Running head: REDUCING SEXUAL PREJUDICE

Interventions to Reduce Sexual Prejudice:

A Study-Space Analysis and Meta-Analytic Review

Sebastian E. Bartoş, Israel Berger, and Peter Hegarty

University of Surrey

Author Note

Sebastian E. Bartoş, Israel Berger, and Peter Hegarty, School of Psychology, University of Surrey, Guildford, Surrey, UK.

We wish to thank Ellie Brodie for her patient help in testing the reliability of inclusion criteria; and Robert Nash, Ilka Gleibs, and Y. Gavriel Ansara for their helpful feedback and methodological suggestions.

This study was conducted in partial fulfillment of the requirements for a doctoral degree by the first author, under the supervision of the third, and with the assistance of the second.

Correspondence concerning this article should be addressed to Sebastian E. Bartoş, School of Psychology, University of Surrey, Guildford, Surrey, GU2 7XH, UK. E-mail: s.bartos@surrey.ac.uk

Interventions to Reduce Sexual Prejudice:

A Study-Space Analysis and Meta-Analytic Review

Abstract

Sexual prejudice is an important threat to the physical and mental wellbeing of lesbians, gay men, and bisexual people. Therefore, we reviewed the effectiveness of interventions designed to reduce such prejudice. A study-space analysis was performed on published and unpublished papers from all over the world to identify well-studied and underexplored issues. Most studies were conducted with North American undergraduates, and were educational in nature. Dissertations were often innovative and well-designed, but were rarely published. We then performed meta-analyses on sets of comparable studies. Education, contact with gay people, and combining contact with education had a medium-size effect on several measures of sexual prejudice. The manipulation of social norms was effective in reducing anti-gay behaviour. Other promising interventions, such as the use of entertainment media to promote tolerance, need further investigation. More research is also needed on populations other than American students, particularly groups who may have higher levels of sexual prejudice.

Keywords: homophobia/heterosexism, sexual prejudice, education, contact hypothesis, meta-analysis, study-space analysis

Interventions to Reduce Sexual Prejudice:
A Study-Space Analysis and Meta-Analytic Review

Since homosexuality was depathologised 40 years ago, psychologists have been ethically committed to reducing sexual prejudice (Conger, 1975). Ethics codes continue to urge psychologists to strive against all forms of prejudice, including that based on sexuality (American Psychological Association, 2008). Empirical evidence in favour of this ethical commitment has also grown considerably after the American Psychiatric Association decided in 1973 that homosexuality would not be considered a mental disorder. Lesbians, gay men, and bisexual (LGB) people have poorer health and wellbeing outcomes than heterosexual people (Cochran, 2001; Meyer, 2007). LGB people are one-and-a-half times as likely to suffer from substance-related, mood, or anxiety disorders as their heterosexual peers (Cochran, Sullivan, & Mays, 2003); and are twice as likely to attempt suicide (King et al., 2008). Experiencing sexual prejudice is strongly associated with poor mental health outcomes (Mays & Cochran, 2001). Most recently, longitudinal studies have confirmed that sexual prejudice is indeed the cause of LGB people's health disadvantage (Hatzenbuehler, Nolen-Hoeksema, & Erickson, 2008; Hatzenbuehler, McLaughlin, Keyes, & Hasin, 2010). To ensure the wellbeing of LGB people, sexual prejudice must be reduced.

However, the effectiveness of psychological interventions in countering sexual prejudice is not well understood. The individual and small-group interventions proposed by psychologists are often seen as ancillary to large-scale social reform: as Morin (1991, p. 245) put it, "the change of society will help more people than an army of psychologists working with them one by one" (see also Ehrlich, 1973). Literature reviews have often emphasised the methodological weaknesses of psychological studies in this domain, and they have consequently shied away from drawing conclusions on the effectiveness of such interventions (e.g., Croteau & Kusek, 1992; Tucker & Potocky-Tripodi, 2006). *The present review aims to*

assess the achievements and shortcomings of psychological science in its pursuit of effective techniques to reduce sexual prejudice. After a series of conceptual clarifications, we proceed to map the well-explored and neglected aspects of research in this area. We then describe the interventions that have been employed to reduce sexual prejudice, and we assess their effectiveness.

Conceptual Clarifications

People with same-gender attractions and relationships have been facing rejection throughout history. In pre-colonial Zimbabwe, for example, liaisons between men were often treated as a misdemeanour (Epprecht, 1998), while in Imperial Korea, same-gender relationships were seen as being at odds with the existing social and religious order (Lim & Johnson, 2001). Even where such relationships are accepted, they are often regarded as ancillary to the heterosexual family (see e.g. Kendall, 1998, on women in precolonial Lesotho; and Dover, 2002, on men in Ancient Greece). Both same-gender love and the rejection thereof have been labelled and described in a number of ways; 19th century Western psychiatry coined the term *homosexuality* to conceptualise same-gender sexual attraction (and, to a lesser extent, behaviour and identity; Sell, 1997). *Homophobia* was later introduced to designate the rejection of homosexuality (Smith, 1971; Weinberg, 1972), and it has now become widely used and accepted (Hegarty, 2006).

The notion of *homophobia* (Smith, 1971; Weinberg, 1972)[1] emerged in the 1970s, when social sciences reconsidered same-gender sexuality; in particular, the idea that the gay community was a marginalised subculture emerged to counter the previous dominant model of homosexuality as a disease (Maher et al., 2009; Pettit, 2011). However, *homophobia* has been criticised for a number of reasons throughout its four-decade history. First, Herek (2004) argues the term is inaccurate: *-phobia* is misleading in this context, since it clusters a social attitude with anxiety disorders. Second, as *homosexuality* with no other specifications is often

used to refer to men, *homophobia* may also focus attention on gay men and render lesbians invisible (Plummer, 1981; Kitzinger, 1987; Herek, 2004). Third, speaking of *homophobia* and *homophobes* focuses research on psychological aspects, concealing social and political implications[2] (Plummer, 1981; Kitzinger, 1987). Given the controversy around *homophobia*, in this paper we opted for the more neutral *sexual prejudice* (see e.g., Herek, 2004).

Herek (2004, 2007) distinguished three levels or facets of the rejection of LGB people: the individual, the socio-political, and the cultural (see e.g., Esses, Semenya & Stelzl, 2004, for a similar approach to other forms of prejudice). He termed these three levels *sexual prejudice*, *heterosexism*, and *sexual stigma*, respectively. Adam (1998) had previously remarked that studies on these three levels often have different philosophical assumptions and are "characterised by considerable disciplinary insularity" (p. 387). As for containing prejudice, students of heterosexism and sexual stigma typically focus on large-scale social and cultural change, while prejudice researchers are concerned with designing individual and small-group interventions. As we explained at the beginning of this paper, the usefulness of the small-scale psychological interventions is unclear in a time of major social and institutional change (see Tucker & Potocki-Tripodi, 2006, for a recent review). Therefore, in this review, we focus on sexual prejudice rather than heterosexism or sexual stigma.

In this paper, we use the acronym LGB (for lesbian, gay, and bisexual) when referring to the targets of sexual prejudice. However, our original sources were often vague: bisexual people are sometimes implied, but rarely named explicitly; other forms of sexuality are only now becoming visible to psychology (see e.g., the *Psychology & Sexuality* special issue on asexuality, volume 4, issue 2, 2013). While it is customary to see transgender people added to this list (esp. in the acronym LGBT and its variations), our article focuses strictly on prejudice based on sexuality. Transgender people who are also LGB may experience sexual prejudice in addition to cisgenderism (i.e., prejudice related to their self-designated gender;

Ansara, 2010).

Previous Reviews

Our systematic review is designed to inform future efforts, within and beyond psychology, to reduce sexual prejudice. When research is not comprehensively integrated, practitioners and policy makers have difficulty using it (Higgins & Green, 2008). The volume of the literature and its inconsistent results often frustrate policy makers, affecting both the prestige and the funding of psychological research (Schmidt & Hunter, 2003), and raising the risk of running unnecessary studies on questions that could be addressed by reassessing previous research (Mulrow, 1994).

There are only two reviews of sexual-prejudice interventions, and neither is comprehensive. Stevenson's (1988) synthesis was thorough but it is now outdated. More recently, Tucker and Potocky-Tripodi (2006) found that no intervention strategy for reducing sexual prejudice was adequately supported by the literature. They speculated that the reticence of funding bodies might have hindered research on sexual prejudice. However, Tucker and Potocky-Tripodi only considered published articles from a ten-year period; the seventeen studies included in their review represent little more than one tenth of the relevant literature (see below our own sample of 157 studies).

Syntheses of prejudice research in general also address sexual prejudice, but with understandable concision. Paluck and Green's (2009) review of over 800 prejudice-reducing interventions did not differentially discuss research on specific types of prejudice. Therefore, this paper did not allow the reader to appraise whether a strategy described as effective was specifically tested in the case of sexual prejudice. Moreover, this impressively broad review still covered less than one-third of the available literature on reducing sexual prejudice.

More focused reviews are available, but they typically confine themselves to such specific interventions as panel discussions (Chonody, Siebert, & Rutledge, 2009; Croteau &

Kusek, 1992) or gay-straight alliances (Hansen, 2007). Pettigrew and Tropp's (2006) meta-analysis on the *contact hypothesis* (Allport, 1954) also found that contact with LGB people reduced heterosexual people's sexual prejudice. Interestingly, the effect of contact on sexual prejudice was slightly stronger than on other forms of prejudice, such as racism. Smith, Axelton, and Saucier (2009) performed a meta-analysis exclusively on contact and sexual prejudice, and they also confirmed the effectiveness of this approach. However, no such review explored other methods of reducing sexual prejudice.

Therefore, a broad synthesis is needed. Meta-analyses and systematic reviews have become standard practice for disentangling the medical literature (Higgins & Green, 2008). These methods have also proved valuable for social psychology (Pettigrew & Tropp, 2006). More recently, study-space analysis has been proposed for identifying underexplored key-issues (Malpass et al., 2008). In the case of sexual prejudice reduction, the volume and diversity of the literature suggest the need for research integration.

The Present Review

The aim of this review is to assess practical strategies to reduce sexual prejudice. We consider studies regardless of disciplinary boundaries (e.g., intervention papers in educational and social work journals), theoretical underpinnings, and study design and setting (e.g., laboratory experiments and classroom interventions). However, as our focus is on intervention studies, we excluded correlational research. We also excluded all qualitative research, the methods and results of which are usually not commensurable with those of quantitative studies. Through these exclusions and restrictions we aimed to review a meaningful and coherent body of studies in a feasible way.

The present study draws on three complementary approaches to assessing and integrating scientific evidence. First, a thorough search of the literature was performed, in accordance with the Cochrane guidelines for systematic reviews (Higgins & Green, 2008).

Second, a study space analysis was performed in order to identify the issues that have been satisfactorily addressed by these studies and the issues that need further research. Malpass et al. (2008) proposed study space analysis as a procedure for "identifying regions of concentration and inattention" (p. 794) in a field of research. A study space is a matrix in which lines and columns represent study characteristics, e.g., whether the research was experimental, or whether the participants were students. Each entry of the matrix represents the number of studies that exhibit the corresponding pair of characteristics, e.g., how many studies were experiments performed on students. An inspection of the study-space matrix can indicate the issues that have been neglected, as the corresponding cells will have visibly low counts; and inferential statistics (e.g., χ^2 tests) can elucidate whether the distribution of the studies across the study space is uneven (see e.g., Memon, Meissner, & Fraser, 2010). Systematic reviews can sometimes point out underexplored issues (e.g., Paluck & Green, 2009), but study-space analyses allow for quantification and increased rigour.

