dimensions that allow for making downward comparisons (Cox & Gallois, 1996).

Within society some social groups are able to define the social status quo because they are more powerful, enjoy more prestige or are otherwise dominant. These social groups are able to determine what factors are key in differentiating groups from one another and hence along which dimensions groups are to be compared. Moreover, their dominant position allows them to choose these dimensions in such a way that make achieving positively valued distinctiveness between the in-group and out-groups easier for their members, but more difficult for members of other devalued groups (Turner, 1981).

Being a member of a devalued group may pose a threat to an individual's (positive) self-esteem as achieving positively valued distinctiveness between the in-group and out-groups is made more difficult. In such cases individuals may try and improve their situation by adopting either an individual strategy or a collective strategy aimed at social change (Hirschman, 1970; Tajfel & Turner, 1979).

SIT predicts that when group boundaries are perceived as permeable individuals will try and improve their position through *individual mobility* (Becker & Barreto, 2019), an "individual-level strategy whereby people may seek to escape, avoid, or deny belonging to a devalued group, and seek instead to be included in (or attempt to "pass" as a member of) a group of higher social standing" (Ellemers & Haslam, 2011, 382). Individual mobility attempts are aimed at leaving the devalued group physically or psychologically (Tajfel, 1981) and reduce group identification (Ellemers, 2001; Reese et al., 2019).

Hypotheses for R1 and R5
Hypotheses for the descriptive questions R1 and R5 cannot be directly derived from the theory and, moreover, the literature does not provide much insights either as this study is the first to consider these matters. Hence I do not formulate hypotheses for these two research questions.

Hypotheses for R2–R4 and R6–R9
In SIT social self-categorisation central involves more than placing oneself into one or more social categories; it also involves "an adoption over time of the normative (prototypical) behaviors, characteristics, and values associated with the particular group membership" (Cox & Gallois, 1996, 11). Thus some degree of congruence between IPP, IPB and IPI is expected in individuals applying one of the terms 'top', 'bottom' or 'versatile' to themselves. Indeed, Wegesin & Meyer-Bahlburg (2000), Hart et al. (2003), Moskowitz et al. (2008), and Wei & Raymond (2011) all found some degree of consistency between IPI and IPB and Swift-Gallant et al. (2017) observed the same for IPP and IPB. As such I expect that the average TBV_I-distribution, the average TBV_P-distribution, and the average TBV_B-distribution do not differ (statistically) from one another. This is, however, a very weak expectation since there is no perfect correlation between IPP, IPB and IPB.

According to hegemonic masculinity theory, heterosexual masculine men occupy a dominant position within society, i.e. they are the dominant social group. What differentiates groups from one another is the degree to which their members meet the gold-standard of manhood. In the hierarchy of masculinities, NGI-MSM rank above GI-MSM. This entails that within a hegemonic masculine society the *group* of NGI-MSM ranks above the *group* of GI-MSM.

Note that members of the group of GI-MSM could, in principle, improve their position through individual mobility by giving up identifying as gay. The fact that these men do not do so may indicate (a) that being a member of these group does not threaten these men's positive self-esteem despite of the fact that achieving positive valued distinctiveness between the in-group and out-group is more difficult to achieve.

Now (a) must be rejected because the group boundaries are permeable according to how permeability is specified in SIT, namely in terms of the *possibility* of moving between groups (Ellemers et al., 1988). There may be various objections against going back into the closet and hence it may not be desirable but this does not diminish that it is in principle possible to move between groups. In case (b) then it seems that being associated

with femininity matters less or is less objectionable to GI-MSM than to NGI-MSM because it is precisely this association with femininity what devalues the group of GI-MSM vis-à-vis the group of NGI-MSM. As engaging in RAI (being a behaviourally-defined bottom) is associated with femininity, it is plausible that GI-MSM find engaging in RAI less objectionable than NGI-MSM and hence that they actually engage more in RAI than NGI-MSM. As such we would expect there to be a positive relationship between the proportion of gay-identified men in samples and the proportion of behaviourally-defined bottoms. Analogously, we expect that the proportion of tops to be higher in samples where the proportion of NGI-MSM is higher, or equivalent, we expect the proportion of tops to be lower in samples where the proportion of GI-MSM is higher. Indeed, such a relationship is found by Hart et al. (2003) which to my knowledge is the only study to ever consider the relation between gay-identification and being top/bottom.

Engaging in RAI and thereby violating a gender norm by doing something un-masculine or feminine is viewed more negative in societies where traditional gender norms are stronger and where masculinity is highly prized. It is therefore expected that in samples taken from such societies there will be proportionally less men who are GI and proportionally less men who engage in RAI only.

Whilst the percentage of GI men and the percentage of behaviourally-defined bottoms in such samples may both be lower, it is possible that the association of being GI and being a behaviourally-defined bottom is *stronger* in societies where traditional gender norms are stronger and masculinity prized more. Indeed, it has been found that in various non-Western societies where traditional gender norms are strong and masculinity is highly prized, engaging in IAI with other men does not lead to one being considered as gay, e.g. in Turkey and Egypt (Dunne, 1998; Simmons, 2014) and Latin America (Murray, 1987; Lancaster, 1988; Carrier, 1989a; Carrier, 1989b; Abramson & Herdt, 1990; Almaguer, 1993; Morales, 1995; Cantú, 2009; Díaz, 2013; Clark et al., 2013). As how one is viewed by society plausibly affects how one views oneself – e.g. many studies find internalising societal homophobia affects homosexual identity formation (Rowen & Malcolm, 2003; Harris et al., 2008; Brown & Trevethan, 2010; Greene & Britton, 2012) – this may entail that in such societies only (or mostly) GI-MSM engage in RAI and, vice versa, only (or mostly) men who engage in RAI regard themselves as gay. As such, the relation between being GI and engaging in RAI (and being a behaviourally-defined bottom) might be stronger in such societies rather than weaker.

I formulate the following hypotheses on the basis of the above considerations, (the numbers correspond to the numbers of the research questions):

H1: –

H2: The average TBV_I-distribution, the average TBV_P-distribution, and the average TBV_B-distribution are not (statistically) different from each other.

H3: The average TBV_I-distribution, the average TBV_P-distribution, and the average TBV_B-distribution differ between the regions Asia, Latin America and the West.

H4: The regional average TBV_P-, TBV_B- and TBV_I-distributions differ from the global TBV_P-, TBV_B- and TBV_I-distributions (here 'global' refers to the combination of the three considered regions).

H5: –

H6: The proportion of GI men in a sample is negatively related to the proportion of behaviourally-defined tops in that sample.

H7: The proportion of GI men in a sample is positively related to the proportion of behaviourally-defined bottoms in that sample.

H8: The effect of the proportion of GI men in a sample on the proportion of behaviourally-defined tops in that sample is moderated by sampling region.

H9: The effect of the proportion of GI men in a sample on the proportion of behaviourally-defined bottoms in that sample is moderated by sampling region.

3 Methods

Data Collection

Search strategy

For this systematic review only electronic databases were searched. Databases used were *Google Scholar* and *China National Knowledge Infrastructure (CNKI)*.

Stage 1

This stage involved only searching *Google Scholar* for sources. As it was known before commencing the study that this area was understudied, no restrictions were placed apart from articles containing information on TBV_X-distributions. Key words used in the search at this stage included: ("top", "bottom", "versatile"); ("homosexual", "role", "anal"); ("insertive", "receptive", "gay"), ("top", "sex", "role"), ("activo", "pasivo", "versatil"). Relevance of articles was judged on the basis of title and a reading the abstract.

Stage 2

This stage involved screening relevant articles. Articles were screened in their entirety, as this would allow for gaining a better overview of what was available in this understudied field already. At this stage, it was found that various authors cited articles that, whilst available on *Google Scholar*, only contained a link to *CNKI*. This website was consequently searched for relevant articles as well, but the search protocol was different. Rather than using the search function, the relevant articles obtained from the English resources were used as a seed, whereby cited sources were systematically checked. What made this possible was the structure of *CNKI* that included direct links to each of the cited articles with the abstract of most these being in English.

Data extraction (articles written in English)

The following data was extracted from articles selected for inclusion in the analysis: (1) name of the authors; (2) publication year; (3) sampling period; (4) sampling country; (5) target population; (6) study base; (7) sampling method; (8) sample size; (9) mean age of the sample or median age of the sample and if available, range or IQR; (10) number of tops, bottoms and/or versatiles in the sample; group ∅ containing all people not using these labels including corrections for missing values; (11) number of men identifying as gay, bisexual and/or heterosexual in the sample; group ∅ containing all people not using these labels including corrections for missing values. In a follow-up stages the following information was also extracted: (12) measure used and response options; (13) e-mail address of the author.

Data extraction (articles written in Chinese)

The following procedure was used for extracting information from articles written in Chinese: (1) the abstract of the article was read and screened for relevance. Only articles containing an abstract written in English were included; (2) relevant information was extracted from the abstract where possible; (3) articles were searched

for relevant information, which was made possible by the systematic structure of the articles and the fact that numbers are easily separated from ordinary characters, e.g. a year "2010" could be located without being able to understand the text. Moreover, as Chinese uses the designations '0', '0.5' and '1' instead of 'bottom', 'versatile' and 'top' respectively, this information could be located; (4) articles were partially translated using *Google Translate* to help locate relevant information; (5) annotations were made on a PDF-version of the document. This included high-lighting relevant texts and including the translations from Google Translate. Questions were also included as notes with the PDF-file; (6) the PDF-file including annotations was sent to a native speaker of Chinese, who verified the information and clarified information; (7) articles were included only if no ambiguities remained after consulting the native Chinese speaker.

Determining what measure was used

When it could not be determined if a behavioural, preference-based or sexual identity measure was used, the author listed as the corresponding author was contacted to request information on the particular item used.

Study selection (research questions R1–R5)

Answering R1–R5 places different demands on what information an article must contain than R6–R9, e.g. the mean age of the sample is not required. Studies were included for use in the analysis if the sample used in the study: (1) was taken from a population of gay men, gay and bisexual men, or MSM; (2) included information about the number of tops, bottoms and versatiles in the sample; (3) either contained explicit information about what measure was used or it was possible to extract this information from the content of the article, e.g. from sentences such as "participants were asked about their preferences regarding anal sex"; (4) the category 'other' comprised strictly less than 1% of the sample size; (5) no study using the same sample or some subset of the sample was included for use in the analysis already. A PRISMA flow diagram of the collection of relevant articles is included in Appendix A (Figure 1).

Study selection (research questions R6–R9)

Answering R6–R9 requires data on the TBV_X-distribution, sexual orientation, sampling period, mean age of the sample, and sampling region. Studies were included for use in the analysis if the following criteria were met: (1) the sample used in the study was taken from a population of MSM; (2) the study included information about the number of tops, bottoms and versatiles in the sample; (3) a behavioural measure was used to determine the number of behavioural tops, behavioural bottoms and behavioural versatiles in the sample; (4) the number of men in the sample identifying as gay can be discerned from the article; (5) either the mean age of the sample or median age and the range/IQR of the sample is reported; (6) sampling took place in Asia, Latin America, or countries traditionally regarded as 'the West' (North America, Europe, Australia); (7) the study does not use some subset of a sample used in another study that has been included for use in the analysis already. If two or more different studies use the same sample or a subset of the same sample, using the data from the study that provides the most accurate information on the variables included in the regression analysis was preferred; (8) the category 'other' comprised strictly less than 10% of the sample size; (9) if no

information was included about the sampling period but the article contained a note indicating the date of first submission, this information was recorded.

Studies about very specific sub-populations of MSM were excluded from the analysis, e.g. meta-amphetamine using MSM. For the purpose of this study men who have sex with men exclusively (MSME), men who have sex with men only (MSMO), men who have sex with men and women (MSMW) and young men who have sex with men (YMSM) were all regarded as MSM and were not excluded from the analysis. A PRISMA flow diagram of the collection of relevant articles for answering R6–R9 is included in Appendix A (Figure 2).

Materials

Data was analysed using IBM SPSS Statistics 26. Statistical power was determined using GPower version 3.1.9.6 (Faul et al., 2020).

Methods for answering R1–R5

Studies were separated on the basis of what type of measure X was used. As before $X = P, B$ or I where 'P' stands for a preference-based measure, 'B' stands for a behavioural measure, and 'I' stands for an identity-based measure. Let $S_{X,r}^i$ denote the sample used in study i, obtained in region r and where a measure of type X was used. Here $r = 1, 2$ or 3 where '1' stands for 'Asia', '2' stands for 'the West' and '3' stands for 'Latin America'. Other regions were not considered because not enough data was available.

Let $T_{X,r}^i, B_{X,r}^i, V_{X,r}^i$ and $O_{X,r}^i$ denote the number of people in the categories top$_X$, bottom$_X$, versatile$_X$ and other$_X$ in sample $S_{X,r}^i$. The subscript 'X' on the terms 'top', 'bottom', 'versatile' and 'other' indicates that these terms are defined according to the measure type X used in the study (which may differ from how these terms are defined in, say, the introduction of the paper). Also, let $N_{X,j}$ denote the total number of studies using measure type X where sampling took place in region j, e.g. $N_{P,1}$ denotes the total number of studies using a preference-based measure where sampling took place in region 1 (Asia) and $N_{B,3}$ denotes the total number of studies using a behavioural measure where sampling took place in region 3 (Latin America).

Because of the imposed inclusion-exclusion criteria, $O_{X,r}^i$ constitutes strictly less than 1% of the participants in the samples $S_{X,r}^i$ used for answering R1–R5. For this reason, and because this is a catch-all category containing people who do not use the terms 'top', 'bottom' and/or 'versatile', people who use different terms, people who did not answer the relevant questions, et cetera, this category was disregarded for the purpose of comparing distributions. That is, each sample $S_{X,r}^i$ was treated as if $O_{X,r}^i = 0$ and the sample size was accordingly taken to be $N_{X,r}^i = T_{X,r}^i + B_{X,r}^i + V_{X,r}^i$.

The percentage of tops$_X$, bottoms$_X$ and versatiles$_X$ in sample $S^i_{X,r}$, denoted $t^i_{X,r}$, $b^i_{X,r}$ and $v^i_{X,r}$ respectively, was calculated as:

$$t^i_{X,r} = \frac{T^i_{X,r}}{T^i_{X,r} + B^i_{X,r} + V^i_{X,r}} \times 100\%, \quad b^i_{X,r} = \frac{B^i_{X,r}}{T^i_{X,r} + B^i_{X,r} + V^i_{X,r}} \times 100\%, \quad v^i_{X,r} = \frac{V^i_{X,r}}{T^i_{X,r} + B^i_{X,r} + V^i_{X,r}} \times 100\%$$

The average TBV$_{X,r}$-distribution (expressed in percentages) for studies using measure type X where sampling took place in region 1 was calculated from the percentage distributions of the individual samples as follows:

$$t_{X,1} = \frac{\sum_{i=1}^{N_{X,1}} t^i_{X,1}}{\sum_{i=1}^{N_{X,1}} (t^i_{X,1} + b^i_{X,1} + v^i_{X,1})}, \quad b_{X,1} = \frac{\sum_{i=1}^{N_{X,1}} b^i_{X,1}}{\sum_{i=1}^{N_{X,1}} (t^i_{X,1} + b^i_{X,1} + v^i_{X,1})}, \quad v_{X,1} = \frac{\sum_{i=1}^{N_{X,1}} v^i_{X,1}}{\sum_{i=1}^{N_{X,1}} (t^i_{X,1} + b^i_{X,1} + v^i_{X,1})}$$

Note that in this way the *normalised* average TBV$_X$-distribution of region 1 is obtained, i.e. the obtained distribution is such that $t_{X,1} + b_{X,1} + v_{X,1} = 100\%$, where $t_{X,1}$ denotes the percentage of tops$_X$, $b_{X,1}$ denotes the percentage of bottoms$_X$ and $v_{X,1}$ denotes the percentage of versatiles$_X$ in the normalised average TBV$_X$-distribution of region 1. Moreover, it is correct to speak of this being an *average* distribution despite the fact that no division by $N_{X,1}$ takes place because dividing each of $t_{X,1}$, $b_{X,1}$, $v_{X,1}$ by $N_{X,1}$ does not change how $t_{X,1}$, $b_{X,1}$, $v_{X,1}$ relate to one another and would be undone when normalising. For brevity I simply refer to this as the average TBV$_X$-distribution of region 1 instead of the normalised average distribution. Also note that the superscript i is dropped because it concerns a regional average and no longer a particular sample i. The average TBV$_{X,r}$-distributions for the other regions were obtained analogously.

Finally, the percentage of tops$_X$, bottoms$_X$, and versatiles$_X$ in the normalised average global TBV$_X$- distribution were calculated from the different TBV$_{X,r}$-distributions of the three regions as follows:

$$t_X = \frac{\sum_{j=1}^{3} t_{X,j}}{\sum_{j=1}^{3} (t_{X,j} + b_{X,j} + v_{X,j})}, \quad b_X = \frac{\sum_{j=1}^{3} b_{X,j}}{\sum_{j=1}^{3} (t_{X,j} + b_{X,j} + v_{X,j})}, \quad v_X = \frac{\sum_{j=1}^{3} v_{X,j}}{\sum_{j=1}^{3} (t_{X,j} + v_{X,j} + v_{X,j})}$$

Using these methods the following distributions are expressed in terms of the percentage of tops$_X$, bottoms$_X$ and versatiles$_X$ in the sample: (a) the TBV$^i_{X,r}$-distribution for every sample $S^i_{X,r}$ meeting the inclusion-criteria for answering R1–R5; (b) the average TBV$_{X,r}$-distribution for all regions r and all measure types X; (c) the average (global) TBV$_X$-distribution for all measure types X. The calculated percentage-distributions are included in Appendix B4.