Third, we performed meta-analytic reviews on clusters of studies that used a similar approach to reduce sexual prejudice. Effect sizes were computed for all reports that provided sufficient information. However, following the advice of Borenstein, Hedges, Higgins, and Rothstein (2009), we only computed summary effect sizes when the studies within a cluster were both sufficiently similar and numerous. In all other cases, we reported and discussed the effect sizes of individual studies.

The Systematic Search of the Literature

Literature Search

Our search for relevant literature followed the recommendations of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins & Greeen, 2008). These guidelines demand a systematic, quasi-exhaustive strategy for collecting both published and unpublished reports; a transparent, a priori protocol for selecting the relevant studies; and a

reliable coding scheme for recording the designs and results of those studies.

We first generated potential keywords for our literature search through brainstorming and by consulting theoretical papers on sexual prejudice (e.g., Herek, 2004). Two lists were compiled: a series of terms representing *intervention* and an array of words and phrases representing *reactions to homosexuality* (see Table 1). Boolean operators and wildcards were employed to facilitate the use of these keywords in search engines.

Ten electronic databases were searched for relevant reports: PsycINFO, Medline, Psychology and Behavioural Sciences Collection, International Bibliography of Social Sciences, Applied Social Sciences Index and Abstracts, Sociological Abstracts, ScienceDirect, Scopus, ERIC, and ISI Web of Knowledge. All English-language reports were retrieved that contained at least one intervention-related phrase and one sexual-prejudice-related phrase in the title, abstract or keywords. In order to retrieve more recent studies, we repeated these searches on 25 March 2012. Studies published after this date were not included.

We made efforts to retrieve relevant studies not identified by searching the databases. The reference lists of several systematic reviews were checked (Hansen, 2007; Paluck & Green, 2009; Pettigrew & Tropp, 2006; Stevenson, 1988; Tucker & Potocky-Tripodi, 2006). Recent volumes of journals likely to publish relevant studies were searched by hand. Specifically, the first author inspected paper copies of the two most recent volumes of the *Journal of Homosexuality* and the *Journal of Sex Research*, and the whole archive of the *Gay and Lesbian Psychology Review/Psychology of Sexualities Review*. Seven additional reports were identified through these supplementary searches.

French and German reports were sought in Persée and at the German National Library and DissOnline, respectively. Moreover, we used French, German, and Spanish versions of our keywords in PsycINFO and Google. We also performed searches in English-language

databases of regional relevance, namely African Journals Online, Central and Eastern Europe Online, and the Indian Citation Index. None of these searches returned any relevant results.

Several strategies were employed to access studies that are unpublished or otherwise difficult to retrieve. First, we performed searches in Google, Google Scholar, Lexis Nexis, and Academia.edu. Second, we searched OpenGrey, a database that indexes unpublished research from across Europe. Third, special attention was granted to theses and dissertations. Although much postgraduate research (especially from the US) is indexed in major databases, we performed supplementary searches in the Index to Theses in Great Britain and Ireland and in the ProQuest Dissertations and Theses Database. Fourth, we consulted the websites of several gay-rights and human-rights organisations: the International Lesbian and Gay Alliance, Stonewall (UK), the National Lesbian and Gay Task Force (US), Global Issues, and the United Nations. The last three strategies returned no relevant results. Except for an unpublished report retrieved through Google, all the other references suggested by these searches were already in our corpus.

We also contacted nineteen key people in the fields of prejudice reduction and sexual prejudice, and asked them to recommend us relevant reports. The list of people to contact was compiled by brainstorming, by consulting relevant handbooks (e.g., Clarke, Ellis, Peel, & Riggs, 2010; Coyle & Kitzinger, 2002), and by searching our own corpus for authors with numerous papers. These experts suggested nine additional reports. One final report was indicated by an anonymous reviewer.

Inclusion Criteria

We defined the boundaries of our review in terms of population, intervention, control, and outcome (PICO; cf. Higgins & Green, 2008). Our specific criteria are described in further detail in the following paragraphs.

Population. Studies were categorized by the age, gender, nationality, and ethnic

composition of their sample; therefore, no study was excluded on such grounds. Only those interventions were included that targeted sexual prejudice in heterosexual people. However, we also included both those studies that had a minority of LGB participants, and those that did not explicitly state their participants' sexuality.

Intervention. Reports were included in the review provided that (a) they described at least one intervention purposefully performed by a person or group, (b) they offered quantitative data reflecting the outcome of that intervention, and (c) the intervention was performed in order to modify reactions to homosexuality. We therefore excluded correlational studies and surveys, but we did not exclude interventions that resulted in an increase in sexual prejudice.

Control. Studies using quantitative methods were included, such as experiments (i.e., comparisons of randomised groups) and quasi-experiments (i.e., comparisons of non-equivalent groups and pretest-posttest studies). As explained in the *Introduction*, we excluded all qualitative research from this review.

Outcomes. Studies with outcome measures that reflected participants' reactions to homosexuality were included. In this context, *homosexuality* could refer to sexual behaviour or desire involving people of the same gender; to individuals and groups to whom such behaviours and desires are attributed (e.g., LGB people, queers); or simply to the term and its individual meanings.

Exclusion Decisions

The database searches returned approximately 40,000 references. The titles and abstracts of these reports were screened based on the inclusion criteria described above (see Figure 1 for a flowchart of the selection process).

Since this screening was performed by the first author alone, we assessed the reliability of the criteria. A batch of 100 articles was compiled using PsycINFO. The first author and a

research assistant independently applied the criteria and rated each article as either *included* or *excluded*. We opted for Gwet's AC_1 statistic over the more popular Cohen's κ because AC_1 gives a better estimate of intercoder agreement when the baseline frequencies of the categories are greatly unequal (Gwet, 2008). In our case, over 90% of the studies were excluded by both coders; if we had used Cohen's method, this would have lead to a substantial overestimation of the probability of random agreements and a subsequent underestimation of the reliability coefficient. Gwet's AC_1 estimates the proportion of random agreements based on binomial probabilities, but it is otherwise identical to Cohen's κ. The inter-rater agreement on exclusion decisions was good, Gwet's $AC_1 = .96$, $SE = .02$, $p < .001$. In a debriefing discussion, the two researchers agreed that the criteria were unambiguous, and that inclusion and exclusion decisions could be effectively made by the first author working alone.

After the literature search, we retained 238 references. Most of these were available through at least one of several academic libraries where we are members. Thirty-one journal articles were obtained through interlibrary loans, six articles were consulted by courtesy of the authors, and five dissertations were purchased. Five reports were deemed irretrievable. Eighty-two reports did not present any relevant intervention and/or outcome, three were dissertations also published as journal articles; and two were duplicates. The resulting corpus comprised 146 reports. The selection process is summarised in Figure 1. The full list of included studies is given in the Online Supplement, Table S1.

The Study Space Analysis

Analytic Strategy

We performed the study space analysis on 146 published and unpublished reports on a total of 159 studies. The aim of this study space analysis was twofold. First, we aimed to describe the studies by looking at the PICO characteristics: the populations sampled, the

interventions tested, the designs employed, and the outcomes examined. We also recorded the publication status of the reports, the year of publication or submission, and whether they received funding.

Second, we explored the associations between study characteristics by cross-tabulating variables to obtain study-space matrices. We then computed Pearson's χ^2s, standardised residuals, and Goodman's γs in order to examine associations between study characteristics. Most tables include cells that are either empty or which have expected values smaller than five. In these instances, χ^2 tests have diminished power, and nonsignificant results should be interpreted with caution (Howell, 2005). We also compared groups of studies on continuous variables such as sample size and mean sample age. None of these continuous variables were normally distributed, all skewness zs > 4.89, ps < .001, and kurtosis zs > 2.94, ps < .01. Therefore, we used nonparametric tests.

Data Coding

We developed a coding scheme in order to systematically extract data about the studies. In addition to information related to the PICO criteria, we also included basic bibliographic data. The variables we constructed in order to extract information from the reports are described in the rest of this section.[3]

The first author coded the studies alone. Fifteen studies (approximately 10% of the corpus) were independently re-coded by the second author to check the reliability of the procedure. For the reasons explained above (under *Exclusion Decisions*), we opted for Gwet's AC_1 coefficient to quantify intercoder agreement on categorical variables. We computed intraclass correlations continuous variables, and a Goodman's γ for our only ordinal measure. Intercoder agreements were good for both categorical and continuous variables (see below, and in the Online Supplement., Tables S2 and S3).

Population. To characterise the sample, we recorded the number of participants, the

proportion of women (between 0 and 1), and the country in which the study took place. We recorded both the average age of the sample, as reported by the authors; and the age group to which the participants belong, classified as children (up to 13 years old), teenagers (13-18), young adults (18-30), adults (30-60), and older adults (older than 60). The presence or absence of data on race and ethnicity was recorded, as well as the proportion of white participants, where available (between 0 and 1). Participants' sexuality was coded either as *all heterosexual*, when heterosexuality was a selection criterion for the study; as *mixed*, when both heterosexual and LGB people participated; or as *unreported*, when this was the case. The intercoder agreement was very good for both continuous (intraclass correlations ranging from .98 to 1, all *p*s < .001) and categorical variables (Gwet's AC_1s ranging from .83 to 1, all *p*s < .001).

Intervention. We classified interventions into fourteen categories. Paluck and Green (2009) described twelve types of intervention to reduce prejudice, although they did not provide a list or comprehensive definitions. Our own operational definitions are given in Table 2. We added a residual category for studies comparing two or more approaches to reducing sexual prejudice, and the cross-over category of contact-plus-education studies. The reliability of classifying the approaches to sexual prejudice reduction was assessed on a sample of 39 reports (approximately one quarter of the corpus). The intercoder agreement was very good, Gwet's $AC_1 = .86$, $p < .001$.

Comparison. To assess research designs, we coded the type of comparison used by each study on a four-point scale. Based on Cook and Campbell's (1979) seminal assessment of experimental and quasi-experimental research, we constructed an ordinal measure of a study's internal validity. Specifically, we coded the type of comparison used by each study on a four-point scale (0 – no comparison; 1 – either pretest-posttest or non-randomised control group; 2 – non-randomised control group with pretest and posttest; 3 – randomly assigned

control group). The intercoder agreement was very good, Goodman's $\gamma = .97$, $p < .001$. The presence of any follow-up test (i.e., post-posttest) was recorded as a separate variable.

Outcomes. We labelled as *attitudinal* all sexual prejudice scales, such as the Attitudes Towards Lesbians and Gay Men scale (ATLG; Herek, 1984). These measures all assessed attitudes, i.e., a *general positive or negative orientation* towards a social object (see e.g., Bohner & Dickel, 2011). When a measure specifically explored behavioural, cognitive, or emotional aspects of prejudice, we classified it as such (see below). Measures of attitudes towards specific issues (e.g., same-gender marriage or employment discrimination) were recorded but not used in this study.

Behavioural measures included not only actual behaviour, but also behavioural intentions. Common examples of behavioural measures included professionals' responses to case vignettes, surveys of intended behaviour, and participants' willingness to help gay people in real-life situations. Verbal behaviour was also classified as a behavioural outcome when participants used speech or writing for a specific end (e.g., to prepare a talk supporting gay rights) rather than to report their own thoughts and feelings.

We classified as *cognitive* all outcome measures of stereotypes and other beliefs about gay, lesbian, and bisexual issues. Knowledge about homosexuality and other measures that explicitly targeted cognition were also included in this category.

Emotional measures included all the instruments assessing participant's feelings towards homosexuality or LGB people. These were typically self-report measures that assessed the extent to which participants felt fear, anger, disgust, or other emotions in response to homosexuality.

Implicit measures attempted to assess participants' reactions without relying on self-report, often in order to bypass their need for favourable self-presentation. Typical implicit measures were implicit associations tests (IAT; see Greenwald et al., 2002) and galvanic skin

responses (GSRs).

Bibliographic information. Each article was identified by the surname of the first author and the year of publication. For unpublished reports, the year of submission or completion was recorded. We also coded the type of the report (journal article, dissertation, conference paper, book chapter, or unclassified research report) and whether it was published. Where applicable, the journal title was recorded. We also noted whether the study was funded.

Population

A total of 19,782 people participated in the 159 studies. The median sample size was of 92 people (range: 18 - 862 participants). Where demographics were reported, participants were mostly female (63%), young (*M* age 22.69), and white (77%). In 8% of the cases, no information was given on participants' gender, and only 56% of the studies described the sample's ethnic composition. One hundred and thirty-eight studies were performed in North America (87%); twelve in Western Europe (8%); four in the Middle East and South Asia (3%); three in Australia (2%); and one in Africa (< 1%). No studies were performed in South America, Eastern Europe, East Asia, or Oceania. Six studies were conducted with teenagers (4%), 18 with adults (11%), and 134 with young adults (84%). No studies were conducted with children or older adults. While the participants' age group was almost always clear, the mean age was only reported in half of the studies. Sample mean ages ranged from 14 to 44 years. One hundred and thirty-nine studies employed undergraduate students as participants (87%). The researchers assembled a confirmed heterosexual sample in only 29% of the studies; 57% of the studies did not report participants' sexuality, while the remaining 14% acknowledged the inclusion of a (usually small) number of LGB participants.

Interventions

We classified the studies according to 12 different types of prejudice-reducing

interventions, following Paluck and Green (2009). Many studies used some form of education as an intervention (n = 63, 40 % of N = 159). Several studies employed LGB guest speakers, and thus combined education with intergroup contact (n = 38, 24 %). Contact with LGB people was also used outside of an educational context in a number of studies (n = 15, 9 %). Other studies examined effects of making tolerance a social norm, either through experts' statements or peers' opinions (n = 22, 14 %), inducing specific emotions (n = 11, 7 %), entertainment media (n = 11, 6 %), priming techniques (n = 7, 4 %) , awareness and suppression (n = 5, 3 %), and accountability (n = 1, < 1 %). Four approaches identified by Paluck and Green (2009) were not represented: cognitive training for children, peer debates, cooperative learning, and the manipulation of social categorisation. Finally, 15 studies (9 %) compared the effectiveness of two or more approaches.

Next, we asked whether studies that examined different forms of prejudice reduction differed in their samples' characteristics. Groups of studies using the same approach to prejudice reduction did not differ in terms of their sample size, Kruskal-Wallis $H(9) = 4.35$, $p > .05$; employment of North American participants, $\chi^2 (9) = 11.13$, $p > .05$; or the proportion of white participants, Kruskal-Wallis $H(7) = 11.22$, $p > .05$. However, different types of interventions had different gender ratios in their samples, Kruskal-Wallis $H (9) = 24.07$, $p < .05$.

Comparison

Recall that the robustness of the study design was operationalised as a four-point ordinal variable ranging from *no comparison (0)* to *comparison of randomised groups (3)*. In 70 studies, participants were randomly assigned to two or more groups (44%); 36 studies had both pretests and non-randomised control groups (23%); 50 studies had either pretests on non-randomised control groups (31 %); and in the remaining 3 studies, only post-intervention data were reported without any term of comparison (2%). Twenty-five studies (16%) reported

follow-up results.

Outcomes

Recall that dependent measures were classified as attitudinal, behavioural, cognitive, emotional, and implicit. Most studies used some form of attitudinal measure (89 % of N = 159). Behavioural, cognitive, and emotional measures were each used in less than one fifth of the studies (16%, 17%, and 18%, respectively). Less than 3% of the studies employed implicit measures, such as implicit associations tests (IAT) and galvanic skin responses (GSRs).

Outcome measures typically referred either to both lesbians and gay men, or to LGB people more generally. Nine studies exclusively dealt with gay men (6%), two studies focused on lesbians (1%), and eight studies compared sexual prejudice directed at men and at women (5%). Only four studies specifically addressed prejudice against bisexual people (3%).

Publication and Funding

The majority of the reports (130 out of 146) were retrieved through searches in electronic databases; the rest were identified through previous reviews (10), key researchers (2), Google searches (2), hand searches of relevant journals (1), and a suggestion from an anonymous reviewer (1). These reports were journal articles (114), unpublished dissertations (30), conference presentations (1), and an unpublished research paper (1). The journal articles appeared in 75 different publications, and were clustered in *Journal of Homosexuality* which published 16 of these papers (14%). No other journal published more than four articles. All dissertations had their abstracts published in *Dissertation Abstracts International*. Almost 15% of the published studies (18) received some form of financial support. While some of the authors of the unpublished dissertations might have received scholarships, we did not find acknowledgements of any other funding.

We examined if published and unpublished studies (n = 124 and 35, respectively) differed on characteristics related to study design. All unpublished studies in this corpus were performed in the US. There were no significant differences between published and unpublished studies in the types of interventions used or in participants' age group, sexuality, or student status; all χ^2s were nonsignificant.

The comparison of published and unpublished studies revealed two unexpected differences. Unpublished studies tended to employ more robust designs than published studies, Goodman's γ = .55, p < .001. Furthermore, 29 % of the unpublished studies followed up on the long-term effectiveness of the intervention, as opposed to only 12 % of the published studies, χ^2 (1) = 5.59, p < .05. Jointly, these two findings suggest the surprising conclusion that unpublished studies are more rigorous in some respects than the published ones.

Finally, we examined the differences between funded and unfunded studies (n = 17 and 142, respectively) on PICO characteristics. Funded studies were more likely to be conducted outside the US, χ^2 (1) = 5.50, p < .05; and to recruit male-only samples, χ^2 (2) = 10.46, p < .01, the standardised residual z = 2.85, p < .01. The design of funded studies was not more robust, Goodman's γ = .37, p > .05. Studies employing different approaches were not equally likely to receive funding, χ^2 (9) = 21.86, p < .01; specifically, contact-plus-education studies were never funded, z = -2.06, p < .05. There were no other meaningful differences between funded and unfunded research in terms of sample size, participants' characteristics, design, or outcome measures, all χ^2s < 3, ps > .05 and Mann-Whitney Zs < 1.96, ps > .05.

Meta-Analytic Reviews

Analytic Strategy

In order to appraise the effectiveness of interventions for reducing sexual prejudice, we used meta-analytic tools. Effect sizes were computed for each study that provided

sufficient information. We grouped studies based on intervention strategies and outcome measures. We computed a summary effect size for every such group of studies, unless there were further reasons to discuss the studies separately. All meta-analytic procedures followed the guidelines of Borenstein et al. (2009) and Field and Gillet (2010).

We proceeded in four stages. First, Cohen's *d* was computed as a measure of effect size for each study. We aimed to compute *d*s using the best available information. When means and standard deviations were not available, we used transformations of the statistical values provided in the report (see Borenstein et al., 2009). If the number of participants in different groups was not provided, the groups were assumed to be equal in size. When a study had more than one type of dependent variable, we computed effect sizes for each outcome. Effect sizes were computed with the online calculator provided by Lipsey and Wilson (http://gunston.gmu.edu/cebcp/EffectSizeCalculator/d/d.html). In situations not covered by this website, we applied Borenstein's et al. (2009) formulae by hand. Our computations were always based on posttest scores; follow-up results (post-posttests) were rare (16 % of the studies) and they were not used in this meta-analysis. In accordance with conventional benchmarks, effect sizes were interpreted as small ($d < 0.30$), medium ($0.30 < d < 0.50$), or large ($d > 0.50$).

Second, we computed the summary effect size[4], relying on a random-effects model. We opted for random effects over fixed effects due to the great variety in our database of studies. While the interventions often relied on similar principles, each team of researchers used a customised set of procedures and interventions. Therefore, we found it more reasonable to assume that the effect sizes reflected a variety of true effects (random effects meta-analysis), rather than all being approximations of a single true effect (fixed effect meta-analysis). All computations were performed using the IBM SPSS Statistics syntax provided by Field and Gillet (2010; http://www.statisticshell.com/meta_analysis/how_to_do_a_meta_analysis.html).

Third, the heterogeneity of the effect sizes was assessed. To achieve this, we computed the weighted sum of squares *Q*, and the proportion of excess dispersion I^2. The *Q* statistic reflects the total variance of the effects subsumed by one summary effect size; it is interpreted as a χ^2 with degrees of freedom equal to the number of studies minus one. If *Q* is statistically significant, the studies are more heterogeneous than expected, and the summary effect size should be interpreted with caution. The I^2 statistic returns the percent on dispersion that exceeds the expected value. Conventionally, values above 25% indicate a noteworthy excess dispersion. If a group of studies is heterogeneous (as indicated by *Q* and I^2), the sources of this heterogeneity should be identified through moderator analyses. None of the groups of studies we meta-analysed showed significant heterogeneity; therefore, no moderation analyses were performed. In order to visualise the dispersion of the effect sizes and their confidence intervals, we constructed forest plots with GraphPad Prism version 6.00 for Windows (GraphPad Software, La Jolla California USA).

Fourth, we assessed the potential effect of publication bias on our results. Publication bias refers to the tendency of researchers and journal editors to publish only significant results, a tendency that leads to the overestimation of effects in meta-analyses. Following Begg and Mazumdar's (1994) method, we computed the correlation between effect sizes and their respective standard errors; a significant correlation would indicate a potential publication bias (see Field & Gillet, 2010, pp. 684-690, for an explanation). We also computed Rosenthal's fail-safe number, i.e., the number studies with nonsignificant results that would be necessary to reduce a summary effect size to 0.

A summary effect size was not always computed. Within certain classes of interventions, studies were too diverse to allow for a meaningful summary effect size. In these cases, only the direction of the effect was considered, and a sign test was performed (see Borenstein et al., 2009).

In the rest of this section, a heading is dedicated to each type of intervention. Since the handful of studies on implicit measures and on prejudice towards bisexual people were scattered across different types of interventions, we review them under a separate heading (as *Neglected Issues*). The only study on accountability is also discussed in that section.

Education

Ignorance is probably the most often cited cause of prejudice (Brown, 2009). Therefore, it is not surprising that education is the most frequently used technique for reducing sexual prejudice. Almost all of the studies in our corpus were performed in an educational setting, whether in a course or workshop or in a university laboratory. However, we defined an intervention as *educational* only when the transfer of information and skills was the main means for reducing sexual prejudice.