Determining what these distributions are answers R1. To answer R2–R4, the appropriate distributions are compared using χ^2-tests; to answer R5, the distributions obtained in samples where the same measure type was used are pair-wise compared to each other and the number of cases where the two distributions statistically differ from one another is determined. Dividing this number by the total number of comparisons made this

gives an estimate of the probability that the $\text{TBV}_{X,r}^{i}$-distribution of sample $S_{X,r}^{i}$ differs from the $\text{TBV}_{X,r}^{j}$-distribution of another sample $S_{X,r}^{i}$ with $j \neq i$ whenever two samples are compared.

Methods for answering R6–R9

Dependent variable

The dependent variable for answering R6 and R8 is the proportion of tops$_X$ in the sample, calculated as the number of tops$_X$ divided by the sample size. The dependent variable for answering R7 and R9 is the proportion of bottoms$_X$ in the sample, calculated as the number of bottoms$_X$ divided by the sample size. Here the sample size includes the category 'other' as in most cases the mean age and/or the median age of the sample are calculated by including this category.

Both the proportion of tops$_X$ and the proportion of bottoms$_X$ are fractions, meaning they must fall within the $[0,1]$ range. Because a fraction is not a binary variable, Probit and Logit are not suitable. Limited dependent variable regression was considered, but ultimately standard multivariate linear regression was opted for as this method mostly becomes problematic in dealing with fractions when values around 0 and around 1 are common (Stock & Watson, 2012; Wooldridge, 2016), which is not the case here as the proportion of tops$_X$ and the proportion of bottoms$_X$ is never close to 0 or 1.

Independent variable: fraction of gay-identified people in the sample

The primary independent variable is *the fraction of gay-identified people in the sample* is calculated as the total number of gay-identified people in the sample divided by the total sample size.

Choice of control variables

As this area remains understudied it is difficult to gather sufficient data for making regression analysis possible. However, many published papers include the sampling period, sampling location, and the mean/median age of the sample. Hence the choice for these control variables is partially motivated by practical reasons.

Control variable 1: sampling year

Societal attitudes concerning homosexuality are not constant but have changed over time, becoming more positive in all three regions (Asia, the West and Latin America) considered (Yang, 1997; Parker & Bhugra, 2000; Loftus, 2001; Gerhards, 2010; Smith et al., 2014; Bowman et al., 2014; Schnabel, 2016; Cheng et al., 2016; Kite & Bryant-Lees, 2016). As mentioned, this affects homosexual identity formation because homophobia in society is internalised (Rowen & Malcolm, 2003; Brown & Trevethan, 2010) and as such the sampling year relates directly to the primary independent variable, but also to both dependent variables as Hart et al. (2003) and Gil (2007) both found tops and bottoms to differ with respect to internalised homophobia. Moreover, both the role and nature of hegemonic masculinity are not constant in time either (Ricciardelli et al., 2010; Hidaka, 2010; French & Rothery, 2011; Connell, 2012; Hearn et al., 2012; Connell, 2016), nor

are gender norms (Seguino, 2007) so that the sampling period is also plausibly correlated with the proportion of tops$_X$ and the proportion of bottoms$_X$ as this bears directly upon the attitudes towards IAI and/or RAI and being a behaviourally-defined top or bottom.

The sampling year is calculated on the basis of the reported sampling period. If sampling took place within a single year, this year is taken as the sampling year. In case the sampling period spanned several years, the sampling year was taken to be the average of the year in which the sampling started and the year in which the sampling ended. In case no information on the sampling period was reported, the year the paper was first submitted was used instead if available; the publication year was used if this was not available. These changes were made manually. Various authors were emailed to enquire about missing sampling periods, but no response was received.

Control variable: sample mean age

The variable *sample mean age* refers to the mean age of the people constituting the sample. Age has been linked to identifying as top, with older people being more likely to identify as top than younger people in Chinese samples (Zheng & Zheng, 2009; Zheng et al., 2012; Zhou et al., 2013; Xu & Zheng, 2018), but also in an international sample (Swift-Gallant et al., 2018) and a Western sample (Moskowitz et al., 2008) which may have something to do with the power dynamic between sexual partners (Ho & Tsang, 2000; Kippax & Smith, 2001; Choi et al., 2003; Carballo-Diéguez et al., 2004; Hoppe, 2011; Arrington-Sanders et al., 2013; Xu & Zheng, 2018). Regardless of the reason, as age also relates to homosexual identity formation (Troiden, 1989) and identifying as gay (Bybee et al., 2009; Floyd & Bakeman, 2006), age must be controlled for.

Most studies reported the mean age of the sample, but some studies reported the median and the range or IQR. In this case methods described in Hozo et al. (2005), Wan et al. (2014) and Luo et al. (2018) were used to estimate the mean age from the median and the range/IQR. The methods described in Wan et al. (2014) are supposed to be an improvement of those in Hozo et al. (2005), and similarly, the formula used in Luo et al. (2018) is supposed to be an improvement of the ones reported in Wan et al. (2014). All three methods were used and the results and it was checked if using a different method affected the analysis. Although effect size and *p*-values changed slightly this did not lead to different decisions regarding the hypotheses.

Moderator: region

Region was determined on the basis of the sampling country, where a distinction is made between countries falling in what is traditionally considered to be *the West*. This concerns in particular the USA, Canada, Australia, and (western) European countries. China and other Asian countries are collapsed into the category *Asia*, and Latin American countries, in particular Peru, constitute the countries falling within the category *Latin America*. Other regions were not considered because sufficient data was not available.

Regionally codes were assigned manually, whereby Asia was coded as '0', Latin America was coded as '1' and the West was coded as '2'. After list-wise deletion was used to remove all studies that had one or more missing values for the variables included in the multivariate regression analysis, the West was dropped from consideration as only two studies remained. In effect this turned the variable 'region' into a dummy variable.

4 Data

Research questions R1–R5

A total of 172 different studies containing information on TBV-subgroups was collected. Of these 172 studies, 11 studies using an identity-based measure, 9 using a preference-based measure, and 37 using a behavioural measure met the inclusion criteria for answering R1–R5. An overview of these studies is given in Appendix B1–B3.

Research questions R6–R9

Of the 172 different studies, 109 did not include information on all the variables required for answering research questions R6–R9. 2 or were otherwise unsuitable for use in the analysis. A total of 43 studies used a behavioural measure and included data on all the relevant variables. Seven more studies were excluded from the analysis because they did not contain enough information to calculate the mean age of the sample from the median and range/IQR using the three different estimation methods. Although these studies did include information on the median that could be used to estimate the mean using formulas from Hozo et al. (2005), other estimation methods – Wan et al. (2014) and Luo et al. (2018) – did require either the range or the IQR too (for samples larger than 25, Hozo et al. (2005) recommends taking the median). As these other methods are improved versions of Hozo et al.'s (2005)'s method, it was opted to include only studies where the mean age could be estimated using these more accurate methods.

A total of 36 studies remained for use in the regression analysis (26 from Asia, 2 from the West, and 8 from Latin America). An overview of these studies is given in Table 5 in Appendix C. Summary statistics and descriptions of the variables used in the analysis are included in Table 6 in Appendix D.

5 Results

Research questions R1–R5

Research question 1

The average TBV_I-, TBV_P- and TBV_B-distributions are included in Table 4 in Appendix B4.

Research question 2

The average TBV_I-, TBV_P- and TBV_B-distributions included in Table 4 in Appendix B4 were pairwise compared. There was no statistically significant difference between the average TBV_I-distribution and the average TBV_P-distribution, $\chi^2(2, N = 200) = .78, p = .677$. There was also no statistically significant difference between the average TBV_I-distribution and the average TBV_B-distribution, $\chi^2(2, N = 200) = 4.72, p = .094$. Finally, the average TBV_P-distribution and the average TBV_B-distribution did also not significantly differ from one another, $\chi^2(2, N = 200) = 2.93, p = .231$. Therefore, H2 is not rejected.

Research question 3

The results of pairwise comparisons are included in Table 10 in Appendix E2. The TBV_I-distributions do not differ significantly between the regions, and neither do the TBV_P-distributions. The TBV_B-distribution for the West differs significantly from the TBV_B-distribution for Asia, $\chi^2(2, N = 200) = 6.54, p = .038$. The TBV_B-distribution for the West also significantly differs from the TBV_B-distribution for Latin America, $\chi^2(2, N = 200) = 13.7, p = .001$. Thus H3 is rejected for identity-based measures and preference-based measures, but not for behavioural measures, although not all distributions differ from one another.

Research question 4

The results of comparing the average regional TBV_I-, TBV_P- and TBV_B-distributions to the average global TBV_I-, TBV_P- and TBV_B-distributions are included in Table 10 in Appendix E4. There are no significant differences between the regional distributions and the global distributions, hence H4 is rejected.

Research question 5

A total of 55 chi-square two sample tests was carried out to pairwise compare the 11 TBV_I-distributions (Table 7, Appendix E1). On 31 occasions, the compared distributions were not statistically different at the 5% significance level. Hence the probability that a TBV_I-distribution in a given sample does not significantly differ from the TBV_I-distribution in another sample obtained is estimated to be $31/55 \approx .56$.

Similarly, a total of 36 chi-square two sample tests was carried out to pairwise compare the 9 TBV_P-distributions (Table 7, Appendix E2). On 10 occasions, the compared distributions were not statistically different at the 5% significance level. The probability that a TBV_P-distribution in a given sample does not significantly differ from the TBV_P-distribution in another sample is therefore estimated to be $10/36 \approx .28$.

Finally, a total of 666 chi-square two sample tests was carried out to pairwise compare the 37 TBV_B-distributions (Table 7, Appendix E3). On 252 occasions, the compared distributions were not statistically different at the 5% significance level. The probability that a TBV_B-distribution in a given sample does not significantly differ from the TBV_B-distribution in another sample is therefore estimated to be $252/666 \approx .38$.

Research questions R6–R9

An overview of the results from the regression analysis is included in Table 11 in Appendix F.

Specification 1T
In Specification 1T, the proportion of gay-identified men in the sample did not significantly predict the proportion of tops, $b = -.008, t(32) = -.013, p = .941$. The sampling period did significantly predict the proportion of tops, $b = -.011, t(32) = -2.65, p = .013$. Together, sampling period, mean age of the sample, and proportion of gay-identified men in the sample did not explain a significant proportion of variance in the proportion of tops, $R^2 = .21, F(3, 32) = 2.77, p = .057$. Running the same regression with A_W slightly changed the effect-size, but sampling period remained statistically significant, $b = -.011, t(32) = -2.65, p = .012$ whilst explained variance remained insignificant, $R^2 = .21, F(3, 32) = 2.79, p = .056$. Using A_H did not alter the findings either, $b = -.011, t(32) = -2.64, p = .013$ and $R^2 = .21, F(3, 32) = 2.78, p = .057$.

Specification 1B
In specification 1B, the proportion of gay-identified men in the sample did not significantly predict the proportion of bottoms, $b = .079, t(32) = .70, p = .491$. Mean age of the sample did significantly predict the proportion of bottoms, $b = -.013, t(32) = -3.20, p = .003$. Together, sampling period, mean age of the sample, and proportion of gay-identified men in the sample explained a significant proportion of variance in the proportion of bottoms $R^2 = .25, F(3, 32) = 4.50, p = .027$. Running the same regression with A_W again had a small effect on the effect-size, but mean age remained statistically significant, $b = -.014, t(32) = -3.45, p = .002$. Explained variance also remained statistically significant, $R^2 = .28, F(3, 32) = 4.06, p = .015$. Using A_H also did not alter findings $b = -.013, t(32) = -2.94, p = .006$ and $R^2 = .22, F(3, 32) = 2.97, p = .046$.

Given that using A_W or A_H instead of A_L only had a minor effect on the effect size but did not lead to different conclusions regarding the statistical significant of results, and that A_L is supposed to be an improvement of A_W and A_H it was decided to only use A_L for any further regression analysis.

Specification 2T
In Specification 2T an interaction term was added to examine if there are differences between regions. No variables were dropped from the regression as statistical insignificance may be precisely due to opposing regional effects (Stock & Watson, 2012; Wooldridge, 2016). As only two studies were conducted in the West

these were excluded from further consideration and focus shifted towards examining potential differences between Asia and Latin America only. In effect this turned R_C into a dummy variable.

Dropping the West from further consideration did not affect earlier findings, as in the combined sample of Asia and Latin America the proportion of gay-identified men in the sample did still not significantly predict the proportion of tops, $b = .014, t(30) = .12, p = .909$. Sampling year remained a significant predictor for the proportion of tops, $b = -.012, t(30) = -2.47, p = .020$; mean age of the sample remained statistically insignificant, $b = -.001, t(30) = -.33, p = .745$.

After adding the interaction term the proportion of gay-identified men G_p remained statistically insignificant, $b = -.002, t(29) = -.018, p = .986$. Mean age of the sample also remained statistically insignificant. The sampling period on the other hand remained a significant predictor of the proportion of tops in the sample, $b = -.011, t(29) = -2.52, p = .017$.

Specification 2B

The same procedure as in Specification 2T was repeated for bottoms. Here too the proportion of gay-identified men in the sample remained statistically insignificant, $b = .063, t(29) = .473, p = .640$. The mean age of the sample remained a significant predictor for the proportion of bottoms, $b = -.012, t(29) = -2.59, p = .015$.

At this stage it seems clear that the proportion of gay-identified men does neither significantly predict the proportion of tops in the sample nor does it significantly predict the proportion of bottoms in the sample. As such, we can reject H6 and H7. Consequently, we also reject H8 and H9.

Specification 3T/3B

Although not central to this paper I decided to drop all the statistically insignificant results from the regression equation and see if there was any moderation effect of region. Doing so showed that the sampling year significantly predicts the proportion of tops in the sample, $b = -.010, t(31) = -2.53, p = .017$ and that this effect is moderated by region, $b = -.008, t(31) = -2.15, p = .039$. Moreover, the mean age of the sample significantly predicts the proportion of bottoms in the sample, $b = -.013, t(31) = -2.94, p = .006$ and this effect too is moderated by region, $b = .003, t(31) = 2.11, p = .043$.

6 Discussion

This paper was the first to systematically analyse TBV_X-distributions. Whilst the findings are novel, much care is needed in interpreting the results.

For answering R1 a systematic overview of studies where TBV_X-distributions were recorded was created (Appendix B1–B3). Furthermore, the global average TBV_I-distribution, the global average TBV_P-distribution, and the global average TBV_B-distribution were calculated (Appendix B4). These distributions should not be interpreted as approximating the TBV_X-distributions in some regional or global population of MSM. I have already touched upon issues of non-probability sampling making this problematic, here I touch upon further issues. Firstly, the average distributions are calculated from a relatively limited number of samples, especially in case of identity-based and preference-based measures (11 and 9 respectively). Secondly, a large percentage of the samples used to calculate the global average TBV_I-, TBV_P-, and TBV_B-distributions were obtained in China (45.5%, 55.6%, and 58.3% respectively). As such the designation 'global' may be somewhat misleading. This cannot be solved by using weights since then, because of the small number of studies available, the global average distribution becomes too dependent upon the distributions in particular samples. A possible resolution is to collect more data, which is a recommendation for future research. In particular it seems fruitful to look at more studies published in Latin America, something I could not feasibly do because of unfamiliarity with the Spanish language. Regardless of these issues, however, the average distributions in Table 4 in Appendix B4 are a positive addition to the current body of knowledge which is lacking an overview combining the results of multiple studies and leads several researchers to adopt a practice of comparing the TBV_X-distribution in their sample to the distributions obtained in only one or two other studies.

That there are no good grounds for such a practice follows from the answer to R5. Indeed, the answer and the probabilities calculated in R5 must only be seen as relevant to assessing such research practice. It shows that if one determines the TBV_X-distribution in one's sample, and one compares it to the TBV_X-distribution in the sample of another randomly chosen study that records this information, the probability that these two distributions do not statistically differ is .56 in case an identity-based measure is used, .28 in case a preference-based measure is used, and .38 in case a behavioural measure is used. This means that on average, a TBV_P-distribution in a given sample is more likely to differ from the TBV_P-distribution in other samples than not, and similarly for the TBV_B-distribution in a given sample. Moreover, there is no real added value to the practice adhered to by some researchers to compare the distribution in their sample to the distribution in one or two other samples; there is always some chance that the distributions do not statistically differ and some chance that they do. However, even if the distributions in the two samples are not statistically different one should not conclude from this that one's sample is somehow representative of the underlying population (as is usually done), for Appendix E1–E3 show that it is possible that two distributions do not statistically differ from one another, but differ from both the global and the regional average distributions, e.g. in Table 9

(Appendix E3) the distributions of study 2 and study 6 do not statistically differ from one another, but both differ from the global and regional average distributions. Whilst the global and regional average distributions need not be representative of the underlying population either, they do provide a way for researchers to see if the distribution in their sample is similar or dissimilar from samples obtained in a number of other studies, which is an improvement from merely making a comparison with one or two other studies.