Thirty-two studies examined the effect of education on sexually prejudiced attitudes. See Figure 2 for a forest-plot of individual effect sizes. The summary effect size was medium, $d = 0.46$, $SE = 0.07$. The heterogeneity of the effect sizes was less than expected for this number of studies, $Q = 25.31$, $p > .05$, $I^2 = 0$. There was no evidence of a publication bias, Begg and Mazumdar's $\tau = .19$, $p > .05$. Rosenthal's fail-safe number was 2094. Education was moderately effective in reducing sexually prejudiced attitudes.

Three studies examined the impact of education on behaviour. Riggs and Fell's (2010) workshop had an average positive impact on psychology students' intended behaviour, $d = 0.55$, $SE = 0.21$; Riggs, Rosenthal, and Smith-Bonahue (2011) obtained a similar result with teacher trainees, $d = 0.61$, $SE = 0.11$. Christensen and Sorensen (1994) achieved a more modest change on students' actual behaviour, $d = 0.36$, $SE = 0.36$.

Five studies tested the effect of education on knowledge about gay people and issues. See Figure 3 for a forest-plot of individual effect sizes. The mean effect size was very large, $d = 1.09$, $SE = 0.13$. The effect sizes of the five studies were not significantly heterogeneous, Q

= 4.19, $p > .05$. The variance among the true effect sizes only accounted for small proportion of the observed variability, $I^2 = 4.54$ %. There was no sign of a publication bias, Begg's and Mazumdar's $\tau = .60$, $p > .05$. Rosenthal's fail-safe number was 184. Furthermore, Boulden (2004) and Scher (2009) both found that educational programmes strongly increased people's self-perception of knowledge, $d = 1.01$, $SE = 0.09$, and $d = 1.21$, $SE = 0.23$, respectively. Unsurprisingly, education was highly effective in increasing knowledge about homosexuality.

Five studies tested the effect of education on emotions. See Figure 4 for a forest-plot of individual effect sizes. The summary effect size was small-to-medium, $d = 0.36$, $SE = 0.05$. There was no evidence for heterogeneity among the effect sizes, since the weighted sum of squares was less than expected, $Q = 3.72$, $p > .05$, $I^2 = 0$. There was no evidence for publication bias, Begg and Mazumdar's $\tau = .20$, $p > .05$. Rosenthal's fail-safe number was 66. Story (1979) examined the effect of a sexuality course on students' comfort with a series of sexual behaviours; different questions yielded different results, *d*s ranging -0.46 to 0.66. Overall, education effectively reduced sexually-prejudiced emotions.

Contact

Intergroup contact is arguably the most researched approach to prejudice reduction. Its results are well-documented (see Pettigrew & Tropp, 2006, for a meta-analysis), and its mechanisms are reasonably well understood (Brown, 2009). Moreover, sexual prejudice may be the prejudice on which intergroup contact has the strongest effect ($r = 0.27$, equivalent of approximately $d = 0.56$; Pettigrew & Tropp, 2006). Heterosexual people's contact with LGB people often occurs through disclosure by friends or family, which is more effective in reducing prejudice than disclosure by new acquaintances (Herek & Capitanio, 1996).

Eight studies examined the effect of contact with lesbians and gay men on sexually prejudiced attitudes. See Figure 5 for a forest-plot of individual effect sizes. The mean effect size of these interventions was medium, $d = 0.56$, $SE = 0.16$. There was no evidence of

heterogeneity among the studies, $Q = 9.41$, $p > .05$. The proportion of true variance was $I^2 = 25.62\%$. There was no evidence for publication bias, Begg and Mazumdar's $\tau = .50$, $p > .05$. Rosenthal's fail-safe number was 115. Intergroup contact was moderately effective in reducing sexually prejudiced attitudes

Three studies investigated the effect of contact with LGB people on emotions. Lance (1987) found that contact greatly reduced students' discomfort with LGB people, $d = 1.07$, $SE = 0.32$. Turner, Crisp, and Lambert (2007) found that imagining an interaction with a gay man has a similarly large effect on straight men's intergroup anxiety, $d = 1.43$, $SE = 0.43$. However, Burke (1995) obtained a much more modest effect by exposing participants to a video of a counter-stereotypical gay man, $d = 0.15$, $SE = 0.19$. Although all three studies found positive effects, they were too few to grant a conclusion, $z = 1.15$, $p = .25$.

Only one study explored the effect of contact on sexually prejudiced cognitions. In their imagined contact study, Turner et al. (2007) achieved a great reduction of straight men's conviction that gay men are all similar (i.e., outgroup homogeneity), $d = 0.84$, $SE = 0.40$.

Contact-plus-Education

Education and intergroup contact were often used together in such a manner that it was impossible to differentiate their effects. The prototype of contact-plus-education interventions is the panel presentation: a group of LGB people are invited to a class or a workshop in order to provide information on sexuality, answer participants' questions, and provide an experience of positive intergroup contact (see Croteau & Kusek, 1992, for an early review).

Twenty-seven studies assessed the effect of contact-plus-education on sexually prejudiced attitudes. See Figure 6 for a forest-plot of individual effect sizes. The mean effect size was medium, $d = 0.41$, $SE = 0.06$. There was no evidence for heterogeneity, $Q = 26.66$, $p > .05$, $I^2 = 2.47$ %. There was also no evidence for a publication bias, Begg and Mazumdar's $\tau = .05$, $p > .05$. Rosenthal's fail-safe number was 1407. Interventions combining contact and

education had a medium effect of sexually prejudiced attitudes.

Three studies examined the impact of contact-plus-education on knowledge. Cramer (1997) found that a workshop on sexuality in which the facilitator disclosed her lesbian identity strongly improved social-work students' understanding of lesbian identity development, $d = 1.09$, $SE = 0.22$. Kelley, Chou, Dibble, and Robertson (2008) found that a workshop that included contact with LGB physicians was moderately effective in dispelling healthcare students' misrepresentations of lesbian and gay health, $d = 0.36$, $SE = 0.12$. Fisher (1996) obtained a similar result through a course for teachers in which contact was provided through videos, $d = 0.40$, $SE = 0.38$. Although all three studies found positive effects, they are too few to grant a conclusion, $z = 1.15$, $p = .25$.

Six studies examined the effect of contact-plus-education on emotions. See Figure 7 for a forest-plot of individual effect sizes. The summary effect size was medium, $d = 0.44$, $SE = 0.08$. There was no evidence for heterogeneity among the effect sizes, since the weighted sum of squares was less than expected, $Q = 4.74$, $p > .05$, $I^2 = 0$. There was no evidence for publication bias, Begg and Mazumdar's $\tau = .33$, $p > .05$. Rosenthal's fail-safe number was 82. Contact-plus-education was effective in reducing sexually prejudiced emotions.

Five studies assessed the effect of contact-plus-education on intended behaviour. See Figure 8 for a forest-plot of individual effect sizes. The summary effect size was small to medium, $d = 0.35$, $SE = 0.09$. There was no evidence for heterogeneity among the effect sizes, since the weighted sum of squares was less than expected, $Q = 2.27$, $p > .05$, $I^2 < 0$. There was no obvious risk of publication bias, Begg and Mazumdar's $\tau = .20$, $p > .05$.[5] Rosenthal's fail-safe number was 21. Two studies that used actual behavioural tasks achieved more modest results. Hugelshoffer (2007) asked students to spend time with allegedly LGB peers; those who had attended a panel presentation were slightly more willing to do so, but the effect size differed by the type of activity proposed, average $d = 0.14$, $SE = 0.02$.

Grutzeck and Gidycz (1997) used a similar behavioural measure, but students who had attended a panel presentation were actually less willing to interact with LGB peers, d = -0.07, SE = 0.19.

Social Norms and Expertise

Prejudice can be reduced if tolerance is set as a norm, either by a reference group or by experts. The norms-or-expertise interventions we review here adopted one of two strategies. Some studies, particularly those drawing on Moscovici's minority-influence paradigm (Moscovici, Lage, & Naffrechoux, 1969), manipulated the source of the norm, i.e., the type of group that advocated tolerance. Other studies manipulated the contents of the norm, i.e., whether sexual prejudice was legitimised or condemned.

Five studies examined the effect of the source of normative influence on sexually-prejudiced attitudes. See Figure 9 for a forest-plot of individual effect sizes.The mean effect size was close to nil, d = - 0.02, SE = 0.01. The effect sizes were not heterogeneous, Q = 0.71, p > .05, I^2 = 0. There was no sign of publication bias[6], Begg and Mazumdar's τ = .20, p > .05. However, effect sizes could not be computed for four relevant studies (three in Alvaro & Crano, 1997; and one in Crano & Alvaro, 1998). These results suggest that norms-or-expertise interventions that rely on the prestige of the source are not effective in reducing sexually prejudiced attitudes.

Four studies tested the effect of manipulating norm contents on sexually-prejudiced behaviour. See Figure 10 for a forest-plot of individual effect sizes. The mean effect size was medium, d = 0.46, SE = 0.13. The effect sizes of the four studies were not significantly heterogeneous, Q = 3.20, p > .05. The variance among the true effect sizes only accounted for a small proportion of the observed variability, I^2 = 6.18 %. There was no sign of a publication bias, Begg and Mazumdar's τ = -.67, p > .05. Rosenthal's fail-safe number was 37. Tolerant social norms had a medium effect on participants' behaviour.

Two studies explored the effect of norms on emotions. Banse, Seise, and Zerbes (2001) found that expert messages did not affect German male students' sexually prejudiced emotions, $d = 0.02$, $SE = 0.32$. Pereira, Monteiro, and Camino (2009) compared Portuguese students' responses in a situation with an anti-discrimination norm and in a control situation; participants exposed to the anti-discrimination norm expressed fewer positive emotions ($d = -0.65$, $SE = 0.21$) and a similar level of negative emotions ($d = 0.07$, $SE = 0.21$) compared to control participants. While little information is available on the issue, norms-or-expertise interventions do not seem to reduce affective sexual prejudice.

Inducing Emotions

Researchers have successfully reduced prejudice by inducing empathy towards a discriminated group or by otherwise manipulating participants' emotions (Paluck & Green, 2009). Certain interventions in our corpus employed empathy-inducing exercises (esp. role playing). Other studies investigated the effect of disgust on sexual prejudice.

Five studies explored the effectiveness of empathy-inducing exercises in reducing sexual prejudice. Both MacLaury (1983) and Israel and Hackett (2004) have obtained some reduction of students' sexually-prejudiced attitudes through such exercises, $d = 0.29$, $SE = 0.23$, and $d = 0.30$, $SE = 0.19$, respectively. Unsurprisingly, the same exercise employed by Israel and Hackett (2004) had a very modest effect on knowledge, $d = 0.05$, $SE = 0.22$. Hillman and Martin (2002) created an exercise named *Alien Nation*, in which students had to imagine living on a planet where all forms of sexuality are forbidden; they obtained a larger reduction of sexually-prejudiced attitudes with this task than with a lecture, $d = 0.17$, $SE = 0.30$. Hodson, Choma, and Costello (2007) also found that *Alien Nation* was more effective than a lecture in reducing negative emotions, $d = 0.45$, $SE = 0.18$. Nevertheless, these results are insufficient to indicate a positive trend, sign test $z = 0.89$, $p = .375$.

The manipulation of disgust was pursued in three studies, with interesting results.

Participants in whom disgust was induced had more prejudiced responses both on the IAT (*d* = - 0.43, *SE* = 0.19; Dasgupta, DeSteno, Williams, & Hunsinger, 2009) and on an emotional thermometer (*d* = - 0.43, *SE* = 0.19; Inbar, Pizarro, & Bloom, 2012). In contrast, disgust was associated with a slight decrease of sexually prejudiced attitudes (*d* = 0.18, *SE* = 0.20; Terrizzi, Shook, & Ventis, 2010).

Entertainment Media

Entertainment media have long been assumed to have an impact on prejudice (see Allport, 1954). Novels, television shows, films, and other forms of entertainment have often been used by activists and policy makers aiming to contain prejudice and counter stereotypes. Empirical evidence on the effectiveness of this approach is mixed, but generally promising (Paluck & Green, 2009).

All 11 studies in this category employed some form of audio-visual entertainment. Books were almost never used, with the notable exception of a qualitative study we excluded from our sample (Smith, 1994). Musical and theatrical performances, such as *The Laramie Project*, were used in a handful of studies, but only to facilitate a broader educational curriculum (see *Education*). The entertainment studies used a wide range of genres, including documentary films, talk-shows, and pornography.

Effect sizes could be computed for 11 studies. The effect sizes ranged from *d* = 0.26 to 0.61, with one study (Duncan, 1988) having an exceptionally large effect of 1.35. The contents of the videos used, as well as the research designs were too heterogeneous to compute a summary effect size. The sample was also too small to explore what differentiates effective and ineffective interventions (i.e., moderators). However, a sign test indicated a tendency for entertainment to have a positive effect, $z = 3.00$, $p = .004$.

Priming Techniques

Priming people on tolerant values has been reported to reduce prejudice both in the

laboratory and in more natural settings (Paluck & Green, 2009). The mechanism behind this effect seems to be people's need to maintain consistency among their attitudes and a sense of positive self-worth (see e.g., Greenwald et al., 2002). Most studies value priming investigated how priming participants on socially conservative values prompted more sexually prejudiced responses.

Five studies tested the effect of value priming on sexually prejudiced attitudes. Lehmiller, Law, and Tormala (2010) performed three studies in which they affirmed participants on the importance of family and on other values. Priming family values induced a small increase in sexually prejudiced attitudes compared to no priming, d = -0.09, SE = 0.20, and d = -0.13, SE = 0.18; and a moderate increase compared to priming participants on humour, d = -0.54, SE = 0.18, and d = -0.53, SE = 0.24. Humour also proved moderately effective in reducing sexual prejudice compared to no priming, d = 0.44, SE = 0.18. Webster and Saucier (2011) performed two studies to test whether thinking about one's mortality increases sexually prejudiced attitudes. The overall effect was close to nil, d < 0.01, SE = 0.06; but there was a complex pattern of interactions. Bonds-Raacke, Cady, Schlegel, Harris, and Firebaugh (2007) found that instructing participants to remember positive gay characters on television moderately improved attitudes towards gay men, d = 0.44, SE = 0.19. (An effect size for attitudes towards lesbians could not be computed.)

Three studies investigated the impact of priming values on affective sexual prejudice. Two studies by Webster and Saucier (2011) found a complex pattern of gender differences, but the overall effect of mortality salience was close to nil, d < 0.01, SE = 0.11. Johnson (2011) found that a lexical priming task with religious content leads to more affective sexual prejudice than the same task with neutral content, d = -0.51, SE = 0.23.

Awareness and Suppression

Becoming aware of one's prejudice and attempting to consciously control it has been

a controversial topic in the history of social psychology. While Allport (1954) was optimistic about this strategy, subsequent experiments have shown paradoxical effects. Attempts to suppress prejudiced thoughts have been shown to induce more prejudiced thoughts and behaviour in some contexts (i.e., a *rebound effect*; Macrae, Bodenhause, Milne, & Jetten, 1994; see Wenzlaff & Wegner, 2000, for a review).

Five studies have examined the effect of awareness and suppression on sexual prejudice. Kennedy (1995) used a self-confrontation technique with a large number of American students. This technique achieved a medium reduction of participants' scores on the ATLG, $d = 0.43$, $SE = 0.03$. Monteith, Spicer, and Tooman (1998) performed two studies examining the rebound effect described above. In both studies, they achieved an average reduction of the number of prejudicial statements by simply instructing participants to avoid them, $d = 0.49$ and 0.50, $SE = 0.05$ and 0.04, respectively. Moreover, neither of the studies found a rebound effect. Banse et al. (2001) similarly found that the conscious suppression of sexual prejudice was very effective for both attitudes, $d = 0.77$, $SE = 0.33$; and emotions, $d = 1.35$, $SE = 0.33$. In an interesting experiment, Gailliot, Peruche, Pant, and Baumeister (2008) offered participants sucrose drinks before writing an essay about a gay character. Although participants did not receive any instructions to suppress prejudice, those who drank the sucrose drink used fewer stereotypes, $d = 0.64$, $SE = 0.08$. The authors interpreted these findings as indicative of the role of the brain's glucose supply in consciously controlling behaviour.

All five studies relying on awareness and suppression achieved an average reduction of sexually prejudiced responses. As these studies were different in their methods and scope, we decided not to compute a summary effect size. A sign test indicated that the probability of five out of five studies having positive results is fairly low; it does not, however, achieve conventional statistical significance, $z = 1.86$, $p = .063$.

Neglected Issues: Accountability Interventions, Implicit Measures, and Prejudice against Bisexual People

Certain interventions and outcomes have received very little attention from researchers. Both prejudice against bisexual people and implicit prejudice have been neglected by the psychological literature at large (see the *General Discussion*). Accountability interventions (addressed by only one study), cooperative learning, and peer debates (not addressed in any report) are also underexplored in general (Paluck & Green, 2009); social categorisation was explored in certain studies as a mechanism of change (e.g., Turner et al., 2007), but no study was dedicated to the effect of manipulating categories on sexual prejudice.

Only one study explored the effect of accountability on sexual prejudice. Pereira et al. (2009) told Portuguese students they would later have to explain their responses to a set of questionnaires[7]. Participants in this condition expressed less sexually prejudiced attitudes than those in a control group, d = 0.53, SE = 0.21. They also expressed the same level of positive emotion, d = - 0.08, SE = 0.21; and less negative emotion, d = 0.38, SE = 0.21.

Four studies have explored the impact of psychological interventions on implicit sexual prejudice. Read (1978) used GSRs to assess the effect of anti-prejudice education. Participants who had listened to a lecture on sexuality had a much lower skin response when an openly gay experimenter touched them (with the pretext of attaching electrodes), d = 0.80, SE = 0.28. Banse et al. (2001) employed the IAT to compare the effect of the intentional suppression of prejudice and that of a pro-gay message by experts; the data was not reported in sufficient detail due to the lack of any significant differences. Dasgupta and Rivera (2008) found that contact with gay people through biographical vignettes had a medium positive impact on sexual prejudice as measured with the IAT, d = 0.30, SE = 0.18. Dasgupta et al. (2009) also found that sexual prejudice IAT scores were increased when disgust was activated, as opposed to anger or no emotion, d = - 0.43, SE = 0.19.

Morin (1974), in what was seemingly the first attempt to reduce sexual prejudice through a psychological intervention, also addressed attitudes towards bisexual people. He combined contact and education to induce a large reduction in social distance to bisexual people, $d = 0.62$, $SE = 0.07$. Hugelshoffer (2006) also performed a contact-plus-education intervention and achieved a small reduction in anti-bisexual prejudice, $d = 0.16$, $SE = 0.10$. Dessel (2010) used a similar approach and achieved a medium effect, $d = 0.42$, $SE = 0.33$. Finally, Bronson (2006) employed empathetic stories to induce tolerance towards bisexual people, but observed the opposite effect, $d = -0.18$, $SE = 0.16$.

General Discussion

The present review examined patterns in the methodology, participant characteristics, and theoretical approaches of interventions to reduce sexual prejudice. Education, contact, contact-plus-education, and norms-or-expertise interventions effectively reduced participants' scores on at least some measures of sexual prejudice. Entertainment with anti-prejudice content produced promising results, but the studies were too diverse to support an overall conclusion. The outcomes of the interventions were typically assessed by the use of self-report sexual-prejudice scales, sometimes accompanied by emotional, cognitive, or behavioural measures, and the use of implicit measures was rare. Participants in these studies were typically young, American women enrolled in education. However, the reports often failed to offer detailed information on participants' characteristics, including participants' sexuality. Most approaches to prejudice reduction were explored in the case of sexual prejudice, but no study in our corpus carried out social-categorisation experiments, cognitive training, or peer debate. Prejudice towards bisexual people was largely neglected. Finally, unpublished postgraduate research showed a number of advantages over published research. Below, we discuss these findings in more detail, looking at both the conclusions we can draw and the issue that are yet to be researched.

The Effectiveness of the Interventions

The meta-analytic review showed that the effectiveness of at least four types of interventions to reduce sexual prejudice is supported by the literature. Educational interventions are highly effective in increasing knowledge about LGB people; their effectiveness in improving attitudes and emotions is more modest, but solid. Contact with LGB people has a moderate positive effect on attitudes. Interventions that integrate contact and education are moderately effective in improving attitudes, emotions, and behavioural intentions in relation to gay people. Finally, inducing tolerant social norms can moderately improve behaviour, but not attitudes. See Table 3 for details.

Overall, the change induced by these four interventions was of about one third to one half of a standard deviation in size, and there was little variation across interventions and outcomes. This is not to say that the characteristics of the interventions do not matter. The effect sizes of individual studies ranged from nil to very large, and it is therefore intuitively likely that there are meaningful differences among the studies. However, the effect sizes were too homogeneous to reveal particular moderators of effectiveness. Future research on more diverse samples may reveal important differences between cultures and between age groups.

Two outcomes, however, do not fit the overall pattern of medium effects. First, educational interventions had an particularly large impact on knowledge about (homo)sexuality. Second, norms-or-expertise interventions had a medium effect on behaviour but had no effect on attitudes. The second of these patterns is difficult to interpret, since researchers who manipulated the source of the message (e.g., a minority versus a majority organisation) typically used attitudinal measures, while researchers who manipulated the contents of the message (i.e., whether tolerance or prejudice was promoted) employed behavioural measures. Future research should explore if behaviour is more susceptible to normative influences, or the contents of a norm is more relevant than its source.

Sampling and Design Issues

This review revealed that scientific knowledge about reducing sexual prejudice has drawn on a very narrow research base. In psychological research, studies are generally conducted with young North American students (Arnett, 2009). Eighty-nine percent of the studies in our review employed North American samples. Arnett (2009) found no APA journal with more than 81% American content between 2003 and 2007; although we searched for studies from all over the world, 83.3% of the samples in our corpus were drawn from the US for the same period (n = 18). The oversampling of American participants is problematic because psychological studies often have substantially different results when conducted with American or non-American populations (Henrich, Heine, & Norenzayan, 2010).