Although H2 was not rejected, this must not be taken to mean that it does not matter what type of measure is used. Firstly, the global average distributions for preference-based and identity-based measures were calculated from only a limited number of distributions. Moreover, there are other reasons for not considering these different measures types as equally suitable. For instance, whilst identity may be relatively constant and independent of situational factors, what anal intercourse behaviour someone engages in (and hence whether someone is a behavioural top, bottom or versatile) may depend much more on situational factors than one's preference or identity, e.g. the HIV status of oneself and/or ones' sexual partner (Wolitski et al., 1998; Van de Ven et al., 2002; Lyons et al., 2013) or the power dynamic between (available) sexual partners (Ho & Tsang, 2000; Kippax & Smith, 2001; Dangerfield et al., 2018).

With regards to R3 and R4 it is noteworthy that only the West differed from other regions, and that the proportion of versatiles in the average regional distributions was higher in the West than in other regions. This confirms the findings of some other studies that did not meet the inclusion criteria, namely that versatility seems to be higher amongst western men (Jameson et al., 2010; Lyons et al., 2011; Lyons et al., 2013), certainly as compared to Latin America where traditionally there is a dichotomy between 'activos' and 'pasivos' (Carrier, 1977; Almaguer, 1993; Morales, 1995; Peinado et al., 2007; Cantú, 2009). Indeed the Mexican term for a versatile, 'internacional', which seems in use since around the year 2000 (Cantu, 2000; Rodriguez, 2019), reflects this non-indigenous nature. That the average regional distributions for the West did not differ from the average global distributions despite differing from the other average regional distributions is possibly due to the fact that the average global distributions are calculated from (amongst others) the average regional distributions from the West. Given the relatively small number of studies conducted in the West, this may nevertheless have affected the average global distributions enough for these to not be significantly different from the average regional Western distributions.

The findings that the proportion of GI men does not significantly predict the proportion of tops in the sample nor the proportion of bottoms in the sample is surprising and goes against the findings of Hart et al. (2003). One of the reasons may be low statistical power. Using Gpower (Faul et al., 2020) the achieved statistical power in Specification 1T was calculated to be only .05, and .12 in Specification 1B. Another possibility is that the MSM included in the studies used for the analysis differ from other MSM in that they are more connected to the gay/MSM-community. Difficulties in sampling MSM are well-known, and the MSM attending MSM-venues where much of the samples used in this study were collected are known to differ from other MSM in many regards (McKirnan et al., 1995; Evans et al., 2007; Zhang et al., 2008; Paz-Bailey et al., 2013). In

particular, these MSM may have felt less restricted leading to a weaker association between being NGI/GI and applying the term 'top'/'bottom' to oneself. However, this remains speculation and there are many other possible reasons, e.g. it is also possible that no effect was found because different studies all used different measures. Indeed, I have grouped the studies by measure *type*, but within each type there were variations, e.g. some studies considered behaviour in the last 6 months, others in the last 9 months, et cetera. Indeed, some studies following gay men over longer periods of time have noticed that when longer time-spans are considered men are more likely to be behaviourally versatile (van Griensven et al., 1987; Lyons et al., 2011; Dangerfield et al., 2018), which may have distorted the results.

Whilst it is not the central focus of this paper, the findings of Specification 3T that sampling year has a negative effect on the proportion of tops in the sample is as we would expect: over time, attitudes towards homosexuality have improved (Cheng et al., 2016; Kite & Bryant-Lees, 2016) and hegemonic masculinity has declined over time (Ricciardelli et al., 2010; Hidaka, 2010; Hearn et al., 2012), hence men may feel less inhibited to having RAI which is viewed more negatively in more hegemonic masculine societies and societies where homosexuality is viewed more negatively.

The findings in Specification 3B are also in accordance with what is expected. The mean age of the sample has a negative effect on the proportion of bottoms in the sample, i.e. as the people in the sample get older, the proportion of bottoms decreases. This has indeed been observed in China and Western samples (Zheng et al., 2012; Swift-Gallant et al., 2018; Moskowitz et al., 2008). The fact that is a positive moderator effect of region (which in this case means: Latin America), is in accordance with the findings of Peinado et al. (2007) who found that 'pasivos' generally are older than 'activos' in Peru, where most of the Latin American samples included in the analysis were obtained.

Although this paper makes some novel contributions to the literature, findings can be improved upon in several ways. Firstly, it seems possible to increase the number of studies included in the analysis (both for answering R1–R5 and for answering R6–R9). This especially concerns publications in Spanish and Chinese. Although I included a number of articles written in Chinese articles, finding these articles could not be done in an efficient manner and I potentially missed many. The same holds true for articles written in Spanish.

Secondly, ideally one does not analyse proportions as I have done, but use statistical methods that allow for predicting TBV_X-distributions. However, statistical techniques allowing for the dependent variable to be a vector (a TBV_X-distribution is a three-component vector). Although there do seem to be ways to use regression techniques to analyse such data (Manca, 2013; Mok & Iz, 2014), these techniques are not developed for use in the social sciences and employing these methods was not deemed feasible. However, future research may benefit from looking into such methods.

Thirdly, future research may benefit from reconsidering how to deal with the category 'other'. In this paper I focused only on TBV_X-distributions and disregarded all studies where the category 'other' contained more than 1% of the sample size, as there were large variations between studies in who was included in this category or what exactly it stood for. A less conservative approach may be recommendable. In particular, this is because now studies where the researcher(s) did not mention the category 'other' and/or did not report how many people this category included, the study was included in the analysis if it met all the other criteria, whereas had the researcher reported how many people fell in the category 'other', the study may have been excluded if this number exceeded 1% of the total sample size even if the study met all the other inclusion criteria. Future research may benefit from contacting various authors so that how to deal with studies where the category 'other' contains a larger number of people can be decided on a case by case basis basis. Although I did contact various authors I either received no reply or did not receive a reply in a timely manner so that for this study this could not be done.

Fourthly, the results can be improved upon greatly if data-sets containing information on individuals could be combined and used. This holds true in particular for answering R6–R9. Whilst there still would be some difficulties in that the measures used may not exactly be the same (even if the measure type is the same), if a sufficient number of data-sets can be combined this could give many more insights than the present study where I had to carry out the regression analysis using summary statistics such as the sample mean, and the proportion of GI men in the sample.

In order to improve research involving gay men and MSM making carrying out such research is very important. Whilst the current research can certainly be improved upon, merely presenting an overview of the TBV_X-distributions is telling in itself: there are large variations in the TBV_X-distributions in the samples used in different studies. This is made more clear in Table 7, Table 8 and Table 9 in Appendix E1, E2 and E3 respectively. If one does not gather data on TBV_X-distributions, as is common practice in much of the research involving gay men and/or MSM, one cannot possibly know if one's sample is somehow similar to the other samples used in similar research. Because much of the research involving these populations relies on non-probability samples, it is perhaps not feasible to know if the sample is representative of the population from which it was drawn, and so knowing if it does or does not differ significantly from other samples used in studying this underlying population whilst not ideal, is still much better than not knowing this and not ever being in a position to know this (because the data was not gathered and cannot retro-actively be gathered). This study shows that whilst it is certainly possible that, per chance, one's sample is similar to other samples, it is also possible, per chance, one's sample differ significantly from most other samples. In the latter case, the findings one obtains may be misleading but are also potentially harmful if this research is used as a basis for, say, informing policy measures aimed at gay men. For instance, if one uses a sample where the proportion of bottoms is unusually high but one does not know this because one did not ask about IPI, IPP, or IPB, one may conclude that internalised homophobia is less of a problem than it really is as internalised homophobia

was found to be lower for bottoms than tops (Hart et al., 2003; Gil, 2007). Similarly, if using a sample where the proportion of tops is unusually high, one may find risky sexual behaviour to be less of a problem than it really is, as various studies found such behaviour to be more prevalent amongst bottoms than amongst tops (Moskowitz et al., 2008; Meng et al., 2015; Zeng et al., 2016; Lei et al., 2018).

Recognising the potential seriousness of this problem it becomes important to focus on understanding what the terms 'top', 'bottom' and 'versatile' mean to the people who use them and, moreover, to develop more accurate measures or, at the very least, to make sure that measures are used correctly. As noted by Donovan (1993), if the criteria for inclusion in a certain population are unclear or different researchers all use different criteria, comparative research is confounded and cumulative understanding is inhibited. Currently, such is the situation when it comes to the terms 'top', 'bottom', and 'versatile' and the corresponding subgroups.

Whilst some authors define these terms consistently in terms of either preferences, behaviour, or as an identity, others mix definitions and, moreover, present research participants with such mixed definitions before asking which of these terms applies to them which may lead to measurement error and biased results that are potentially harmful because they are misleading to policymakers, researchers and/or (health) practitioners (Harry, 1986; Ramirez-Valles et al., 2005; Meyer & Wilson, 2009). Thus, similarly to how not controlling for these subgroups may be harmful, controlling for subgroups may be harmful if not done correctly.

7 Conclusion

Previous studies leave little doubt that tops, bottoms and versatiles (irrespective of whether these are behavioural-defined, preference-defined, or identity-defined) differ significantly in a number of aspects – e.g. the degree of internalised homophobia – that directly or indirectly affect the lives of gay men or MSM more generally by making certain (negative) outcomes more or less likely. Despite of this, only a few researchers have focused explicitly on studying these subgroups and none have systematically looked at the distribution of these subgroups in different samples.

This study explicitly considered such distributions leaving little doubt about the fact that whilst the distribution of tops, bottoms, and versatiles in a sample may well be similar to the distributions in other samples, it also may be very different. In some cases the distribution of tops, bottoms and versatiles in a sample differed neither from the average regional and global distributions, nor from (most of) the individual distributions in other samples. In other cases the distribution in a particular sample differed from both the global and regional average distributions, and from the distributions in (most) other samples.

Together, the findings over others that there seem to be significant differences between tops, bottoms and versatiles and the findings of this study that the distribution of tops, bottoms and versatiles more often than not differs between samples. Regardless, if participants are not asked about IPP, IPB or IPI there is no way to know what the distribution of tops, bottoms and versatiles in a given sample is, and hence no way to know if this sample is very different from other samples of not. Not only does this make comparing between studies difficult if not impossible, it also threatens to undermine research findings (which might have been very different if tops, bottoms and versatiles had been distributed differently in the sample used), it is also potentially very harmful to society if these research findings are used to inform policy for instance.

There is only one way to eliminate this risk, namely to control for these subgroups, which future research therefore must do. However, first accurate measures must be developed, for using inaccurate measures or using inconsistent measures creates much of the same problems as not controlling for these subgroups: it carries the risk of misrepresenting and undermining research findings and creates a false sense of studies being comparable and inhibits cumulative understanding. There thus is a need for (more) qualitative research in order to better understand what the terms 'top', 'bottom' and 'versatile' mean to the people using them and what they conceptualise, which may differ between societies.

References

Abrams, D. Psychology of social identity. In *International Encyclopedia of the Social & Behavioral Sciences*, pages 14306–14309. Elsevier Science and Technology, 2001.

Abramson, P. R. and Herdt, G. The assessment of sexual practices relevant to the transmission of AIDS: A global perspective. *Journal of Sex Research*, 27(2):215–232, 1990.

Acaba, J. Meanings, preferences, and power among men having sex with men in Manila. *Philippine Social Sciences Review*, 65(2):101–121, 2013.

Adelman, M. Stigma, gay lifestyles, and adjustment to aging: A study of later-life gay men and lesbians. *Journal of Homosexuality*, 20(3-4):7–32, 1991.

Almaguer, T. Chicano men: A cartography of homosexual identity and behavior. *The lesbian and gay studies reader*, pages 255–73, 1993.

Altman, D. et al. Globalization and the international gay/lesbian movement. *Handbook of lesbian and gay studies*, pages 415–425, 2002.

Arrington-Sanders, R., Leonard, L., Brooks, D., Celentano, D., and Ellen, J. Older partner selection in young African-American men who have sex with men. *Journal of Adolescent Health*, 52(6):682–688, 2013.

Ayoub, P. M. and Garretson, J. Getting the message out: Media context and global changes in attitudes toward homosexuality. *Comparative Political Studies*, 50(8):1055–1085, 2017.

Bai, H., Huan, X., Tang, W., Chen, X., Yan, H., Liu, X., Yang, H., Peng, Z., Zhao, X., Yu, R., Hao, Y., and Feng, C. A survey of HIV infection and related high-risk factors among men who have sex with men in Suzhou, Jiangsu, China. *Journal of biomedical research*, 25(1):17–24, 2011.

Banse, R., Seise, J., and Zerbes, N. Implicit attitudes towards homosexuality: Reliability, validity, and controllability of the IAT. *Zeitschrift für experimentelle Psychologie*, 48(2):145–160, 2001.

Bauer, G. R. and Brennan, D. J. The problem with 'behavioral bisexuality': Assessing sexual orientation in survey research. *Journal of Bisexuality*, 13(2):148–165, 2013.

Becker, J. C. and Barreto, M. Personal, collective, and group-distancing motives underlying confrontation of prejudice. In *Confronting Prejudice and Discrimination*, pages 141–158. Elsevier, 2019.

Berggren, N., Bjørnskov, C., and Nilsson, T. What aspects of society matter for the quality of life of a minority? Global evidence from the new gay happiness index. *Social Indicators Research*, 132(3):1163–1192, 2017.

Besen, Y., Zicklin, G., et al. Young men, religion and attitudes towards homosexuality. *Journal of Men, Masculinities and Spirituality*, 1(3):250–266, 2007.

Bjekić, M., Šipetić-Grujičić, S., Vlajinac, H., and Dunić, I. Does penis size influence sexual behaviour of men who have sex with men? *Acta facultatis medicae Naissensis*, 35(4):311–319, 2018.

Boellstorff, T. But do not identify as gay: A proleptic genealogy of the msm category. *Cultural Anthropology*, 26(2): 287–312, 2011.

Bowleg, L. Love, sex, and masculinity in sociocultural context: HIV concerns and condom use among African American men in heterosexual relationships. *Men and masculinities*, 7(2):166–186, 2004.

Bowman, K., Sims, H., and Marsico, J. Polls on attitudes on homosexuality & gay marriage. *AEI Paper & Studies*, 2014.

Brown, J. and Trevethan, R. Shame, internalized homophobia, identity formation, attachment style, and the connection to relationship status in gay men. *American journal of men's health*, 4(3):267–276, 2010.

Bybee, J. A., Sullivan, E. L., Zielonka, E., and Moes, E. Are gay men in worse mental health than heterosexual men? The role of age, shame and guilt, and coming-out. *Journal of adult development*, 16(3):144–154, 2009.

Cai, R., Zhao, J., Cai, W., Chen, L., Richardus, J. H., and de Vlas, S. J. HIV risk and prevention behaviors in men who have sex with men and women: a respondent-driven sampling study in Shenzhen, China. *AIDS and Behavior*, 18(8):1560–1568, 2014.

Cantu, L. Entre hombres/between men: Latino masculinities and homosexualities. *Research on Men and Masculinities Series*, 12:224–246, 2000.

Cantú, L. *The sexuality of migration: Border crossings and Mexican immigrant men*, volume 5. NYU Press, 2009.

Cao, Y., Li, S.-Y., Lu, L., Wang, P.-W., Chen, X.-G., and Yan, H. Comparative analysis on the characteristics of bisexually- and gay-identified men in Nanchang City. *Journal of Public Health and Preventive Medicine*, 24(4):62 – 66, 2013. Original (Chinese): 曹越, 李十月, 路亮, 王佩韦, 陈心广, 燕虹. (2013). 南昌市不同性取向男男性行为人群特征比较分析. 公共卫生与预防医学, 24(4), 62 – 66.

Carballo-Diéguez, A., Dolezal, C., Nieves, L., Díaz, F., Decena, C., and Balan, I. Looking for a tall, dark, macho man …sexual-role behaviour variations in latino gay and bisexual men. *Culture, Health & Sexuality*, 6(2):159–171, 2004.

Carrier, J. M. "sex-role preference" as an explanatory variable in homosexual behavior. *Archives of Sexual Behavior*, 6 (1):53–65, 1977.

Carrier, J. M. Gay liberation and coming out in Mexico. *Journal of Homosexuality*, 17(3–4):225–252, 1989a.

Carrier, J. M. Sexual behavior and spread of AIDS in Mexico. *Medical anthropology*, 10(2–3):129–142, 1989b.

Chen, K., Zhao, X., Chen, L., Zou, J., Li, Q., Teng, G., and Ding, X. Analysis on prevalence of HIV and syphilis and its risk behavior characteristics among men who have sex with men in Yongchuan district of Chongqing. *Chongqing Medical Journal*, 41(17):1730–1733, 2012. Original (Chinese): 陈克江, 赵希友, 陈亮, 邹静波, 李琴, 滕高松 & 丁贤彬. (2012). 永川区男男性行为人群 HIV 梅毒感染率及高危行为特征分析. 重庆医学, 41(17), 1730 − 1733.

Cheng, Y.-h. A., Wu, F.-C. F., and Adamczyk, A. Changing attitudes toward homosexuality in Taiwan, 1995–2012. *Chinese Sociological Review*, 48(4):317–345, 2016.