Focusing on young, educated Americans is especially problematic in prejudice research. The US is among the less sexually prejudiced nations. According to the World Value Survey (Inglehart, 2008), 31.3 % of Americans stated that "homosexuality is never justifiable," as opposed to 90% of Georgians and 99.2% of Jordanians (while only 4.1% of Swedes agreed with this statement). Predominantly researching young people is also problematic: North American youths tend to be more accepting of homosexuality than their elder (see e.g., Andersen & Fetner, 2008). The oversampling of women (approximately 76% of the participants) and the failure to report the sample's gender composition (in 16% of the studies) further troubles the generalisation of findings from these studies. Men have been shown to be more sexually prejudiced than women in multiple studies, and this difference is especially large among college students (Kite & Whitley, 1996). Students who volunteer to participate in sexuality related research also have more sexual experiences and less restrictive values than their peers (Wiederman, 1999). In conclusion, the extant literature has studied sexual prejudice on a population that is comparatively unlikely to hold such prejudice. Consequently, research has addressed intervention strategies that may not be easily

transferable to other populations where such interventions are needed the most. We strongly urge the diversification of this field of research in order to guide prejudice-reduction efforts in other populations.

Promising approaches to prejudice reduction were also left unexplored by the studies in our corpus. Social categorisation, cognitive training, and peer debate have had promising results in reducing prejudice based on race and ethnicity (Paluck & Greene, 2009). However, no study seems to have explored the utility of any of these approaches in combating sexual prejudice. Intergroup contact has a particularly large effect on sexual prejudice (Pettigrew & Tropp, 2006), and there is a possibility that other approaches to prejudice reduction would also be very effective. Future studies will need to investigate whether this is the case.

Like other reviews (e.g., Kite & Whitley, 1996), we found that researchers in this area did not always record their participants' sexualities. While LGB people may foster negative thoughts and feelings about their sexuality (Szymanski, Kashubeck-West, & Meyer, 2008), they are still, on average, vastly more positive about homosexuality than their heterosexual peers (see e.g., Herek, Gillis, & Cogan, 2009). Researchers often rely on the assumption that LGB people are few in number, and therefore unlikely to participate in their studies or to affect their statistical conclusions (see Bonds-Raacke et al., 2007). However, there are both theoretical and empirical reasons to insist on accounting for participants' sexuality in such research. Lesbians, gay men, and bisexual people can be surprisingly common among volunteers for sexuality research (e.g., 16% in Hylton's, 2006, sample). Assuming by default that people are heterosexual is central to heterosexism and sexual stigma (Herek, Kimmel, Amaro, Melton, 1991; Warner, 1993; Herek, 2007): ironically, this assumption is frequently made in conducting the very studies that aim to reduce sexual prejudice.

Sexual prejudice has most often been operationalised in terms of specific sexual prejudice scales, while implicit measures such as the IAT have been used very rarely. The use

of standardised scales has obvious advantages, but it can have unintended effects on the way sexual prejudice is understood. Self-report scales rely on the assumption that people can and will express their prejudice, which is not always the case (Steffens, 2005). While sexual prejudice scales largely overlap (Rye & Meaney, 2010), they tend to obscure specific aspects of prejudice such as fear of outgroups (Stephan & Stephan, 1985) or the rejection of bisexual people (see our subsection on *Neglected Issues*). The neglect of anti-bisexual prejudice is particularly concerning, since bisexual people are subject to more stress than their lesbian and gay peers (Meyer, 2006).

The Value of Unpublished Studies

One of the most surprising findings from our study-space analysis is that methodologically strong studies often go unpublished. There is approximately one dissertation for every seven journal articles archived on PsycINFO in general. However, in our corpus, there is one dissertation for every three articles. These unpublished reports often present significant differences, and a "publication bias" in favour of positive results (Rosenthal, 1979; Ferguson & Brannick, 2012) does not seem to explain this pattern. A general bias against postgraduate research is a second possibility. Of course, postgraduate students do not all publish their work, but this explanation cannot account for the prevalence of the methodologically *stronger* studies in the unpublished literature. A third possibility is that scholars in this particular field are affected by *courtesy stigma*; psychologists doing research on sexuality often face "stigma by association," and they may be automatically labelled as LGB themselves (Minton, 2002; Coyle, 2004). Younger researchers may be particularly affected by courtesy stigma, such that good quality dissertations on sexual prejudice are not developed for publication. This explanation is consistent with observations that postgraduate researchers in LGB psychology are concerned about the effect of courtesy stigma on their future career (Biaggio, Orchard, Larson, Petrino, & Mihara, 2003). Such

young researchers find experiences of mentorship in LGB psychology to be surprising and transformative, but such mentorship may be difficult to access (Curtin, Hegarty, & Stewart, 2012). Whatever the explanation, much good research on sexual prejudice seems to remain unpublished, and this is particularly concerning in a field that remains small (Lee & Crawford, 2007; 2012) and which carries ethical obligations from psychology's past.

Finally, a larger proportion of non-U.S. studies than U.S.-based studies were funded. Studies performed outside the US were comparatively rare (about 10% of our corpus), and this finding may be a statistical artefact. Alternatively, funding bodies in other countries may be more willing to fund research on reducing sexual prejudice than their U.S. counterparts. Conversely, there may be so little support for this topic outside the US that research is hardly ever completed or published, apart from the handful of projects that manage to secure funding.

Limitations and Future Research

No review can be complete, but we took several precautions to assure that we included as many of the relevant studies as possible. We sampled dissertations, performed Google searches, and translated our keywords to several widely-spoken languages. However, several interesting interventions might have escaped our attention. Most countries lack-grey literature databases, and none of the dissertations we retrieved were from outside the US. Yet numerous interventions are performed without research in mind, and therefore no data are collected in these contexts. For example, several large-scale campaigns against sexual prejudice took place in South America in the early 2000s, but none of them yielded data on their psychological impact (Pan American Health Organization, 2008). Those who perform such interventions in the future should seek to rejoin practice and research, especially outside the US.

The study-space analysis pointed out several directions for future research. The

outcomes of these interventions were overwhelmingly assessed with attitudes scales. There was comparatively little information available on cognitions, emotions, and behaviours, and almost none on implicit prejudice. Most importantly, almost 90% of these studies were performed in the US. The question as to whether these interventions are similarly effective in other cultures remains open.

In addition to data collection, more research integration is also necessary. For the sake of coherence, we limited our review to exclude follow-up studies and qualitative research, but our searches suggested that both bodies of work could be reviewed in the future. Sexual prejudice itself has many intertwined aspects that are beyond the scope of our review, although we recognise their importance. Specifically, our knowledge of how to reduce sexual prejudice would be more complete if we better understood how to reduce LGB people's prejudice towards themselves (i.e., internalised heterosexism; Szymanski, Kashubek-West, & Meyer, 2008), as well as the strategies they use to cope with prejudice and discrimination (Moradi, Mohr, Worthington, & Fassinger, 2009). It is equally important to understand prejudice directed towards heterosexual people who combat sexual prejudice; as we have seen above, courtesy stigma may actually be hindering research in this field. We hope that well-synthesised research on all these issues will emerge in the near future.

Lastly but importantly, our review remained silent on the theoretical underpinnings of interventions to reduce sexual prejudice. As this review reveals, the development of practical anti-prejudice strategies has often had a loose relationship with theory and research. Educational interventions, for example, are often informed by our society's view of prejudice as rooted in ignorance rather than a more sophisticated theory of how prejudice works (Bartoş & Hegarty, in press). However, it is not uncommon for intervention studies to proceed with theoretical research following years later. Allport's (1954) contact hypothesis gained prominence during the 1960s struggle for African-American civil rights (Brown, 2008), three

decades before its underlying mechanisms were clarified by Gaertner, Mann, Dovidio, Murrell, and Pomare (1990). Nevertheless, understanding the psychological mechanisms behind each of these strategies is of both scientific and practical importance, since increasing the effectiveness of a practical technique requires theoretical understanding (Michie, 2008). Intervention mapping provides tools for synthesising research and integrating it with theory, with excellent results in health psychology interventions (Bartholomew, Parcel, & Kok, 1998). Therefore, we feel that intervention mapping performed on different practical strategies could bring major advances in our understanding of reducing sexual prejudice; we intended our review as a first step toward such deeper inquiries.

Conclusion

The first 40 years of psychological research on reducing sexual prejudice has produced reliable knowledge, but it has also neglected several promising approaches and many relevant demographic groups. Future research should explore cultural and age differences systematically, in order to design anti-sexual prejudice interventions for populations that are more in need of them than are typical American college students. Filling in the gaps of this literature is obviously intertwined with issues of funding and dissemination. Limited resources are the typical reason for performing research on convenience samples (Dasgupta & Hunsinger, 2008). The neglect of certain approaches and certain outcome measures may have similar underpinnings: cognitive training and implicit prejudice are comparatively resource-intensive to research. Moreover, postgraduate researchers seem to face particular difficulties in completing and publishing their work on this topic. We therefore conjecture that the current weaknesses in our knowledge about sexual prejudice may be due to a lack of systematic support for research in this area, which may be partially due to sexual prejudice itself.

While we agree with other reviewers that the literature on reducing sexual prejudice has

serious limitations, we have reason to see this field in a brighter light. While Tucker and Potocki-Tripodi (2006) found a handful of studies, many of which had questionable designs, we managed to identify over one hundred and fifty studies, almost half of which were randomised experiments. Most of these studies were successful, to some extent, in reducing sexual prejudice, and meta-analyses show that effect sizes were typically in the medium range. Much research was conducted by postgraduate students, often without the recognition that comes with publication. While the limitations discussed above commend caution and future investigations, the literature we have reviewed also evidences psychology's ethical commitment to understand and reduce sexual prejudice.

References

References marked with an asterisk indicate studies included in the study space-analysis and meta-analysis.

Adam, B. D. (1998). Theorizing homophobia. *Sexualities, 1*, 387-404.

Allport, G. W. (1954). *The nature of prejudice*. Oxford, England: Addison-Wesley.

*Alvaro, E. M., & Crano, W. D. (1997). Indirect minority influence: Evidence for leniency in source evaluation and counterargumentation. *Journal of Personality and Social Psychology, 72*, 949-964.

American Psychological Association. (2008). Ethical principles of psychologists and code of conduct. In D. N. Bersoff, & D. N. Bersoff (Eds.), *Ethical conflicts in psychology* (pp. 14-30). Washington, DC: Author.

American Psychological Association. (2012). Guidelines for psychological practice with lesbian, gay, and bisexual clients. *American Psychologist, 67*, 10-42.

Andersen, R., & Fetner, T. (2008). Cohort differences in tolerance of homosexuality: Attitudinal change in Canada and the United States, 1981-2000. *Public Opinion Quarterly, 72*, 311-330.

Ansara, Y. G. (2010). Beyond cisgenderism: Counselling people with non-assigned gender identities. In L. Moon (Ed.), *Counselling ideologies: Queer challenges to heteronormativity* (pp. 167-200). Aldershot, England: Ashgate.

Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist, 63*, 602-614.

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

Bartholomew, L.K., Parcel, G.S., & Kok, G. (1998). Intervention mapping: A process for developing theory- and evidence-based health education programs. *Health Education & Behavior, 25*, 545-563.

Bartoș, S.E., & Hegarty, P. (in press). Gender, race, and ethnic relations. In C. Kinvall, T. Capelos, H. Dekker, & P. Nesbitt-Larking (Eds.), *Political psychology in Europe and the World*. London, England: Palgrave.

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50*, 1088-1101.

Biaggio, M., Orchard, S., Larson, J., Petrino, K., & Mihara, R. (2003). Guidelines for gay/lesbian/bisexual-affirmative educational practices in graduate psychology programs. *Professional Psychology: Research and Practice, 34*, 548-554.