Chmiel, J. S., Detels, R., Kaslow, R. A., Van Raden, M., Kingsley, L. A., Brookmeyer, R., and Group, M. A. C. S. Factors associated with prevalent human immunodeficiency virus (HIV) infection in the multicenter AIDS cohort study. *American journal of epidemiology*, 126(4):568–575, 1987.

Choi, K.-H., Operario, D., Gregorich, S. E., and Han, L. Age and race mixing patterns of sexual partnerships among Asian men who have sex with men: Implications for HIV transmission and prevention. *AIDS Education and Prevention*, 15(1 Supplement):53–65, 2003.

Clark, J., Salvatierra, J., Segura, E., Salazar, X., Konda, K., Perez-Brumer, A., Hall, E., Klausner, J., Caceres, C., and Coates, T. Moderno love: sexual role-based identities and HIV/STI prevention among men who have sex with men in Lima, Peru. *AIDS and Behavior*, 17(4):1313–1328, 2013.

Clark, J. L., Segura, E. R., Oldenburg, C. E., Rios, J., Montano, S. M., Perez-Brumer, A., Villaran, M., Sanchez, J., Coates, T. J., and Lama, J. R. Expedited partner therapy (EPT) increases the frequency of partner notification among MSM in Lima, Peru: a pilot randomized controlled trial. *BMC medicine*, 15(1):94, 2017.

Clark, J. L., Segura, E. R., Oldenburg, C. E., Salvatierra, H. J., Rios, J., Perez-Brumer, A. G., Gonzales, P., Sheoran, B., Sanchez, J., and Lama, J. R. Traditional and web-based technologies to improve partner notification following syphilis diagnosis among men who have sex with men in Lima, Peru: pilot randomized controlled trial. *Journal of medical Internet research*, 20(7):e232, 2018.

Cochran, S. D. and Mays, V. M. Relation between psychiatric syndromes and behaviorally defined sexual orientation in a sample of the US population. *American journal of epidemiology*, 151(5):516–523, 2000.

Connell, R. Masculinity research and global change. *Masculinities & Social Change*, 1(1):4–18, 2012.

Connell, R. Masculinities in global perspective: Hegemony, contestation, and changing structures of power. *Theory and Society*, 45(4):303–318, 2016.

Connell, R. W. *Masculinities*. Berkeley: University of California press CA, 1995.

Connell, R. W. and Messerschmidt, J. W. Hegemonic masculinity: Rethinking the concept. *Gender & society*, 19(6): 829–859, 2005.

Cox, S. and Gallois, C. Gay and lesbian identity development: A social identity perspective. *Journal of Homosexuality*, 30(4):1–30, 1996.

Crawford, I. and Solliday, E. The attitudes of undergraduate college students toward gay parenting. *Journal of Homosexuality*, 30(4):63–77, 1996.

Dai, L.-P., Lin, Y.-X., Liu, Q., Jiang, Y., Gong, C., and Xia, J. Characteristics of sexual behavior and HIV infection among 1472 MSM with different gender role in Guangzhou city. *Chinese Journal of Public Health*, 29(2):256 – 258, 2013. Original (Chinese): 戴丽萍, 林玉虾, 刘奇, 姜宜海, 龚诚华 & 厦嘉杰. (2013). 广州市不同性角色 MSM 行为特征及艾滋病感染分析. 中国公共卫生, 29(2), 256 – 258.

Dai, Z., Zhong, X., Peng, B., Zhang, Y., Liang, H., Zou, Y., Luo, J., Peng, H., and Huang, A. Comparative analysis on characteristics of gay-identified and bisexual-identified men in western China. *Journal of Third Military Medical University*, 36(19):2025–2029, 2014. Original (Chinese): 戴涢涢, 钟晓妮, 彭斌, 张燕, 梁浩, 邹云峰, 罗锦坤, 彭鸿斌 & 黄爱龙. (2014). 中国西部地区男同性恋人群与男双性恋人群相关特征比较分析. 第三军医大学学报, 36(19), 2025–2029.

Dang, J., Liu, S.-J., and Liu, L.-H. Influencing factors analysis of holding time with regular sex partners among men who have sex with men in Shijiazhuang. *China Tropical Medicine*, 15(12):1448–1451, 2015. Original (Chinese): 党静, 刘淑君, 刘丽花. (2015). 石家庄市男男性接触者同性固定性伴保持时间影响因素. 中国热带医学, 15(12), 1448–1451.

Dangerfield, D. T., Gravitt, P., Rompalo, A. M., Tai, R., and Lim, S. H. Correlates of anal sex roles among Malay and Chinese MSM in Kuala Lumpur, Malaysia. *International journal of STD & AIDS*, 27(4):313–320, 2016.

Dangerfield, D. T., Smith, L. R., Anderson, J. N., Bruce, O. J., Farley, J., and Bluthenthal, R. Sexual positioning practices and sexual risk among black gay and bisexual men: a life course perspective. *AIDS and Behavior*, 22(6):1919–1931, 2018.

Decena, C. U. *Queering the heights: Dominican transnational identities and male homosexuality in New York City*. New York University, 2004.

Díaz, R. M. *Latino gay men and HIV: Culture, sexuality, and risk behavior*. Routledge, 2013.

Donovan, J. M. Homosexual, gay, and lesbian: Defining the words and sampling the populations. *Journal of homosexuality*, 24(1-2):27–47, 1993.

Duan, C., Wei, L., Cai, Y., Chen, L., Yang, Z., Tan, W., Gan, Y., Au, W. W., and Zhao, J. Recreational drug use and risk of HIV infection among men who have sex with men: a cross-sectional study in Shenzhen, China. *Drug and alcohol dependence*, 181:30–36, 2017.

Dube, E. M. The role of sexual behavior in the identification process of gay and bisexual males. *Journal of Sex Research*, 37(2):123–132, 2000.

Dunne, B. Power and sexuality in the Middle East. *Middle East Report*, pages 8–37, 1998.

Ellemers, N. Legitimacy of intergroup relations. *The psychology of legitimacy: Emerging perspectives on ideology, justice, and intergroup relations*, page 205, 2001.

Ellemers, N. and Haslam, S. A. Social Identity Theory. In Van Lange, P. A., Kruglanski, A. W., and Higgins, E. T., editors, *Handbook of theories of social psychology: Volume 2*, pages 379–398. Sage London, UK, 2011.

Ellemers, N., Van Knippenberg, A., De Vries, N., and Wilke, H. Social identification and permeability of group boundaries. *European Journal of Social Psychology*, 18(6):497–513, 1988.

Evans, A. R., Wiggins, R. D., Mercer, C. H., Bolding, G. J., and Elford, J. Men who have sex with men in Great Britain: comparison of a self-selected internet sample with a national probability sample. *Sexually Transmitted Infections*, 83(3):200–205, 2007.

Fan, S., Lu, H., Ma, X., Sun, Y., He, X., Li, C., Raymond, H., McFarland, W., Sun, J., Ma, W., et al. Behavioral and serologic survey of men who have sex with men in Beijing, China: implication for HIV intervention. *AIDS patient care and STDs*, 26(3):148–155, 2012.

Faul, F., Buchner, A., Erdfelder, E., and Lang, A. Gpower version 3.1.9.6. *Germany: Universitat Kiel*, 2020.

Feng, Y., Bu, K., Li, M., Zhang, X., Jin, S., and Wang, L. Meta-analysis of HIV infection incidence and risk factors among men who have sex with men in China. *Chinese Journal of Epidemiology*, 36(7):752–758, 2015. in Chinese.

Feng, Y., Wu, Z., Detels, R., Qin, G., Liu, L., Wang, X., Wang, J., and Zhang, L. HIV/STD prevalence among MSM in Chengdu, China and associated risk factors for HIV infection. *Journal of acquired immune deficiency syndromes (1999)*, 53(Supplement 1):S74, 2010.

Floyd, F. J. and Bakeman, R. Coming-out across the life course: Implications of age and historical context. *Archives of sexual behavior*, 35(3):287–296, 2006.

French, H. and Rothery, M. Hegemonic Masculinities? Assessing change and processes of change in elite masculinity, 1700–1900. In *What is Masculinity?*, pages 139–166. Springer, 2011.

Frost, D. M. and Meyer, I. H. Internalized homophobia and relationship quality among lesbians, gay men, and bisexuals. *Journal of counseling psychology*, 56(1):97–109, 2009.

Frost, D. M., Lehavot, K., and Meyer, I. H. Minority stress and physical health among sexual minority individuals. *Journal of behavioral medicine*, 38(1):1–8, 2015.

Galea, J. T., León, S. R., Peinado, J., Calvo, G., Zamora, J., Sánchez, H., and Brown, B. J. HPV knowledge, burden and genital wart location among heterosexually identified versus homosexually identified men who have sex with men in Lima, Peru: cross-sectional results from a cohort study. *BMJ open*, 7(10):e017338, 2017.

García, M. C., Meyer, S. B., and Ward, P. Elevated HIV prevalence and risk behaviours among men who have sex with men (MSM) in Vietnam: a systematic review. *BMJ Open*, 2(5), 2012.

Gates, G. J. How many people are lesbian, gay, bisexual and transgender? *The Williams Institute*, 2011.

Gerhards, J. Non-discrimination towards homosexuality: The European Union's policy and citizens' attitudes towards homosexuality in 27 european countries. *International Sociology*, 25(1):5–28, 2010.

Gil, S. A narrative exploration of gay men's sexual practices as a dialectical dialogue. *Sexual and Relationship Therapy*, 22(1):63–75, 2007.

Goodreau, S. M., Goicochea, L. P., and Sanchez, J. Sexual role and transmission of HIV type 1 among men who have sex with men in Peru. *The Journal of infectious diseases*, 191(Supplement 1):S147–S158, 2005.

Greene, D. C. and Britton, P. J. Stage of sexual minority identity formation: The impact of shame, internalized homophobia, ambivalence over emotional expression, and personal mastery. *Journal of Gay & Lesbian Mental Health*, 16(3):188–214, 2012.

Grov, C., Parsons, J. T., and Bimbi, D. S. The association between penis size and sexual health among men who have sex with men. *Archives of Sexual Behavior*, 39(3):788–797, 2010.

Guadamuz, T. E., Wimonsate, W., Varangrat, A., Phanuphak, P., Jommaroeng, R., Mock, P. A., Tappero, J. W., and van Griensven, F. Correlates of forced sex among populations of men who have sex with men in Thailand. *Archives of sexual behavior*, 40(2):259–266, 2011.

Halkitis, P. N., Shrem, M. T., and Martin, F. W. Sexual behavior patterns of methamphetamine-using gay and bisexual men. *Substance Use & Misuse*, 40(5):703–719, 2005.

Hammack, P. L. and Cohler, B. J. *The story of sexual identity: Narrative perspectives on the gay and lesbian life course*. Oxford University Press, 2009.

Han, C.-s. A qualitative exploration of the relationship between racism and unsafe sex among Asian Pacific Islander gay men. *Archives of Sexual Behavior*, 37(5):827–837, 2008.

Harada, M. Japanese male gay and bisexual identity. *Journal of Homosexuality*, 42(2):77–100, 2002.

Harris, J. I., Cook, S. W., and Kashubeck-West, S. Religious attitudes, internalized homophobia, and identity in gay and lesbian adults. *Journal of Gay & Lesbian Mental Health*, 12(3):205–225, 2008.

Harry, J. Sampling gay men. *Journal of Sex Research*, 22(1):21–34, 1986.

Hart, T. A., Wolitski, R. J., Purcell, D. W., Gómez, C., Halkitis, P., Team, S. U. M. S., et al. Sexual behavior among HIV-positive men who have sex with men: What's in a label? *Journal of Sex Research*, 40(2):179–188, 2003.

Hearn, J., Nordberg, M., Andersson, K., Balkmar, D., Gottzén, L., Klinth, R., Pringle, K., and Sandberg, L. Hegemonic Masculinity and beyond: 40 years of research in Sweden. *Men and masculinities*, 15(1):31–55, 2012.

Heckman, T. G., Kelly, J. A., Bogart, L. M., Kalichman, S. C., and Rompa, D. J. HIV risk differences between African-American and white men who have sex with men. *Journal of the National Medical Association*, 91(2):92, 1999.

Hidaka, T. *Salaryman masculinity: Continuity and change in hegemonic masculinity in Japan*. Brill, 2010.

Hirschman, A. O. *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*, volume 25. Harvard university press, 1970.

Ho, P. S. Y. and Tsang, A. K. T. Negotiating anal intercourse in inter-racial gay relationships in Hong Kong. *Sexualities*, 3(3):299–323, 2000.

Hodge, S. No fags out there: gay men, identity and suburbia. *journal of interdisciplinary gender studies*, 1(1):41–48, 1995.

Hogg, M. and Abrams, D. *Social Identifications*. London: Routledge, 1988.

Hogg, M. A. and Reid, S. A. Social identity, self-categorization, and the communication of group norms. *Communication theory*, 16(1):7–30, 2006.

Hoppe, T. Circuits of power, circuits of pleasure: Sexual scripting in gay men's bottom narratives. *Sexualities*, 14(2): 193–217, 2011.

Hozo, S. P., Djulbegovic, B., and Hozo, I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC medical research methodology*, 5(1):13, 2005.

Hu, Y., Qian, H.-Z., Sun, J., Gao, L., Yin, L., Li, X., Xiao, D., Li, D., Sun, X., Ruan, Y., et al. Anal human papillomavirus infection among HIV-infected and uninfected men who have sex with men in Beijing, China. *Journal of acquired immune deficiency syndromes (1999)*, 64(1):103, 2013.

Hu, Y., Lu, H., Raymond, H. F., Sun, Y., Sun, J., Jia, Y., He, X., Fan, S., Xiao, Y., McFarland, W., et al. Measures of condom and safer sex social norms and stigma towards HIV/AIDS among Beijing MSM. *AIDS and Behavior*, 18(6):1068–1074, 2014.

Huang, S.-W., Wang, S.-F., Cowó, Á. E., Chen, M., Lin, Y.-T., Hung, C.-P., Chen, Y.-H., Yang, J.-Y., Tang, H.-J., and Chen, Y.-M. A. Molecular epidemiology of HIV-1 infection among men who have sex with men in Taiwan in 2012. *PLoS one*, 10(6):e0128266, 2015.

Jameson, D. R., Celum, C. L., Manhart, L., Menza, T. W., and Golden, M. R. The association between lack of circumcision and HIV, HSV-2, and other sexually transmitted infections among men who have sex with men. *Sexually transmitted diseases*, 37(3):147–152, 2010.

Jewkes, R., Morrell, R., Hearn, J., Lundqvist, E., Blackbeard, D., Lindegger, G., Quayle, M., Sikweyiya, Y., and Gottzén, L. Hegemonic masculinity: combining theory and practice in gender interventions. *Culture, health & sexuality*, 17(sup2):112–127, 2015.

Jin, X., Xu, J., Smith, M. K., Xiao, D., Rapheal, E. R., Xiu, X., Ding, Z., Zhang, Y., Jie, Y., Liao, Y., Cao, N., Whu, H., and Bao, Y. An Internet-Based self-testing model (Easy Test): Cross-sectional survey targeting men who have sex with men who never tested for HIV in 14 provinces of China. *Journal of medical Internet research*, 21(5): e11854, 2019.

Johns, M. M., Pingel, E., Eisenberg, A., Santana, M. L., and Bauermeister, J. Butch tops and femme bottoms? Sexual positioning, sexual decision making, and gender roles among young gay men. *American journal of men's health*, 6(6):505–518, 2012.

Joseph, K. M., Adib, S. M., Joseph, J. G., and Tal, M. Gay identity and risky sexual behavior related to the AIDS threat. *Journal of Community Health*, 16(6):287–297, 1991.

Kelley, J. et al. Attitudes towards homosexuality in 29 nations. *Australian Social Monitor*, 4(1):15–22, 2001.

Kinsey, A. C. Criteria for a hormonal explanation of the homosexual. *The Journal of Clinical Endocrinology*, 1(5): 424–428, 1941.

Kinsey, A. C., Pomeroy, W. B., Martin, C. E., and Sloan, S. *Sexual behavior in the human male*, volume 1. Saunders Philadelphia, 1948.

Kippax, S. and Smith, G. Anal intercourse and power in sex between men. *Sexualities*, 4(4):413–434, 2001.

Kite, M. E. and Bryant-Lees, K. B. Historical and contemporary attitudes toward homosexuality. *Teaching of Psychology*, 43(2):164–170, 2016.

Kojima, N., Park, H., Konda, K. A., Davey, D. L. J., Bristow, C. C., Brown, B., Leon, S. R., Vargas, S. K., Calvo, G. M., Caceres, C. F., et al. The PICASSO cohort: baseline characteristics of a cohort of men who have sex with men and male-to-female transgender women at high risk for syphilis infection in Lima, Peru. *BMC infectious diseases*, 17(1):255, 2017.

Lama, J. R., Agurto, H. S., Guanira, J. V., Ganoza, C., Casapia, M., Ojeda, N., Ortiz, A., Zamalloa, V., Suarez-Ognio, L., Cabezas, C., et al. Hepatitis B infection and association with other sexually transmitted infections among men who have sex with men in Peru. *The American journal of tropical medicine and hygiene*, 83(1):194–200, 2010.

Lancaster, R. N. Subject honor and object shame: The construction of male homosexuality and stigma in Nicaragua. *Ethnology*, 27(2):111–125, 1988.

Laner, M. R. and Laner, R. H. Personal style or sexual preference: Why gay men are disliked. *International Review of Modern Sociology*, pages 215–228, 1979.

Lau, J. F., Zhao, J.-K., Wu, X.-B., Gu, J., and Hao, C. Gay saunas and the risks of HIV and syphilis transmissions in China – results of a meta-analysis. *The journal of sexual medicine*, 10(3):642–652, 2013.

Leaper, C. More similarities than differences in contemporary theories of social development? A plea for theory bridging. In *Advances in child development and behavior*, volume 40, pages 337–378. Elsevier, 2011.

Lei, Y., Zhang, K., Xiao, X., Tang, C., Li, X., and Wang, H. Sexual roles, risk sexual behaviours, and HIV prevalence among men who have sex with men seeking HIV testing in changsha, china. *Current HIV research*, 16(2): 174–181, 2018.

Lick, D. J. and Johnson, K. L. Intersecting race and gender cues are associated with perceptions of gay men's preferred sexual roles. *Archives of sexual behavior*, 44(5):1471–1481, 2015.

Liu, J. X. and Choi, K. Experiences of social discrimination among men who have sex with men in shanghai, china. *AIDS and Behavior*, 10(1):25–33, 2006.

Loftus, J. America's liberalization in attitudes toward homosexuality, 1973 to 1998. *American sociological review*, pages 762–782, 2001.

Lovaas, K. E. and Jenkins, M. M. *Sexualities and communication in everyday life: A reader*. Sage, 2007.

Luhtanen, R. and Crocker, J. A collective self-esteem scale: Self-evaluation of one's social identity. *Personality and social psychology bulletin*, 18(3):302–318, 1992.

Luo, D., Wan, X., Liu, J., and Tong, T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. *Statistical methods in medical research*, 27(6):1785–1805, 2018.

Lyons, A., Pitts, M., Smith, G., Grierson, J., Smith, A., McNally, S., and Couch, M. Versatility and HIV vulnerability: investigating the proportion of Australian gay men having both insertive and receptive anal intercourse. *The journal of sexual medicine*, 8(8):2164–2171, 2011.

Lyons, A., Pitts, M., and Grierson, J. Versatility and HIV vulnerability: Patterns of insertive and receptive anal sex in a national sample of older Australian gay men. *AIDS and Behavior*, 17(4):1370–1377, 2013.

Manca, V. *Infobiotics: Information in Biotic Systems*. Emergence, Complexity and Computation. Springer Berlin Heidelberg, 2013.

Mao, L., Templeton, D. J., Crawford, J., Imrie, J., Prestage, G. P., Grulich, A. E., Donovan, B., Kaldor, J. M., and Kippax, S. C. Does circumcision make a difference to the sexual experience of gay men? Findings from the Health in Men (HIM) cohort. *The journal of sexual medicine*, 5(11):2557–2561, 2008.

McKirnan, D. J., Stokes, J. P., Doll, L., and Burzette, R. G. Bisexually active men: Social characteristics and sexual behavior. *Journal of Sex Research*, 32(1):65–76, 1995.

McLaren, S. The interrelations between internalized homophobia, depressive symptoms, and suicidal ideation among Australian gay men, lesbians, and bisexual women. *Journal of homosexuality*, 63(2):156–168, 2016.

McLean, S. A., Galea, J. T., Prudden, H. J., Calvo, G., Sánchez, H., and Brown, B. Association between sexual role and HIV status among Peruvian men who have sex with men seeking an HIV test: a cross-sectional analysis. *International journal of STD & AIDS*, 27(9):783–789, 2016.

Meng, X., Zou, H., Fan, S., Zheng, B., Zhang, L., Dai, X., Deng, M., Zhang, X., and Lu, B. Relative risk for HIV infection among men who have sex with men engaging in different roles in anal sex: a systematic review and meta-analysis on global data. *AIDS and Behavior*, 19(5):882–889, 2015.

Meyer, I. H. Minority stress and mental health in gay men. *Journal of health and social behavior*, pages 38–56, 1995.

Meyer, I. H. and Dean, L. Internalized homophobia, intimacy, and sexual behavior among gay and bisexual men. *Psychological perspectives on lesbian and gay issues*, 4:160–186, 1998.

Meyer, I. H. and Wilson, P. A. Sampling lesbian, gay, and bisexual populations. *Journal of counseling psychology*, 56 (1):23, 2009.

Mok, T. and Iz, H. B. Vector regression introduced. *Journal of geodetic science*, 2014.

Morales, J. Chapter 6: Gay latinos and aids: A framework for hiv/aids prevention curriculum. *Journal of Gay & Lesbian Social Services*, 2(3-4):89–106, 1995.

Moskowitz, D. A. and Garcia, C. P. Top, bottom, and versatile anal sex roles in same-sex male relationships: Implications for relationship and sexual satisfaction. *Archives of sexual behavior*, 48(4):1217–1225, 2019.

Moskowitz, D. A. and Hart, T. A. The influence of physical body traits and masculinity on anal sex roles in gay and bisexual men. *Archives of Sexual Behavior*, 40(4):835–841, 2011.

Moskowitz, D. A., Rieger, G., and Roloff, M. E. Tops, bottoms and versatiles. *Sexual and Relationship Therapy*, 23(3): 191–202, 2008.

Murray, S. O. *Male Homosexuality in Central and South America*. Number 5. Instituto Obregón, 1987.

Ohlert, J. and Zepp, C. Theory-based team diagnostics and interventions. In *Sport and exercise psychology research*, pages 347–370. Elsevier, 2016.

Oldmeadow, J. A. and Fiske, S. T. Social status and the pursuit of positive social identity: Systematic domains of intergroup differentiation and discrimination for high-and low-status groups. *Group Processes & Intergroup Relations*, 13 (4):425–444, 2010.

Pan, S., Xu, J.-J., Han, X.-X., Zhang, J., Hu, Q.-H., Chu, Z.-X., Hai, Y.-Q., Mao, X., Yu, Y.-Q., Geng, W.-Q., et al. Internet-based sex-seeking behavior promotes HIV infection risk: a 6-year serial cross-sectional survey to MSM in Shenyang, China. *BioMed research international*, 2016, 2016.

Parker, A. and Bhugra, D. Attitudes of British medical students towards male homosexuality. *Sexual and relationship therapy*, 15(2):141–149, 2000.

Pathela, P., Hajat, A., Schillinger, J., Blank, S., Sell, R., and Mostashari, F. Discordance between sexual behavior and self-reported sexual identity: a population-based survey of new york city men. *Annals of Internal Medicine*, 145(6):416–425, 2006.

Paz-Bailey, G., Miller, W., Shiraishi, R. W., Jacobson, J. O., Abimbola, T. O., and Chen, S. Y. Reaching men who have sex with men: a comparison of respondent-driven sampling and time-location sampling in Guatemala City. *AIDS and Behavior*, 17(9):3081–3090, 2013.

Peinado, J., Goodreau, S. M., Goicochea, P., Vergara, J., Ojeda, N., Casapia, M., Ortiz, A., Zamalloa, V., Galvan, R., and Sanchez, J. R. Role versatility among men who have sex with men in urban Peru. *Journal of sex research*, 44 (3):233–239, 2007.

Perez-Brumer, A. G., Konda, K. A., Salvatierra, H. J., Segura, E. R., Hall, E. R., Montano, S. M., Coates, T. J., Klausner, J. D., Caceres, C. F., and Clark, J. L. Prevalence of HIV, STIs, and risk behaviors in a cross-sectional community- and clinic-based sample of men who have sex with men (MSM) in Lima, Peru. *PloS one*, 8(4), 2013.

Perez-Brumer, A. G., Oldenburg, C. E., Segura, E. R., Sanchez, J., Lama, J. R., and Clark, J. L. Anonymous partnerships among MSM and transgender women (TW) recently diagnosed with HIV and other STIs in Lima, Peru: an individual-level and dyad-level analysis. *Sexually transmitted infections*, 92(7):554–557, 2016.

Plug, E. and Berkhout, P. Effects of sexual preferences on earnings in the Netherlands. *Journal of Population Economics*, 17(1):117–131, 2004.

Qian, H.-Z., Ruan, Y., Liu, Y., Milam, D. F., Spiegel, H. M., Yin, L., Li, D., Shepherd, B. E., Shao, Y., and Vermund, S. H. Lower HIV risk among circumcised men who have sex with men in China: Interaction with anal sex role in a cross-sectional study. *Journal of acquired immune deficiency syndromes (1999)*, 71(4):444–451, 2016.

Ramirez-Valles, J., Heckathorn, D. D., Vázquez, R., Diaz, R. M., and Campbell, R. T. From networks to populations: the development and application of respondent-driven sampling among IDUs and Latino gay men. *AIDS and Behavior*, 9(4):387–402, 2005.

Ravenhill, J. P. and de Visser, R. O. "it takes a man to put me on the bottom": Gay men's experiences of masculinity and anal intercourse. *The Journal of Sex Research*, 55(8):1033–1047, 2018.

Reese, G., Rosenmann, A., and Cameron, J. E. *The Psychology of Globalization: Identity, Ideology, and Action*. Academic Press, 2019.

Ricciardelli, R., Clow, K. A., and White, P. Investigating hegemonic masculinity: Portrayals of masculinity in men's lifestyle magazines. *Sex Roles*, 63(1-2):64–78, 2010.

Rietmeijer, C. A., Wolitski, R. J., Fishbein, M., Corby, N. H., and Cohn, D. L. Sex hustling, injection drug use, and non-gay identification by men who have sex with men: Associations with high-risk sexual behaviors and condom use. *Sexually transmitted diseases*, 25(7):353–360, 1998.

Rodriguez, N. S. "Gay" Men. *Intercultural Communication, Identity, and Social Movements in the Digital Age*, 2019.

Rosario, M., Schrimshaw, E. W., Hunter, J., and Braun, L. Sexual identity development among lesbian, gay, and bisexual youths: Consistency and change over time. *Journal of sex research*, 43(1):46–58, 2006.

Rowen, C. J. and Malcolm, J. P. Correlates of internalized homophobia and homosexual identity formation in a sample of gay men. *Journal of Homosexuality*, 43(2):77–92, 2003.

Ruan, Y., Qian, H.-Z., Li, D., Shi, W., Li, Q., Liang, H., Yang, Y., Luo, F., Vermund, S. H., and Shao, Y. Willingness to be circumcised for preventing HIV among Chinese men who have sex with men. *AIDS patient care and STDs*, 23(5):315–321, 2009.

Rubin, A. and Babbie, R. *Essential Research Methods for Social Work*. Cengage Learning, 2009.

Sandfort, T. G., Lane, T., Dolezal, C., and Reddy, V. Gender expression and risk of HIV infection among black South African men who have sex with men. *AIDS and Behavior*, 19(12):2270–2279, 2015.

Scheepers, D. and Ellemers, N. Social identity theory. In *Social Psychology in Action*, pages 129–143. Springer, 2019.

Schnabel, L. Gender and homosexuality attitudes across religious groups from the 1970s to 2014: Similarity, distinction, and adaptation. *Social Science Research*, 55:31–47, 2016.

Seguino, S. Plusça change? Evidence on global trends in gender norms and stereotypes. *Feminist Economics*, 13(2):1–28, 2007.

She, M., Zhang, H., Wang, J., Xu, J., Zhang, Z., Fan, Y., Sun, Y., and Zhang, C. Associated factors for HIV and syphilis infection among men who have sex with men only and men who have sex with both men and women in cities of China. *International journal of STD & AIDS*, 24(4):293–300, 2013.

Simmons, H. Dying for love: homosexuality in the Middle East. *Human rights and human welfare (Issue on Human Rights in the Middle East and North Africa)*, pages 160–172, 2014.

Smith, T. W. Cross-national differences in attitudes towards homosexuality. Technical report, GSS Cross-National Report No. 31, 2011.

Smith, T. W., Son, J., and Kim, J. Public attitudes toward homosexuality and gay rights across time and countries. *NORC at the University of Chicago*, 2014.

Song, Y., Cai, Y., and Hong, F. Marital status and its effect on syphilis/HIV infection among men who have sex with men in Shenzhen. *Chinese Journal of AIDS & STD*, 22(10):806–809, 2016. Original (Chinese): 宋亚娟, 蔡于茂 & 洪福昌. (2016). 深圳市男男性行为者婚姻状况及其对梅毒和 HIV 感染的影响. 中国艾滋病性病, 22(10), 806–809.

Starks, T. J., Rendina, H. J., Breslow, A. S., Parsons, J. T., and Golub, S. A. The psychological cost of anticipating HIV stigma for HIV-negative gay and bisexual men. *AIDS and Behavior*, 17(8):2732–2741, 2013.

Stock, J. H. and Watson, M. W. *Introduction to econometrics: Global edition*. Pearson Education Boston, MA, 2012.

Swift-Gallant, A., Coome, L. A., Monks, D. A., and VanderLaan, D. P. Handedness is a biomarker of variation in anal sex role behavior and recalled childhood gender nonconformity among gay men. *PLoS one*, 12(2):e0170241, 2017.

Swift-Gallant, A., Coome, L. A., Skorska, M. N., Monks, D. A., and VanderLaan, D. P. Pubertal onset and anal sex role among gay men. *Psychology of Sexual Orientation and Gender Diversity*, 6(1):107–112, 2018.

Tajfel, H. and Turner, J. C. The social identity theory of intergroup behavior. In Austin, W. G. and Worchel, S., editors, *The social psychology of intergroup relations*, pages 33–47. Monterey; CA: Brooks/Cole, 1979.

Tajfel, H. *Human groups and social categories: Studies in social psychology*. Cup Archive, 1981.

Tajfel, H. E. Social identity and intergroup behaviour. *Information (International Social Science Council)*, 13(2):65–93, 1974.

Tajfel, H. E. Interindividual behaviour and intergroup behaviour. In *Differentiation between social groups: Studies in the social psychology of intergroup relations*, pages 27–60. Academic Press, London, 1978a.

Tajfel, H. E. Social categorization, social identity, and social comparison. In *Differentiation between social groups: Studies in the social psychology of intergroup relations*, pages 61–76. Academic Press, London, 1978b.

Tan, J. Y., Pratto, F., Operario, D., and Dworkin, S. L. Sexual positioning and race-based attraction by preferences for social dominance among gay Asian/Pacific Islander men in the United States. *Archives of sexual behavior*, 42 (7):1233–1239, 2013.

Tang, E. C., Galea, J. T., Kinsler, J. J., Gonzales, P., Sobieszczyk, M. E., Sanchez, J., and Lama, J. R. Using conjoint analysis to determine the impact of product and user characteristics on acceptability of rectal microbicides for HIV prevention among Peruvian men who have sex with men. *Sex Transm Infect*, 92(3):200–205, 2016a.

Tang, W., Best, J., Zhang, Y., Liu, F.-y., Tso, L. S., Huang, S., Yang, B., Wei, C., and Tucker, J. D. Gay mobile apps and the evolving virtual risk environment: a cross-sectional online survey among men who have sex with men in China. *Sex Transm Infect*, 92(7):508–514, 2016b.

Thompson, E. H., Pleck, J. H., and Ferrera, D. L. Men and masculinities: Scales for masculinity ideology and masculinity-related constructs. *Sex roles*, 27(11-12):573–607, 1992.

Toepoel, V. *Doing Surveys Online*. SAGE Publications, 2015.

Troiden, D. R. R. The formation of homosexual identities. *Journal of homosexuality*, 17(1-2):43–74, 1989.

Tskhay, K. O. and Rule, N. O. Accuracy in categorizing perceptually ambiguous groups: A review and meta-analysis. *Personality and Social Psychology Review*, 17(1):72–86, 2013.

Turner, J. C. The experimental social psychology of intergroup behaviour. *Intergroup behaviour*, 66, 1981.

Turner, J. C. The analysis of social influence. In Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., and Wetherell, M. S., editors, *Rediscovering the social group: A self-categorization theory*. Basil Blackwell, 1987.

Valdiserri, R. O., Lyter, D., Leviton, L. C., Callahan, C. M., Kingsley, L. A., and Rinaldo, C. R. Variables influencing condom use in a cohort of gay and bisexual men. *American Journal of Public Health*, 78(7):801–805, 1988.

Valentova, J. V., Stulp, G., Třebickỳ, V., and Havlíček, J. Preferred and actual relative height among homosexual male partners vary with preferred dominance and sex role. *PloS one*, 9(1):e86534, 2014.

Van de Ven, P., Kippax, S., Crawford, J., Rawstorne, P., Prestage, G., Grulich, A., and Murphy, D. In a minority of gay men, sexual risk practice indicates strategic positioning for perceived risk reduction rather than unbridled sex. *AIDS care*, 14(4):471–480, 2002.

Van Griensven, F., Thanprasertsuk, S., Jommaroeng, R., Mansergh, G., Naorat, S., Jenkins, R. A., Ungchusak, K., Phanuphak, P., Tappero, J. W., Group, B. M. S. S., et al. Evidence of a previously undocumented epidemic of HIV infection among men who have sex with men in Bangkok, Thailand. *Aids*, 19(5):521–526, 2005.

van Griensven, G. J. P., Tielman, R. A. P., Goudsmit, J., Van der Noordaa, J., De Wolf, F., De Vroome, E. M. M., and Coutinho, R. A. Risk factors and prevalence of HIV antibodies in homosexual men in the Netherlands. *American Journal of Epidemiology*, 125(6):1048–1057, 1987.