Bohner, G., & Dickel, N. (2011). Attitudes and attitude change. *Annual Review of Psychology, 62*, 391-417.

*Bonds-Raacke, J., Cady, E. T., Schlegel, R., Harris, R. J., & Firebaugh, L. (2007). Remembering gay/lesbian media characters: Can Ellen and Will improve attitudes toward homosexuals? *Journal of Homosexuality, 53*, 19-34.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to*

meta-analysis. Chichester, England: John Wiley & Sons.

*Boulden, W. T. (2004). Evaluation of the Anytown leadership institute programming on lesbian, gay, bisexual and transgender issues. *Journal of Gay & Lesbian Social Services: Issues in Practice, Policy & Research, 17,* 17-38.

*Bronson, J. C. (2006). Examining the efficacy of fact-based and personal-story-based information on attitudes towards bisexuality. *Dissertation Abstracts International: Section B. Sciences and Engineering, 66,* 4532.

Brown, R. (2010). *Prejudice: Its social psychology.* Chichester, England: Wiley-Blackwell.

*Burke, T. B. (1997). The effect of gay male sex role adherence on others' homosexual stress and attitudes toward gay men. *Dissertation Abstracts International: Section B. Sciences and Engineering, 57,* 6635.

*Chonody, J. M., Siebert, D. C., & Rutledge, S. E. (2009). College students' attitudes toward gays and lesbians. *Journal of Social Work Education, 45,* 499-512.

*Christensen, S., & Sorenson, L. M. (1994). Effects of a multi-factor education program on the attitude of child and youth worker students toward gays and lesbians. *Child & Youth Care Forum, 23,* 119-133.

Clarke, V., Ellis, S. J., Peel, E., & Riggs, D. W. (2010). *Lesbian, gay, bisexual, trans and queer psychology: An introduction*. New York, NY: Cambridge University Press.

Cochran, S. D. (2001). Emerging issues in research on lesbians' and gay men's mental health: Does sexual orientation really matter? *American Psychologist, 56,* 931-947.

Cochran, S. D., Sullivan, J. G., & Mays, V. M. (2003). Prevalence of mental disorders, psychological distress, and mental services use among lesbian, gay, and bisexual adults in the United States. *Journal of Consulting and Clinical Psychology, 71,* 53-61.

Conger, J. J. (1975). Proceedings of the American Psychological Association, incorporated, for the year 1974: Minutes of the annual meeting of the council of

representatives. *American Psychologist, 30,* 620-651.

Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field studies.* Boston, MA: Houghton Mifflin.

Coyle, A. (2004). Heterosexual people working in lesbian and gay psychology. *Lesbian & Gay Psychology Review, 5*, 54.

Coyle, A. & Kitzinger, C. (Eds.). (2002). *Lesbian and gay psychology: New perspectives*. Malden, England: Blackwell.

*Cramer, E. P. (1997). Effects of an educational unit about lesbian identity development and disclosure in a social work methods course. *Journal of Social Work Education, 33,* 461-472.

Crandall, C. S., Eshleman, A., & O'Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalization. *Journal of Personality and Social Psychology, 82,* 359-378.

*Crano, W. D., & Alvaro, E. M. (1998). Indirect minority influence: The leniency contract revisited. *Group Processes & Intergroup Relations, 1,* 99-115.

Croteau, J. M., & Kusek, M. T. (1992). Gay and lesbian speaker panels: Implementation and research. *Journal of Counseling & Development, 70,* 396-401.

Curtin, N., Hegarty, P., & Stewart, A.J. (2012). Expanding the research community in LGBT psychology: Collaborative studies from the international institute. *Psychology and Sexuality, 3*, 187-194.

*Dasgupta, N., DeSteno, D., Williams, L. A., & Hunsinger, M. (2009). Fanning the flames of prejudice: The influence of specific incidental emotions on implicit prejudice. *Emotion, 9*, 585-591.

Dasgupta, N., & Hunsinger, M. (2008). The opposite of a great truth is also true: When do student samples help versus hurt the scientific study of prejudice? *Psychological*

Inquiry, 19, 90-98.

*Dasgupta, N., & Rivera, L. M. (2008). When social context matters: The influence of long-term contact and short-term exposure to admired outgroup members on implicit attitudes and behavioral intentions. *Social Cognition, 26*, 112-123.

*Dessel, A. B. (2010). Effects of intergroup dialogue: Public school teachers and sexual orientation prejudice. *Small Group Research, 41*, 556-592.

Dover, K.J. (2002). Classical Greek attitudes to sexual behaviour. In L.K. McClure (Ed.), *Sexuality and gender in the classical world: Readings and sources* (pp. 19-33). Oxford, England: Blackwell.

*Duncan, D. F. (1988). Effects on sexual prejudice of viewing a gay-themed film. *Psychological Reports, 63*, 46.

Ehrlich, H. J. (1973). *The social psychology of prejudice: A systematic theoretical review and propositional inventory of the American social psychological study of prejudice*. Oxford, England: John Wiley & Sons.

Epprecht, M. (1998). Good God Almighty. In S.O. Murray & W. Roscoe (Eds.), *Boy-wives and female husbands: Studies in African homosexualities* (pp. 197-222). New York, NY: Palgrave.

Esses, V.M., Smenya, A.H., & Stelzl, M. (2004). Prejudice and discrimination. In C. Spielberger (Ed.), *Encyclopedia of applied psychology* (Vol. 3, pp. 101-106). Amsterdam, Netherlands: Elsevier.

Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods*, *17*, 120-128.

Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology, 63*, 665-694.

*Fisher, J. B. (1996). The effect of an educational program on teacher and school counselor knowledge, attitudes, and beliefs regarding homosexuality and gay youth. *Dissertation Abstracts International Section A. Humanities and Social Sciences, 57*, 1569.

Gaertner, S. L., Mann, J. A., Dovidio, J. F., Murrell, A. J., & Pomare, M. (1990). How does cooperation reduce intergroup bias? *Journal of Personality and Social Psychology, 59*, 692-704.

*Gailliot, M. T., Peruche, B. M., Plant, E. A., & Baumeister, R. F. (2009). Stereotypes and prejudice in the blood: Sucrose drinks reduce prejudice and stereotyping. *Journal of Experimental Social Psychology, 45*, 288-290.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3-25.

*Grutzeck, S., & Gidyck, C. A. (1997). The effects of a gay and lesbian speaker panel on college students' attitudes and behaviors: The importance of context effects. *Imagination, Cognition and Personality, 17*, 65-81.

Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology, 61*, 29-48.

Hansen, A. L. (2007). School-based support for GLBT students: A review of three levels of research. *Psychology in the Schools, 44*, 839-848.

Hatzenbuehler, M. L., McLaughlin, K. A., Keyes, K. M., & Hasin, D. S. (2010). The impact of institutional discrimination on psychiatric disorders in lesbian, gay, and bisexual populations: A prospective study. *American Journal of Public Health, 100*, 452-459.

Hatzenbuehler, M. L., Nolen-Hoeksema, S., & Erickson, S. J. (2008). Minority stress predictors of HIV risk behavior, substance use, and depressive symptoms: Results from a prospective study of bereaved gay men. *Health Psychology, 27*, 455-462.

Hegarty, P. (2006). Where's the sex in sexual prejudice? *Lesbian & Gay Psychology Review, 7*, 264-275.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*, 61-83.

Herek, G. M. (1984). Attitudes toward lesbians and gay men: A factor analytic study. *Journal of Homosexuality, 10*, 39-51.

Herek, G. M. (2004). Beyond 'sexual prejudice': Thinking about sexual prejudice and stigma in the twenty-first century. *Sexuality Research & Social Policy: A Journal of the NSRC, 1*, 6-24.

Herek, G. M. (2007). Confronting sexual stigma and prejudice: Theory and practice. *Journal of Social Issues, 63*, 905-925.

Herek, G. M., & Capitanio, J. P. (1996). 'Some of my best friends': Intergroup contact, concealable stigma, and heterosexuals' attitudes toward gay men and lesbians. *Personality and Social Psychology Bulletin, 22*, 412-424.

Herek, G. M., Gillis, J. R., & Cogan, J. C. (2009). Internalized stigma among sexual minority adults: Insights from a social psychological perspective. *Journal of Counseling Psychology, 56*, 32-43.

Herek, G. M., Kimmel, D. C., Amaro, H., & Melton, G. B. (1991). Avoiding heterosexist bias in psychological research. *American Psychologist, 46*, 957-963.

Higgins, J. P., & Green, S. (Eds.). (2008). *Cochrane handbook for systematic reviews of interventions*. New York, NY: Wiley-Blackwell.

*Hillman, J., & Renee A. (2002). Lessons about gay and lesbian lives: A spaceship exercise. *Teaching of Psychology, 29*, 308-311.

*Hodson, G., Choma, B. L., & Costello, K. (2009). Experiencing Alien-Nation: Effects of a simulation intervention on attitudes toward homosexuals. *Journal of Experimental Social*

Psychology, 45, 974-978.

Howell, D. C. (1992). *Statistical methods for psychology* (3rd ed.). Boston, MA: PWS-Kent.

*Hugelshofer, D. S. (2007). The effectiveness of lesbian, gay, and bisexual speaker panels in facilitating attitude and behavior change among heterosexual university students. *Dissertation Abstracts International: Section B. Sciences and Engineering, 68*, 623.

*Hylton, M. E. (2006). Online versus classroom-based instruction: A comparative study of learning outcomes in a diversity course. *The Journal of Baccalaureate Social Work, 11*, 102-114.

*Inbar, Y., Pizarro, D. A., & Bloom, P. (2012). Disgusting smells cause decreased liking of gay men. *Emotion, 12*, 23-27.

Inglehart, R. F. (2008). Changing values among western publics from 1970 to 2006. *West European Politics, 31*, 130-146.

*Israel, T., & Hackett, G. (2004). Counselor education on lesbian, gay, and bisexual issues: Comparing information and attitude exploration. *Counselor Education and Supervision, 43*, 179-191.

*Kelley, L., Chou, C. L., Dibble, S. L., & Robertson, P. A. (2008). A critical intervention in lesbian, gay, bisexual, and transgender health: Knowledge and attitude outcomes among second-year medical students. *Teaching and Learning in Medicine, 20*, 248-253.

Kendall. (1998). When a woman loves a woman. In S.O. Murray & W. Roscoe (Eds.), *Boy-wives and female husbands: Studies in African homosexualities* (pp. 223-242). New York, NY: Palgrave.

*Kennedy, S. A. (1996). Changing attitudes toward lesbian women and gay men through self-confrontation. *Dissertation Abstracts International: Section B. Sciences and Engineering, 57*, 2219.

King, M., Semlyen, J., Tai, S. S., Killaspy, H., Osborn, D., Popelyuk, D., & Nazareth, I.

(2008). A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. *BMC Psychiatry, 8*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18706118

Kite, M. E., & Whitley, B. E. J. (1996). Sex differences in attitudes toward homosexual persons, behaviors, and civil rights: A meta-analysis. *Personality and Social Psychology Bulletin, 22*, 336-353.

Kitzinger, C. (1987). *The social construction of lesbianism*. Thousand Oaks, CA: Sage.

Kitzinger, C., & Peel, E. (2005). The de-gaying and re-gaying of AIDS: Contested sexual prejudices in lesbian and gay awareness training. *Discourse & Society, 16,* 173-197.