Vansintejan, J., Vandevoorde, J., and Devroey, D. The GAy MEn Sex StudieS: Design of an online registration of sexual behaviour of men having sex with men and preliminary results (GAMESSS-study). *Central European journal of public health*, 21(1):48, 2013.

Wan, X., Wang, W., Liu, J., and Tong, T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC medical research methodology*, 14(1):135, 2014.

Wang, J., Xu, J., Dou, Z., Ruan, Y., Min, X., Li, F., Zhang, L., Liu, W., Limei, S., Zhuang, M., Feng, L., and Wu, Z. HIV, syphilis and HSV-2 prevalence in relation to receptive，insertive and both sex roles among MSM in eight cities of China. *Chinese Journal of AIDS & STD*, 18(5):306 – 309, 2012. Original (Chinese): 王珏, 徐杰, 斗智, 阮玉华, 闵向东, 李凡, 张灵麟, 刘伟, 申莉梅, 庄鸣华, 冯连贵 & 吴尊友. (2012). 八城市男男性行为人群不同性角色高危行为 及 HIV 梅毒 生殖器疱疹感染状况分析. 中国艾滋病性病, 18(5), 306 – 309.

Wang, Y., Li, L.-L., Zhang, G.-G., Fan, J., Zhao, X.-H., Shu, Y., Jia, S.-G., and Zhou, L. Anal sex roles and related behaviors among men who have sex with men (MSM)and the HIV/syphilis infection in Mianyang city of China. *Chinese Journal of Viral Diseases*, 4(1):43–48, 2014. Original (Chinese): 王毅, 李六林, 张光贵, 樊静, 赵西和, 舒亚, 贾蜀光 & 周力. (2014). 男男性行为者不同性角色 HIV/梅毒感染 及与相关行为特征的关系研究. 中国病毒病杂志, 4(1), 43–48.

Wang, Y., Li, L.-L., Fan, J., Zhao, X.-H., Zhou, L., and Wang, S. Analysis on condoms carrying and its influence factors among MSM of Mianyang in Sichuan. *Chinese Journal of AIDS & STD*, 21(10):861–864, 2015. Original (Chinese): 王毅, 李六林, 樊静, 赵西和, 周力 & 王思源. (2015). 四川省绵阳市 MSM 安全套携带现状及影响因素. 中国艾滋病性病, 21(10), 861–864.

Wang, Z., Zhang, F., Zeng, X., and Zheng, Y. Social acceptance and sexual behavior characteristics of MSM young students with different sexual roles in Haikou. *China Tropical Medicine*, 19(6):542–547, 2019. Original (Chinese): 王召乾, 张帆, 曾小妹 & 郑钰. (2019). 海口市不同性角色 MSM 青年学生社会接纳和性行为特征. 中国热带医学, 19(6), 542–547.

Washington, T. A., Galai, N., Cohn, S., Celentano, D. D., Vlahov, D., and Strathdee, S. A. The relationship between self-reported sexual orientation and behavior in a sample of middle-aged male injection drug users. *Archives of Sexual Behavior*, 35(1):67–74, 2006.

Weatherburn, P., Hickson, F., Reid, D., Davies, P., and Crosier, A. Sexual hiv risk behaviour among men who have sex with both men and women. *AIDS care*, 10(4):463–471, 1998.

Wegesin, D. J. and Meyer-Bahlburg, H. F. Top/bottom self-label, anal sex practices, HIV risk and gender role identity in gay men in New York City. *Journal of Psychology & Human Sexuality*, 12(3):43–62, 2000.

Wei, C. and Raymond, H. F. Preference for and maintenance of anal sex roles among men who have sex with men: Sociodemographic and behavioral correlates. *Archives of sexual behavior*, 40(4):829–834, 2011.

Weinrich, J. D., Grant, I., Jacobson, D. L., Robinson, S. R., McCutchan, J. A., Group, H., et al. Effects of recalled childhood gender nonconformity on adult genitoerotic role and AIDS exposure. *Archives of Sexual Behavior*, 21(6):559–585, 1992.

Weisberg, H. *The Total Survey Error Approach: A Guide to the New Science of Survey Research.* University of Chicago Press, 2009.

Wen, G. and Zheng, L. The influence of internalized homophobia on health-related quality of life and life satisfaction among gay and bisexual men in China. *American journal of men's health*, 13(4), 2019.

Whitehead, T. L. Urban low-income african american men, hiv/aids, and gender identity. *Medical Anthropology Quarterly*, 11(4):411–447, 1997.

Wohl, A. R., Johnson, D. F., Lu, S., Jordan, W., Beall, G., Currier, J., and Simon, P. A. HIV risk behaviors among African American men in Los Angeles County who self-identify as heterosexual. *JAIDS-HAGERSTOWN MD-*, 31(3): 354–360, 2002.

Wolitski, R. J., Rietmeijer, C., Goldbaum, G., and Wilson, R. HIV serostatus disclosure among gay and bisexual men in four American cities: general patterns and relation to sexual practices. *AIDS care*, 10(5):599–610, 1998.

Wong, D. Hybridization and the emergence of "gay" identities in Hong Kong and in China. *Visual Anthropology*, 24(1-2): 152–170, 2011.

Wooldridge, J. M. *Introductory econometrics: A modern approach.* Nelson Education, 2016.

Xie, Y. and Peng, M. Attitudes toward homosexuality in China: Exploring the effects of religion, modernizing factors, and traditional culture. *Journal of homosexuality*, 65(13):1758–1787, 2018.

Xu, J.-J., Qian, H.-Z., Chu, Z.-X., Zhang, J., Hu, Q.-H., Jiang, Y.-J., Geng, W.-Q., Zhang, C. M., and Shang, H. Recreational drug use among Chinese men who have sex with men: a risky combination with unprotected sex for acquiring HIV infection. *BioMed research international*, 2014, 2014.

Xu, W., Zheng, Y., and Kaufman, M. R. Predictors of recent HIV testing among Chinese men who have sex with men: a barrier perspective. *AIDS patient care and STDs*, 32(10):408–417, 2018.

Xu, Y. and Zheng, Y. The influence of power and intimacy sexual motives on sexual position preference among men who have sex with men in China. *Archives of sexual behavior*, 47(1):245–258, 2018.

Yang, A. S. Trends: Attitudes toward homosexuality. *The Public Opinion Quarterly*, 61(3):477–507, 1997.

Yang, H.-M., Wang, W., Chen, X.-F., Yan, H., and Li, S.-Y. Multiple partners and its correlative factors among young gay men. *Modern Preventive Medicine*, 45(13):2460–2463, 2018. Original (Chinese): 杨惠民, 王威, 陈向凡, 燕虹 & 李十月. (2018). 青少年男男性行为人群多性伴与相关因素分析. 现代预防医学, 45(13), 2460–2463.

Yarhouse, M. A., Kays, J. L., Poma, H., and Atkinson, A. N. Characteristics of mixed orientation couples: An empirical study. *Edification: The Transdisciplinary Journal of Christian Psychology*, 4(2), 2011.

Ye, M. and Giri, M. Prevalence and correlates of HIV infection among men who have sex with men: a multi-provincial cross-sectional study in the southwest of China. *HIV/AIDS (Auckland, NZ)*, 10:167, 2018.

Zeng, X., Zhong, X., Peng, B., Zhang, Y., Kong, C., Liu, X., and Huang, A. Prevalence and associated risk characteristics of HIV infection based on anal sexual role among men who have sex with men: a multi-city cross-sectional study in Western China. *International Journal of Infectious Diseases*, 49:111–118, 2016.

Zepp, C., Kleinert, J., and Liebscher, A. Inhalte und Strukturen prototypischer Merkmale in Fußballmannschaften. *Sportwissenschaft*, 43(4):283–290, 2013.

Zhang, D., Bi, P., Lv, F., Zhang, J., and Hiller, J. E. Differences between internet and community samples of MSM: implications for behavioral surveillance among MSM in China. *AIDS care*, 20(9):1128–1137, 2008.

Zhang, H.-B., Zhao, J., Au, W. W., Liu, S.-C., Wu, Y., and Zhang, X.-X. Characteristics of sexual behavior and related factors among MSM with different sexual roles in Shenzhen. *China Tropical Medicine*, 16(5):470–474, 2016. Original (Chinese): 张海波, 赵锦, William W. Au, 刘少础, 吴宇 & 张潇侠. (2016). 深圳市不同性角色 MSM 性行为特征及影响因素分析. 中国热带医学, 16(5), 470–474.

Zhang, H., Lu, H., Pan, S. W., Xia, D., Zhao, Y., Xiao, Y., He, X., Yue, H., Sun, Z., Xu, Y., et al. Correlates of unprotected anal intercourse: the influence of anal sex position among men who have sex with men in Beijing, China. *Archives of sexual behavior*, 44(2):375–387, 2015.

Zhang, J., Zheng, L., and Zheng, Y. Consistency in preferences for masculinity in faces, bodies, voices, and personality characteristics among homosexual men in China. *Personality and Individual Differences*, 134:137–142, 2018.

Zhang, J.-Y., Li, Q.-Q., Han, D.-L., Wang, X.-D., Wu, C.-L., He, Q.-Y., Liu, S., and Luan, R.-S. The sexual role status and its relationship with condom use and HIV-testing in men who have sex with men. *Chinese Journal of Disease Control & Prevention*, 17(5):400–403, 2013a. Original (Chinese): 张静瑜, 李青青, 韩德琳, 王晓东, 吴春霖, 何勤英, 刘寿 & 栾荣生. (2013). 男男性行为人群性角色状况及其与 安全套使用和 HIV 检测的关系. 中华疾病控制杂志, 17(5), 400–403.

Zhang, L., Xiao, Y., Lu, R., Wu, G., Ding, X., Qian, H.-z., McFarland, W., Ruan, Y., Vermund, S. H., and Shao, Y. Predictors of HIV testing among men who have sex with men in a large Chinese city. *Sexually transmitted diseases*, 40(3):235, 2013b.

Zhang, W., Shi, P., Shi, W., Li, G., Yang, T., Mo, Y., and Xu, Y. Characteristics of sexual orientation and sexual behavior among men who have sex both with men and women. *Journal of Huazhong University of Science and Technology*, 41(3):375–383, 2012. Original (Chinese): 张万宏, 石萍, 石卫东, 李刚, 杨涛, 莫运滨 & 徐业华. (2012). 男性双性性行为者性取向及群体性行为特征分析. 华中科技大学学报 (医学版), 41(3), 375–383.

Zheng, J., Zhang, X., Lu, C., Fu, J., and Ma, W. Characteristics of homosexual behaviors and relationship between high-risk sexual behaviors and seeking sex partners through the internet in men who have sex with men among university students. *Journal of Shandong University (Health Sciences)*, 45(7):115–123, 2017. Original (Chinese): 郑杰滔, 张晓菲, 吕翠霞, 傅继华 & 马伟. (2017). 高校学生男男性行为特征及其高危性行为与互联网交友的关联. 山东大学学报 (医学版), 55(7), 115–123.

Zheng, L. and Zheng, Y. Sexual self-labels and preference for partner's sex role among Chinese gay men and women. *Chinese Mental Heath Journal*, 23(3):77–80, 2009.

Zheng, L. and Zheng, Y. Preferences for masculinity across faces, bodies, and personality traits in homosexual and bisexual Chinese men: Relationship to sexual self-labels and attitudes toward masculinity. *Archives of sexual behavior*, 45(3):725–733, 2016.

Zheng, L. and Zheng, Y. Sexual satisfaction in Chinese gay and bisexual men: relationship to negative sexual minority identity and sexual role preference. *Sexual and Relationship Therapy*, 32(1):75–88, 2017.

Zheng, L., Hart, T. A., and Zheng, Y. The relationship between intercourse preference positions and personality traits among gay men in China. *Archives of Sexual Behavior*, 41(3):683–689, 2012.

Zheng, L., Hart, T. A., and Zheng, Y. Attraction to male facial masculinity in gay men in China: Relationship to intercourse preference positions and sociosexual behavior. *Archives of Sexual Behavior*, 42(7):1223–1232, 2013.

Zheng, L., Hart, T. A., and Zheng, Y. Top/bottom sexual self-labels and empathizing–systemizing cognitive styles among gay men in China. *Archives of sexual behavior*, 44(5):1431–1438, 2015.

Zhou, C., Raymond, H. F., Ding, X., Lu, R., Xu, J., Wu, G., Feng, L., Fan, S., Li, X., McFarland, W., et al. Anal sex role, circumcision status, and HIV infection among men who have sex with men in Chongqing, China. *Archives of sexual behavior*, 42(7):1275–1283, 2013.

Zunt, J. R., La Rosa, A. M., Peinado, J., Lama, J. R., Suarez, L., Pun, M., Cabezas, C., and Sanchez, J. Risk factors for HTLV-II infection in Peruvian men who have sex with men. *The American journal of tropical medicine and hygiene*, 74(5):922–925, 2006.

Appendix A1
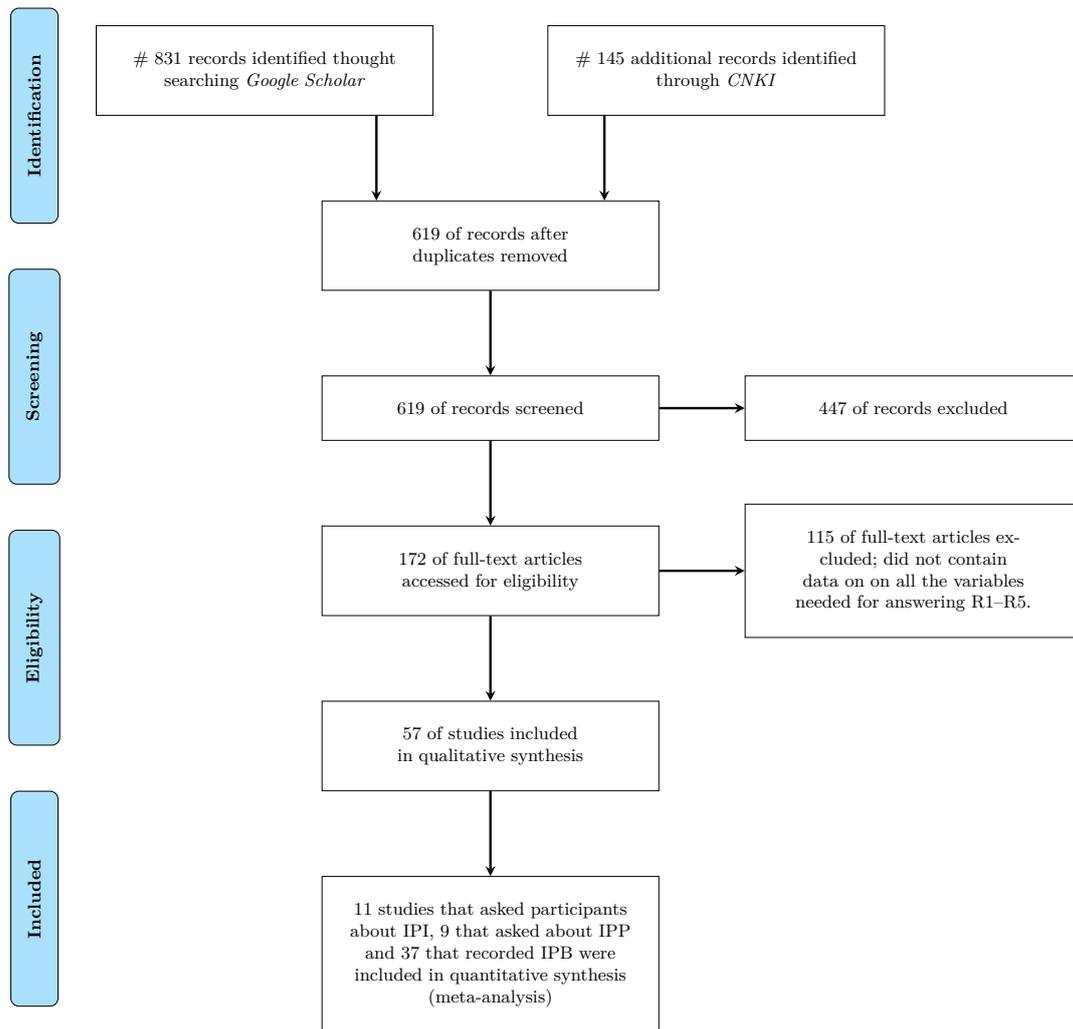
Figure 1. PRISMA Flow Diagram for selecting studies for answering R1–R5.

Appendix A2
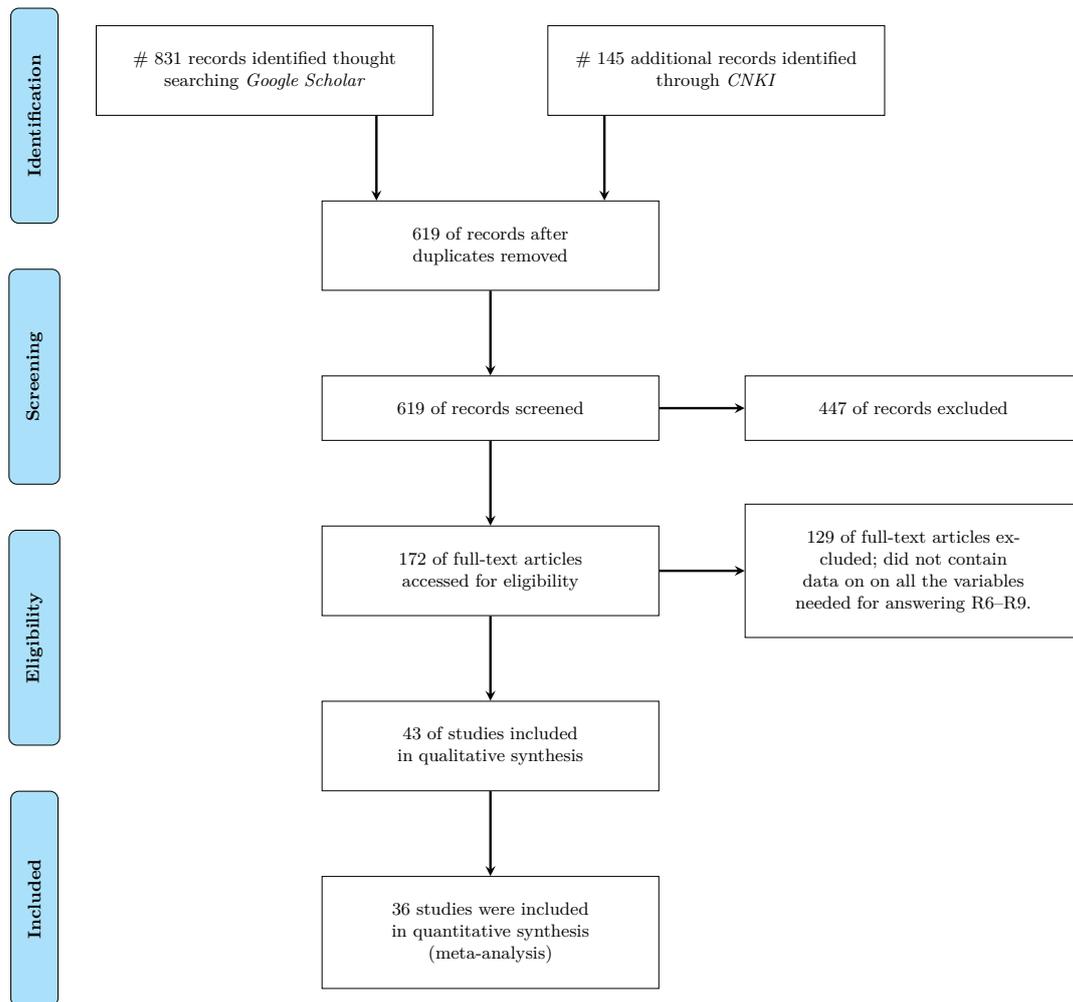
Figure 2. PRISMA Flow Diagram for selecting studies for answering R6–R9.

Table 1: Overview of studies used for calculating the TBV_1-distribution.

					Study design			Sample		Subgroups				TBV_1-distribution (%)			
ID	Author(s) + publication year	IP	Period	Country	Target population	Study base	Method*	Size	Mean age	T	B	V	∅	t	b	v	∅
1	Goodreau et al. (2005)	I	N/A	Peru	MSM	Multiple	Multiple	254	N/A	45	98	110	1	17.8	38.7	43.5	0.4
2	Peinado et al. (2007)	I	2002	Peru	MSM	MSM venues	Multiple	2655	24 (18–71)[†]	1127	1095	432	1	42.5	41.3	16.3	0.0
3	Moskowitz et al. (2008)	I	N/A	International	Gay men	MSM Website	SS	145	32.8 (8.9)	27	50	68	0	18.6	34.5	46.9	0.0
4	Guadamuz et al. (2011)	I	2005	Thailand	MSM	MSM venues	VDTS	2049	24.8	651	628	770	0	31.8	30.6	37.6	0.0
5	Zheng et al. (2012)	I	2007	China	Gay men	MSM websites	SSS	220	23.5 (4.7)	60	76	82	2	27.5	34.9	37.6	0.9
6	Zheng et al. (2013)	I	N/A	China	Gay men	MSM websites	SSS	447	24.7 (5.7)	91	205	151	0	20.4	45.9	33.8	0.0
7	Starks et al. (2013)	I	2008–2009	USA	Gay and bisexual men	Community	CS	683	37.9 (13.1)	232	157	294	0	34.0	23.0	43.0	0.0
8	Clark et al. (2013)	I	2007	Peru	MSM	Multiple	Multiple	532	28 (23–35)[‡]	170	179	183	0	32.0	33.6	34.4	0.0
9	Zheng & Zheng (2016)	I	SL	China	Gay and bisexual men	MSM websites	SSS	484	23 (5.7)	97	235	152	0	20.0	48.6	31.4	0.0
10	Zhang et al. (2018)	I	SL	China	Gay and bisexual men	MSM websites	SSS	334	24.99 (6.37)	54	191	99	0	15.7	55.5	28.8	0.0
11	Wen & Zheng (2019)	I	2018–2018	China	Gay and bisexual men	MSM websites	SSS	483	28.55	125	147	211	0	25.9	43.7	30.4	0.0

† median and range; ‡ median and IQR.

∗ VDTS = Venue Day Time Sampling; SSS = Self Selected Sample; CS = Convenience sample; TLS = Time Location Sampling; RDS = Respondent Driven Sampling; SBS = Snowball Sampling; SCS = Stratified Cluster Sampling.

Percentages may not add up to 100% due to rounding errors.

Appendix B2

Table 2: Overview of studies used for calculating the TBV$_P$-distribution.

			Study design					*Sample*		*Subgroups*				*TBV$_P$-distribution (%)*			
ID	Author(s) + publication year	IP	Period	Country	Target population	Study base	Method[*]	Size	Mean age	T	B	V	∅	t	b	v	∅
1	Wei & Raymond (2011)	P	2008	USA	MSM	MSM venues	TLS	386	35.3 (9.7)	144	82	160	0	37.3	21.2	41.5	0.0
2	Moskowitz & Hart (2011)	P	2008	N/A	Gay and bisexual men	Craigslist.org	SS	429	40.88 (12.03)	148	144	137	0	34.5	33.6	31.9	0.0
3	Zhou et al. (2013)	P	2010	China	MSM	Clinic	RDS	491	24 (18–65)	112	62	317	0	22.8	12.6	64.6	0.0
4	Wang et al. (2015)	P	2012–2013	China	MSM	Multiple	SBS	397	26 (median)	91	113	192	1	23.0	28.5	48.5	0.0
5	Tang et al. (2016b)	P	2014	China	MSM	MSM websites	SSS	1424	N/A	524	621	279	0	36.8	43.6	19.6	0.0
6	Swift-Gallant et al. (2018)	P	2015	International	Gay and straight men	LGBT events	CS	271	33.55 (14.2)	43	100	128	0	15.9	36.9	47.2	0.0
7	Xu et al. (2018)	P	2014–2015	China	MSM	MSM websites	SSS	1100	24.79 (6.10)	283	681	136	0	25.7	61.9	12.4	0.0
8	Ye & Giri (2018)	P	2013–2014	China	MSM	Multiple	CS	1245	31.60 (9.02)	565	337	332	11	45.8	27.3	26.9	0.9
9	Moskowitz & Garcia (2019)	P	2015	International	Gay and bisexual men	MSM websites	SSS	169	39.14 (11.65)	66	64	39	0	39.1	37.9	23.1	0.0

† median and range; ‡ median and IQR.

∗ VDTS = Venue Day Time Sampling; SSS = Self Selected Sample; CS = Convenience sample; TLS = Time Location Sampling; RDS = Respondent Driven Sampling; SBS = Snowball Sampling; SCS = Stratified Cluster Sampling.

Percentages may not add up to 100% due to rounding errors.

Table 3: Overview of studies used for calculating the TBV_B-distribution.

ID	Author(s) + publication year	IP	Period	Country	Target population	Study base	Method*	Size	Mean age	T	B	V	∅	t	b	v	∅
1	Chmiel et al. (1987)	B	1984–1985	USA	Gay and bisexual men	Multiple	CS	4943	N/A	653	943	3336	11	13.2	19.1	67.6	0.2
2	Valdiserri et al. (1988)	B	1986–1987	USA	Gay and bisexual men	Multiple	Multiple	955	N/A	105	76	774	0	11.0	8.0	81.0	0.0
3	Van Griensven et al. (2005)	B	2003	Thailand	MSM	MSM venues	VDTS	1121	26.9	525	197	376	3	47.8	17.9	34.2	0.3
4	Zunt et al. (2006)	B	N/A	Peru	MSM	Multiple	CS	2655	24.1 (8.9)	1127	1095	432	1	42.5	41.3	16.3	0.0
5	Mao et al. (2008)	B	2001–2004	Australia	MSM	Community	CS	1307	N/A	331	175	801	0	25.3	13.4	61.3	0.0
6	Ruan et al. (2009)	B	2007	China	MSM	Multiple	Multiple	284	27 (24–32)	63	22	199	0	22.2	7.7	70.1	0.0
7	Jameson et al. (2010)	B	2001–2006	USA	MSM	Clinic	CS	3828	N/A	928	455	2445	0	24.2	11.9	63.9	0.0
8	Wei & Raymond (2011)	B	2008	USA	MSM	MSM venues	TLS	386	35.3 (9.7)	117	55	214	0	30.3	14.2	55.4	0.0
9	Moskowitz & Hart (2011)	B	2008	N/A	Gay and bisexual men	Craigslist.org	SS	429	40.88 (12.03)	100	125	204	0	23.3	29.1	47.6	0.0
10	Wang et al. (2012)	B	2009	China	MSM	Multiple	SBS	3227	27.4 (7.8)	983	621	1623	0	30.5	19.2	50.3	0.0
11	She et al. (2013)	B	2008	China	MSM/M	Multiple	SBS	1390	28.3 (8.6)	353	245	792	0	33.7	6.9	59.4	0.0
12	She et al. (2013)	B	2008	China	MSMW	Multiple	SBS	303	31.4 (8.7)	102	21	180	0	25.4	17.6	57.0	0.0
13	Zhang et al. (2013a)	B	2011	China	MSM	Multiple	SBS	191	30.4 (11.3)	57	60	74	0	29.8	31.4	38.7	0.0
14	Cao et al. (2013)	B	2011	China	MSM	Multiple	Multiple	266	N/A	80	67	119	0	30.1	25.2	44.7	0.0
15	Dai et al. (2013)	B	2010	China	MSM	Multiple	Multiple	1472	28	478	659	335	0	32.5	44.8	22.8	0.0
16	Vansintejan et al. (2013)	B	2008	Belgium	MSM	Multiple	Multiple	1830	35 (12)	191	123	1516	0	10.4	6.7	82.8	0.0
17	Wang et al. (2014)	B	2011–2012	China	MSM	Multiple	SBS	402	28.48 (10.5)	92	115	195	0	22.9	28.6	48.5	0.0
18	Xu et al. (2014)	B	2011	China	MSM	Multiple	SBS	625	N/A	322	181	122	0	29.0	19.5	51.5	0.0
19	Dang et al. (2015)	B	2014	China	MSM	Multiple	Multiple	294	28.64 (0.438)	114	68	112	0	38.8	23.1	38.1	0.0
20	Zhang et al. (2015)	B	2009–2010	China	MSM	Multiple	SBS	475	29 (25–35)	164	98	213	0	36.9	23.3	39.8	0.0
21	Zhang et al. (2015)	B	2010–2011	China	MSM	Multiple	SBS	352	27 (23–33)	130	82	140	0	39.7	25.6	34.7	0.0
22	Zhang et al. (2015)	B	2011–2012	China	MSM	Multiple	SBS	403	28 (24–34)	160	103	140	0	34.5	25.6	34.7	0.0
23	Sandfort et al. (2015)	B	2011–2013	South Africa	MSM	Multiple	RDS	480	24.5 (5.3)	142	170	165	3	29.8	35.6	34.6	0.6
24	Zhang et al. (2016)	B	2012	China	MSM	Multiple	RDS	675	29.23 (7.01)	180	132	363	0	26.7	19.6	53.8	0.0
25	McLean et al. (2016)	B	2012–2013	Peru	MSM	Clinic	CS	366	23 (16–62)	101	69	196	0	27.6	18.9	53.6	0.0
26	Tang et al. (2016a)	B	2014	Peru	MSM	Multiple	SSS	1008	N/A	314	425	265	4	31.3	42.3	26.4	0.4
27	Zheng et al. (2017)	B	2016	China	MSM	Multiple	CS	300	20.6 (1.6)	111	151	38	0	37.0	50.3	12.7	0.0
28	Duan et al. (2017)	B	2015	China	MSM	Multiple	Multiple	1935	29.3 (8.4)	986	631	318	0	51.0	32.6	16.4	0.0
29	Kojima et al. (2017)	B	2013–2014	Peru	MSM + TGW	Clinic	PS	312	29 (23.3–37.4)	87	68	157	0	27.9	21.8	50.3	0.0
30	Galea et al. (2017)	B	2012–2013	Peru	MSM	Multiple	Multiple	341	26.91 (6.07)	108	132	99	2	31.9	38.9	29.2	0.6

† median and range; ‡ median and IQR.

∗ VDTS = Venue Day Time Sampling; SSS = Self Selected Sample; CS = Convenience sample; TLS = Time Location Sampling; RDS = Respondent Driven Sampling; SBS = Snowball Sampling; SCS = Stratified Cluster Sampling.

Percentages may not add up to 100% due to rounding errors.

Table 3: Overview of studies used for calculating the TBV$_B$-distribution.

			Study design					Sample		Subgroups				TBV$_B$-distribution (%)			
ID	Author(s) + publication year	IP	Period	Country	Target population	Study base	Method*	Size	Mean age	T	B	V	∅	t	b	v	∅
31	Xu & Zheng (2018)	B	N/A	China	MSM	MSM websites	SSS	228	24.64 (6.20)	48	60	120	0	21.1	26.3	52.6	0.0
32	Yang et al. (2018)	B	2013	China	YMSM	Multiple	Multiple	373	22.1 (2.22)	79	90	204	0	21.2	24.1	54.7	0.0
33	Lei et al. (2018)	B	2014	China	MSM	NGO	CS	601	28.2 (6.6)	164	92	345	0	27.3	15.3	57.4	0.0
34	Swift-Gallant et al. (2018)	B	2015	International	Gay and straight men	LGBT events	CS	253	34.33 (14.2)	69	108	76	0	27.3	42.7	30.0	0.0
35	Bjekić et al. (2018)	B	2014	Serbia	MSM	Clinic	CS	319	N/A	101	77	141	0	31.7	24.1	44.2	0.0
36	Wang et al. (2019)	B	2017	China	YMSM	Multiple	RDS	150	20	42	60	48	0	28.0	40.0	32.0	0.0
37	Jin et al. (2019)	B	2017	China	MSM	Multiple	SSS	879	28 (24 – 34)	295	252	332	0	33.6	28.7	37.8	0.0

† median and range; ‡ median and IQR.

∗ VDTS = Venue Day Time Sampling; SSS = Self Selected Sample; CS = Convenience sample; TLS = Time Location Sampling; RDS = Respondent Driven Sampling; SBS = Snowball Sampling; SCS = Stratified Cluster Sampling.

Percentages may not add up to 100% due to rounding errors.

Appendix B4

Table 4: Normalised average distributions separated by measure type.

	$tops_X$	$bottoms_X$	$versatiles_X$
TBV$_I$-distribution	26.0	39.1	34.9
Asia	23.1	44.2	32.7
Latin America	30.7	37.9	31.4
the West	†	†	†
TBV$_P$-distribution	31.2	33.7	35.1
Asia	30.8	34.8	34.4
Latin America	†	†	†
the West	31.7	32.4	35.9
TBV$_B$-distribution	29.4	24.8	45.9
Asia	31.8	24.9	43.3
Latin America	32.2	32.6	35.2
the West	20.6	18.1	61.2

† could not be determined because not enough studies were available.

Numbers indicate percentages.

Percentages may not add up to 100% due to rounding errors.

Appendix C

Table 5: Overview of studies used for answering R6–R9.

ID	Author(s) + publication year	IP	Period	Country	Target population	Study base	Method[*]	Size	Mean age	T	B	V	∅	Ho	Bi	He	∅
								Sample		*Subgroups*				*Sexual orientation*			
1	Van Griensven et al. (2005)	B	2003	Thailand	MSM	MSM venues	VDTS	1121	26.9	525	197	376	3	852	–	–	269
2	Zunt et al. (2006)	B	N/A	Peru	MSM	Multiple	CS	2655	24.1 (8.9)	1127	1095	432	1	1438	936	273	8
3	Lama et al. (2010)	B	2002–2003	Peru	MSM	Multiple	Multiple	2703	24 (20–30)[‡]	1132	1095	432	44	1438	938	276	51
4	Feng et al. (2010)	B	2007	China	MSM	Multiple	SBS	513	24 (16.8–44.5)[‡]	226	164	113	10	347	130	36	0
5	Grov et al. (2010)	B	2006	USA	MSM	LGBT events	CS	1065	37.9 (12.5)	354	270	397	44	950	100	3	12
6	Bai et al. (2011)	B	2008	China	MSM	Multiple	Multiple	280	28.1 (8.8)	87	39	129	25	165	115	0	0
7	Guadamuz et al. (2011)	B	2005	Thailand	MSM	MSM venues	VDTS	2049	24.8	681	776	418	174	989	309	419	332
8	Wei & Raymond (2011)	B	2008	USA	MSM	MSM venues	TLS	386	35.3 (9.7)	117	55	214	0	345	38	3	0
9	Fan et al. (2012)	B	2009	China	MSM	Multiple	RDS	500	29 (18–71)[†]	133	66	291	10	346	147	5	2
10	Chen et al. (2012)	B	2010	China	MSM	Multiple	SBS	218	30.9 (10.7)	59	89	67	3	161	55	2	0
11	Zhang et al. (2012)	B	N/A	China	MSMW	MSM venues	SCS	473	34.1	178	77	213	5	247	172	30	24
12	Perez-Brumer et al. (2013)	B	2007	Peru	MSM	Multiple	CS	560	28 (23–35)[‡]	170	179	183	28	198	60	154	148
13	Hu et al. (2013)	B	2010–2011	China	MSM	Multiple	Multiple	671	30.1 (7.9)	122	112	409	28	443	201	9	18
14	She et al. (2013)	B	2008	China	MSMO	Multiple	SBS	1390	28.3 (8.6)	353	245	792	0	1059	285	10	36
15	She et al. (2013)	B	2008	China	MSMW	Multiple	SBS	303	31.4 (8.7)	102	21	180	0	84	194	9	16
16	Cai et al. (2014)	B	2010	China	MSMW	Multiple	RDS	649	28.3 (7.0)	269	156	177	47	454	145	–	50
17	Hu et al. (2014)	B	2010	China	MSM	Multiple	RDS	500	28 (9.4)	208	157	118	17	275	225	0	0
18	Dai et al. (2014)	B	2009–2010	China	MSM	Multiple	CS	1275	26.47	304	206	743	22	974	301	0	0
19	Wang et al. (2014)	B	2011–2012	China	MSM	Multiple	SBS	402	28.48 (10.5)	92	115	195	0	316	86	0	0
20	Huang et al. (2015)	B	2012	Taiwan	MSM	MSM venues	CS	1208	28.6 (6.4)	268	224	607	109	969	182	32	25
21	Zhang et al. (2015)	B	2009–2010	China	MSM	Multiple	SBS	475	29 (25–35)[‡]	164	98	213	0	321	–	–	154
22	Zhang et al. (2015)	B	2010–2011	China	MSM	Multiple	SBS	352	27 (23–33)[‡]	130	82	140	0	245	–	–	107
23	Zhang et al. (2015)	B	2011–2012	China	MSM	Multiple	SBS	403	28 (24–34)[‡]	160	103	140	0	289	–	–	114
24	Perez-Brumer et al. (2016)	B	2012	Peru	MSM	Community	CS	395	30.5 (9.3)	67	141	182	5	235	66	23	71
25	McLean et al. (2016)	B	2012–2013	Peru	MSM	Clinic	CS	366	23 (16–62)[†]	101	69	196	0	289	74	3	0
26	Qian et al. (2016)	B	2010–2011	China	MSM	Multiple	Multiple	1140	30.52	480	358	262	40	759	355	12	14
27	Song et al. (2016)	B	2009–2015	China	MSM	Multiple	Multiple	5221	30.78 (8.32)	2198	1063	1850	110	3006	2161	53	1
28	Zhang et al. (2016)	B	2012	China	MSM	Multiple	RDS	675	29.23 (7.01)	180	132	363	0	416	219	40	0
29	Clark et al. (2017)	B	2012–2014	Peru	MSM	MSM venues	CS	171	26 (23–31)[‡]	39	49	77	6	108	45	9	9
30	Galea et al. (2017)	B	2012–2013	Peru	MSM	Multiple	Multiple	341	26.91 (6.07)	108	132	99	2	284	0	57	0

† median and range; ‡ median and IQR.

∗ VDTS = Venue Day Time Sampling; SSS = Self Selected Sample; CS = Convenience sample; TLS = Time Location Sampling; RDS = Respondent Driven Sampling; SBS = Snowball Sampling; SCS = Stratified Cluster Sampling.

The same author may appear more than once because different measures were used or the study reports data on more than one sample.

Appendix C (continued)

Table 5 (continued): Overview of studies used for answering R6–R9.

ID	Author(s) + publication year	IP	Period	Country	Target population	Study base	Method*	Size	Mean age	T	B	V	∅	Ho	Bi	He	∅
			Study design					Sample		Subgroups				Sexual orientation			
31	Zheng et al. (2017)	B	2016	China	MSM	Multiple	CS	300	20.6 (1.6)	111	151	38	0	266	19	3	12
32	Clark et al. (2018)	B	2012–2014	Peru	MSM	MSM venues	CS	370	27.5	51	106	196	17	260	80	8	22
33	Yang et al. (2018)	B	2013	China	YMSM	Multiple	Multiple	373	22.1 (2.22)	79	90	204	0	289	69	15	0
34	Xu & Zheng (2018)	I	N/A	China	MSM	MSM websites	SSS	228	24.64 (6.20)	48	60	120	0	175	52	1	0
35	Lei et al. (2018)	B	2014	China	MSM	NGO	CS	601	28.2 (6.6)	164	92	345	0	376	128	49	48
36	Jin et al. (2019)	B	2017	China	MSM	Multiple	SSS	879	28 (24 – 34)‡	295	252	332	0	791	–	–	88

† median and range; ‡ median and IQR.

∗ VDTS = Venue Day Time Sampling; SSS = Self Selected Sample; CS = Convenience sample; TLS = Time Location Sampling; RDS = Respondent Driven Sampling; SBS = Snowball Sampling; SCS = Stratified Cluster Sampling.

The same author may appear more than once because different measures were used or the study reports data on more than one sample.

49

Appendix D

Table 6: Definitions and descriptive statistics of variables used in regression analysis (*N* = 36).

Variable	Definition	Mean (SD)	Range
Dependent variable			
T	Proportion of behaviourally-defined tops in the sample. Calculated as the number of tops in the sample divided by the sample size.	0.314 (0.086)	0.14–0.47
B	Proportion of behaviourally-defined bottoms in the sample. Calculated as the number of bottoms in the sample divided by the sample size.	0.256 (0.098)	0.07–0.50
Primary Independent Variable			
G_p	Proportion of gay-identified men in the sample. Calculated as the total number of men identifying as gay in the sample used in the study divided by the sample size	0.68 (0.14)	0.28–0.90
Control Variables			
A_H	Mean age of participants in all the samples. Formulas described in Hozo et al. (2005) are used to estimate the mean in case the median and range or the median and IQR are reported.	28.1 (3.4)	20.6–37.9
A_W	Mean age of participants in all the samples. Formulas described in Wan et al. (2014) are used to estimate the mean in case the median and range or the median and IQR are reported.	28.8 (3.8)	20.6–39.3
A_L	Mean age of participants in all the samples. Formulas described in Luo et al. (2018) are used to estimate the mean in case the median and range or the median and IQR are reported.	28.7 (3.8)	20.6–39.9
Y_c	Measures when a sample was taken relative to when the first sample in this "sample of samples" was taken. Calculated as (year in which sampling started + year in which sampling ended)/2-2002.5. For articles that did not include the sampling period, year of first submission was used if available. If not available, publication year was used instead.	7.7 (3.6)	0.0–15.5
Moderator			
R_C	Code for the region in which the sample was collected. Three regions are distinguished: Asia, Latin America and 'the West". Asia is coded as '0', Latin America as '1' and the West as '2'.	–	–

Appendix E1

Table 7: *p*-values of pairwise chi-square two sample tests.

H_0 : the TBV_I-distribution in the two samples is the same.

	1	2	3	4	5	6	7	8	9	10	11	Region
1	1.000											.067
2	.000	1.000										.038
3	.761	.000	1.000									†
4	.077	.003	.111	1.000								.130
5	.260	.002	.261	.775	1.000							.394
6	.011	.000	.038	.461	.186	1.000						‡
7	.070	.014	.067	.837	.779	.198	1.000					.828
8	.420	.001	.141	.052	.218	.002	.102	1.000				.874
9	.203	.001	.049	.023	.112	.001	.060	.890	1.000			.763
10	.053	.000	.007	.001	.009	.000	.004	.403	.615	1.000		.229
11	.131	.017	.047	.149	.361	.007	.341	.584	.586	.147	1.000	.849
Global	.321	.004	.203	.438	.815	.049	.613	.503	.339	.050	.710	–

† Not computed as the international sample covered multiple regions.

‡ Not computed as only one studied was conducted in the West.

Numbers refer to the numbers of the studies in Appendix B1.

p-values above .05 have been coloured red.

Appendix E2

Table 8: *p*-values of pairwise chi-square two sample tests.

H_0 : the TBV$_P$-distribution in the two samples is the same.

	1	2	3	4	5	6	7	8	9	Region
1	1.000									.227
2	.116	1.000								†
3	.005	.000	1.000							.000
4	.078	.057	.013	1.000						.126
5	.000	.124	.000	.000	1.000					.075
6	.001	.009	.000	.327	.000	1.000				†
7	.000	.000	.000	.000	.031	.000	1.000			.000
8	.100	.220	.000	.001	.048	.000	.000	1.000		.093
9	.007	.361	.000	.001	.706	.000	.003	.252	1.000	†
Global	.131	.872	.000	.164	.052	.036	.000	.093	.164	–

† Sampling region not uniquely defined.

Numbers refer to the numbers of the studies in Appendix B2.

p-values above .05 have been coloured red.

Table 9: *p*-values of pairwise chi-square two sample tests.

H_0 : the TBV_B-distribution in the two samples is the same.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	Region
1	1.000																																					.318
2	.056	1.000																																				.007
3	.000	.000	1.000																																			.067
4	.000	.000	.000	1.000																																		.010
5	.071	.009	.000	.000	1.000																																	.563
6	.033	.107	.000	.000	.369	1.000																																.000
7	.083	.022	.000	.000	.938	.561	1.000																															.479
8	.012	.000	.008	.000	.670	.097	.473	1.000																														.302
9	.016	.000	.001	.000	.021	.000	.009	.036	1.000																													.121
10	.009	.000	.027	.000	.263	.011	.138	.608	.218	1.000																												.495
11	.000	.000	.001	.000	.202	.167	.198	.257	.000	.038	1.000																											.002
12	.091	.001	.001	.000	.626	.068	.444	.611	.180	.627	.044	1.000																										.138
13	.000	.000	.019	.002	.002	.000	.001	.010	.382	.120	.000	.026	1.000																									.637
14	.002	.000	.033	.000	.036	.000	.014	.129	.517	.584	.002	.222	.585	1.000																								.946
15	.000	.000	.000	.248	.000	.000	.001	.000	.001	.000	.000	.000	.034	.002	1.000																							.003
16	.024	.933	.000	.000	.003	.059	.008	.000	.000	.000	.000	.000	.000	.000	.000	1.000																						.002
17	.019	.000	.001	.000	.024	.000	.011	.039	.997	.223	.000	.196	.346	.500	.001	.000	1.000																					.340
18	.016	.000	.014	.000	.290	.013	.157	.566	.286	.970	.029	.731	.120	.585	.000	.000	.296	1.000																				.461
19	.000	.000	.414	.001	.004	.000	.001	.039	.050	.203	.001	.024	.305	.397	.003	.000	.045	.146	1.000																			.582
20	.000	.000	.284	.000	.009	.000	.003	.071	.096	.330	.002	.052	.381	.574	.002	.000	.088	.254	.949	1.000																		.758
21	.000	.000	.334	.005	.001	.000	.000	.009	.034	.076	.000	.006	.354	.260	.015	.000	.029	.053	.854	.730	1.000																	.422
22	.001	.000	.150	.000	.051	.001	.020	.251	.146	.695	.012	.192	.255	.695	.000	.000	.140	.579	.640	.804	.347	1.000																.770
23	.000	.000	.006	.009	.000	.000	.000	.001	.156	.019	.000	.003	.747	.199	.170	.000	.135	.019	.125	.141	.219	.061	1.000															†
24	.038	.000	.005	.000	.374	.021	.221	.547	.313	.853	.026	.879	.084	.466	.000	.000	.330	.947	.075	.146	.025	.392	.012	1.000														.329
25	.029	.000	.007	.000	.427	.027	.255	.665	.232	.904	.042	.872	.068	.427	.000	.000	.245	.960	.084	.157	.026	.428	.009	.978	1.000													.018
26	.000	.000	.001	.132	.000	.000	.000	.000	.006	.000	.000	.000	.000	.118	.009	.000	.005	.000	.013	.011	.045	.002	.409	.000	.000	1.000												.298
27	.000	.000	.000	.470	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.182	.000	.000	.000	.000	.000	.001	.000	.000	.062	1.000											.000
28	.000	.000	.004	.421	.000	.000	.000	.000	.000	.000	.000	.000	.001	.000	.024	.000	.000	.000	.002	.001	.010	.000	.002	.000	.000	.015	.049	1.000										.000
29	.015	.000	.013	.000	.168	.005	.083	.353	.474	.000	.011	.598	.228	.770	.000	.000	.483	.929	.177	.304	.078	.589	.048	.877	.832	.001	.000	.000	1.000									.078
30	.000	.000	.003	.076	.000	.000	.000	.000	.022	.002	.000	.000	.294	.037	.571	.000	.018	.002	.049	.044	.132	.011	.690	.001	.001	.867	.020	.014	.005	1.000								.588
31	.089	.000	.000	.000	.073	.003	.041	.074	.778	.252	.001	.378	.125	.323	.000	.000	.816	.356	.018	.040	.008	.096	.032	.464	.351	.001	.000	.000	.491	.003	1.000							.188
32	.147	.000	.000	.000	.141	.007	.087	.122	.595	.301	.002	.538	.074	.271	.000	.000	.636	.432	.014	.033	.005	.096	.015	.572	.453	.000	.000	.000	.516	.001	.943	1.000						.152
33	.042	.001	.003	.000	.837	.138	.635	.892	.063	.581	.154	.842	.011	.131	.000	.000	.069	.608	.022	.045	.005	.180	.001	.679	.760	.000	.000	.000	.407	.000	.146	.240	1.000					.087
34	.000	.000	.000	.023	.000	.000	.000	.000	.027	.001	.000	.000	.194	.019	.498	.000	.022	.001	.010	.011	.029	.003	.561	.000	.000	.753	.012	.002	.003	.728	.003	.001	.000	1.000				‡
35	.001	.000	.069	.000	.032	.000	.012	.141	.347	.599	.003	.184	.533	.953	.002	.000	.331	.556	.563	.750	.370	.843	.175	.406	.394	.008	.000	.000	.692	.036	.202	.173	.124	.015	1.000			†
36	.000	.000	.001	.017	.000	.000	.000	.000	.066	.003	.000	.000	.387	.057	.362	.000	.055	.003	.032	.035	.074	.012	.815	.002	.001	.665	.006	.002	.010	.808	.010	.004	.000	.909	.046	1.000		.069
37	.000	.000	.075	.003	.002	.000	.000	.014	.194	.139	.000	.021	.850	.568	.027	.000	.173	.121	.598	.649	.679	.390	.569	.073	.066	.093	.000	.002	.205	.255	.058	.036	.011	.107	.617	.241	1.000	.719
Global	.004	.000	.022	.000	.045	.001	.018	.141	.597	.608	.002	.271	.539	.986	.001	.000	.583	.632	.314	.478	.196	.632	.175	.532	.479	.007	.000	.000	.829	.029	.408	.350	.154	.017	.899	.050	.484	–

† Not computed as the international sample covered multiple regions.

‡ Not computed as only one studied was conducted in the West.

Numbers refer to the numbers of the studies in Appendix B3.

p-values above .05 have been coloured red.

Appendix E4

Table 10: *p*-values of pairwise chi-square two sample tests.

H$_0$: the TBV$_P$-distribution in the two samples is the same.

Measure Type		0	1	2	Global
	0	1.000			.762
I	1	.430	1.000		.430
	2	†	†	†	†
	0	1.000			.986
P	1	†	†		†
	2	.902	†	1.000	.956
	0	1.000			.883
B	1	.382	1.000		.253
	2	.038	.001	1.000	.104

† Sampling region not uniquely defined.

Numbers refer to the numbers of the studies in Appendix B2.

'0' indicates Asia, '1' Latin America and '2' the West.

p-values above .05 have been coloured red.

Appendix F

Table 11: Results from multivariate linear regression of T and B on independent variables.

Specification	Dependent Variable	Constant	A_L	Y_c	G_p	$R_c \times A_L$	$R_c \times Y_c$	$R_c \times G_p$
1T	T	**.464** (.001)	−.002 (.554)	**−.011** (.013)	.006 (.951)	–	–	–
1B	B	**.595** (.000)	**−.013** (.003)	−.003 (.477)	.079 (.491)	–	–	–
2T	T	**.503** (.004)	−.003 (.499)	**−.011** (.017)	−.0.002 (.986)	–	–	−.082 (.124)
2B	B	**.570** (.003)	**−.012** (.015)	−.002 (.665)	−.063 (.640)	–	–	.098 (.092)
3T	T	**.402** (.012)	–	**−.010** (.017)	–	–	**−.008** (.039)	–
3B	B	**.605** (.000)	**−.013** (.006)	–	–	**.003** (.043)	–	–

Values are non-standardised coefficients. Values underneath are *p*-values. Statistically significant results are indicated in boldface.