*Lance, L. M. (1987). The effects of interaction with gay persons on attitudes toward homosexuality. *Human Relations, 40,* 329-336.

Lee, I., & Crawford, M. (2007). Lesbians and bisexual women in the eyes of scientific psychology. *Feminism & Psychology, 17,* 109-127.

Lee, I., & Crawford, M. (2012). Lesbians in empirical psychological research: A new perspective for the twenty-first century? *Journal of Lesbian Studies, 16*, 4-16.

*Lehmiller, J. J., Law, A. T., & Tormala, T. T. (2010). The effect of self-affirmation on sexual prejudice. *Journal of Experimental Social Psychology, 46,* 276-285.

Lim, H. S., & Johnson, M. M. (2001). Korean social work students' attitudes towards homosexuals. *Journal of Social Work Education, 37*, 545-554.

*MacLaury, S. D. (1984). A comparison of three methods of teaching about human sexuality to determine their effectiveness in positively-modifying attitudes about homosexuality. *Dissertation Abstracts International: Section B. Sciences and Engineering, 45,* 755.

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology, 67,* 808-817.

Maher, M. J., Landini, K., Emano, D. M., Knight, A. M., Lantz, G. D., Parrie, M., … Sever, L.M. (2009). Hirschfeld to Hooker to Herek to high schools: A study of the history and development of GLBT empirical research, institutional policies, and the relationship between the two. *Journal of Homosexuality, 56,* 921-958.

Malpass, R. S., Tredoux, C. G., Compo, N. S., McQuiston-Surrett, D., Maclin, O. H., Zimmerman, L. A., & Topp, L.D. (2008). Study space analysis for policy development. *Applied Cognitive Psychology, 22,* 789-801.

Memon, A., Meissner, C. A., & Fraser, J. (2010). The cognitive interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law, 16,* 340-372.

Meyer, I. H. (2007). Prejudice and discrimination as social stressors. In I. H. Meyer, & M. E. Northridge (Eds.), *The health of sexual minorities: Public health perspectives on lesbian, gay, bisexual and transgender populations* (pp. 242-267). New York, NY: Springer.

Michie, S. (2008). What works and how? Designing more effective interventions needs answers to both questions. *Addiction,103,* 886-887.

Minton, H. L. (2002). *Departing from deviance: A history of homosexual rights and emancipatory science in America.* Chicago, IL: University of Chicago Press.

*Monteith, M. J., Spicer, C. V., & Tooman, G. D. (1998). Consequences of stereotype suppression: Stereotypes on and not on the rebound. *Journal of Experimental Social Psychology, 34,* 355-377.

Moradi, B., Mohr, J.J., Worthington, R.L., & Fassinger, R.E. (2009). Counseling psychology research on sexual (orientation) minority issues: Conceptual and methodological challenges and opportunities. *Journal of Counseling Psychology, 56,* 5-22.

*Morin, S. F. (1974). Educational programs as a means of changing attitudes toward gay people. *Homosexual Counseling Journal, 1,* 160-165.

Morin, S. F. (1991). Removing the stigma: Lesbian and gay affirmative counseling. *The Counseling Psychologist, 19*, 245-247.

Moscovici, S., Lage, E., & Naffrechoux, M. (1969). Influence of a consistent minority on the responses of a majority in a color perception task. *Sociometry, 32*, 365-380.

Mulrow, C. D. (1994). Rationale for systematic reviews. *BMJ (Clinical Research Ed.), 309*, 597-599.

Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual Review of Psychology, 60*, 339-367.

Pan American Health Organisation. (2008). *Campaigns against sexual prejudice in Argentina, Brazil, Colombia and Mexico.* Washington, DC: Author.

*Pereira, A., Benedicta Monteiro, M., & Camino, L. (2009). Social norms and prejudice against homosexuals. *The Spanish Journal of Psychology, 12*, 576-584.

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology, 90*, 751-783.

Pettit, M. (2011). The SPSSI task force on sexual orientation, the nature of sex, and the contours of activist science. *Journal of Social Issues, 67*, 92-105.

Plummer, K. (1981). *The making of the modern homosexual*. Totowa, NJ: Barnes & Noble.

*Read, D. V. (1980). A touching case of attitude change toward the homosexual experimenter. *Dissertation Abstracts International: Section B. Sciences and Engineering, 40*, 5081.

*Riggs, A. D., Rosenthal, A. R., & Smith-Bonahue, T. (2011). The impact of a combined cognitive–affective intervention on pre-service teachers' attitudes, knowledge, and anticipated professional behaviors regarding homosexuality and gay and lesbian issues. *Teaching and Teacher Education, 27*, 201-209.

*Riggs, D. W., & Fell, G. R. (2010). Teaching cultural competency for working with lesbian, gay, bisexual and trans clients. *Psychology Learning & Teaching, 9*, 30-38.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86,* 638-641.

Rye, B. J., & Meaney, G. J. (2010). Measuring homonegativity: A psychometric analysis. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement, 42,* 158-167.

*Scher, L. J. (2009). Beyond acceptance: An evaluation of the safe zone project in a clinical psychology doctoral program. *Dissertation Abstracts International: Section B. Sciences and Engineering, 69,* 6453.

Schmidt, F. L., & Hunter, J. E. (2003). Meta-analysis. In I.B. Weiner (Series Ed.), *Handbook of psychology: Vol. 2. Research methods in psychology* (pp. 533-554). Hoboken, NJ: John Wiley & Sons.

Sell, R. L. (1997). Defining and measuring sexual orientation: A review. *Archives of Sexual Behavior, 26,* 643-658.

Smith, D. R. (1994). *The responses of gay and straight readers to YA novels about homosexuality.* Retrieved from ERIC database. (ED378576)

Smith, K. T. (1971). Sexual prejudice: A tentative personality profile. *Psychological Reports, 29,* 1091-1094.

Smith, S. J., Axelton, A. M., & Saucier, D. A. (2009). The effects of contact on sexual prejudice: A meta-analysis. *Sex Roles, 61,* 178-191.

Steffens, M. C. (2005). Implicit and explicit attitudes towards lesbians and gay men. *Journal of Homosexuality, 49,* 39-66.

Stephan, W.G., & Stephan, C.W. (1985). Intergroup anxiety. *Journal of Social Issues, 41,* 157–175

Stevenson, M. R. (1988). Promoting tolerance for homosexuality: An evaluation of intervention strategies. *Journal of Sex Research, 25,* 500-511.

*Story, M. D. (1979). A longitudinal study of the effects of a university human sexuality course on sexual attitudes. *Journal of Sex Research, 15*, 184-204.

Szymanski, D. M., Kashubeck-West, S., & Meyer, J. (2008). Internalized heterosexism: A historical and theoretical overview. *The Counseling Psychologist, 36*, 510-524.

*Terrizzi, J. A., Jr., Shook, N. J., & Ventis, W. L. (2010). Disgust: A predictor of social conservatism and prejudicial attitudes toward homosexuals. *Personality and Individual Differences, 49*, 587-592.

Tucker, E. W., & Potocky-Tripodi, M. (2006). Changing heterosexuals' attitudes toward homosexuals: A systematic review of the empirical literature. *Research on Social Work Practice, 16*, 176-190.

*Turner, R. N., Crisp, R. J., & Lambert, E. (2007). Imagining intergroup contact can improve intergroup attitudes. *Group Processes and Intergroup Relations, 10*, 427-441.

Warner, M. (1993). Introduction. In M. Warner (Ed.), *Fear of a queer planet: Queer politics and social theory* (pp. i-xxxi). Minneapolis, MN: University of Minnesota Press.

*Webster, R. J., & Saucier, D. A. (2011). The effects of death reminders on sex differences in prejudice toward gay men and lesbians. *Journal of Homosexuality, 58*, 402-426.

Weinberg, G. (1972). *Society and the healthy homosexual.* New York, NY: St. Martin's Press.

Wenzlaff, R. M., & Wegner, D. M. (2000). Thought suppression. *Annual Review of Psychology, 51*, 59-91.

Wiederman, M. W. (1999). Volunteer bias in sexuality research using college student participants. *Journal of Sex Research, 36*, 59-66.

Table 1

Keywords for Searching Studies on Reducing Sexual Prejudice

Intervention	Sexual Prejudice
Challenge	Anti-gay/ anti homosexual/ sexual prejudice
Change	Anti-gay/ homophobic etc. violence/
Educate/ation	sentiment/ bullying/ harassment
Improvement/improve	Attitudes towards gay/lesbian people etc.
Intervention	Biphobia
Modification/modify	Gay/ lesbian/ homosexual etc. stereotypes
Prevention/prevent	Heterocentrism/heterocentric
Reduction/reduce	Heterosexism/heterosexist
	Homonegativity/homonegative
	Homophobia/homophobic
	Hostility towards gays/ lesbians etc.
	Lesbophobia
	Monosexism
	Prejudice against gay/ lesbian people etc.
	Sexual prejudice
	Sexual/ anti-gay stigma(tization)

Table 2

Classification of Interventions to Reduce Sexual Prejudice

Approach	Description
Education	Information on (homo)sexuality, LGB lives, or prejudice, through either lectures, educational films, scientific readings, or a combination of these in the form of a course or workshop.
Intergroup contact	Contact with lesbians, gay men, or bisexual people in an organized setting, e.g., a panel presentation; it does not imply physical presence: contact may be imagined, vicarious, or otherwise mediated.
Contact-plus-education	Education and intergroup contact used together.
Norms or expertise	Information on how prejudice is viewed by either experts (e.g., evolutionary psychologists) or a significant group (e.g., public opinion, peers).
Inducing emotions	Exercises that directly target participants' emotions towards LGB people, including the facilitation of empathy.
Priming techniques	Participants' identity or values made salient in a certain situation; what is primed may be directly relevant to prejudice (e.g., tolerance) or more general (e.g., self-worth).
Awareness or suppression	Participants instructed (or otherwise prompted) to either recognize or suppress their prejudice.
Accountability	Participants prompted to explain their prejudiced beliefs or behaviours, or are otherwise held responsible for them.
Entertainment	Recreational books, films, or shows with content expected to

	influence prejudice.
Cooperative learning	Participants and LGB people studying together, esp. in a jigsaw-classroom setting.
Peer debate	Participants discussing their beliefs and feelings with peers.
Cognitive training	Exercises to retrain stereotypes.
Manipulation of categories	Specifically-devised situations that prompt participants to change the way they categorise others (e.g., acknowledge that one persons belongs to multiple categories)
Comparison of approaches	Two or more of the above approaches compared in the same study.

Note. This classification is based on Paluck and Green (2009).

Table 3

Results of the Meta-Analytic Reviews by Type of Intervention and Outcome Measure

Intervention	Outcome	*k*	*d*	95% confidence interval for *d*	
Education	Attitude	32	0.46	0.33	0.59
	Knowledge	5	1.09	0.52	1.88
	Emotions	5	0.36	0.26	0.47
Contact	Attitudes	8	0.56	0.25	0.89
Contact-plus-education	Attitudes	27	0.41	0.28	0.52
	Emotions	6	0.44	0.28	0.62
	Intentions	5	0.35	0.18	0.52
Norms and expertise	Attitudes	5	- 0.02	-0.16	0.19
	Behaviour	4	0.46	0.21	0.72

k = the number of studies on which *d* is based.

Figure 1. Flowchart of searching and selecting studies.

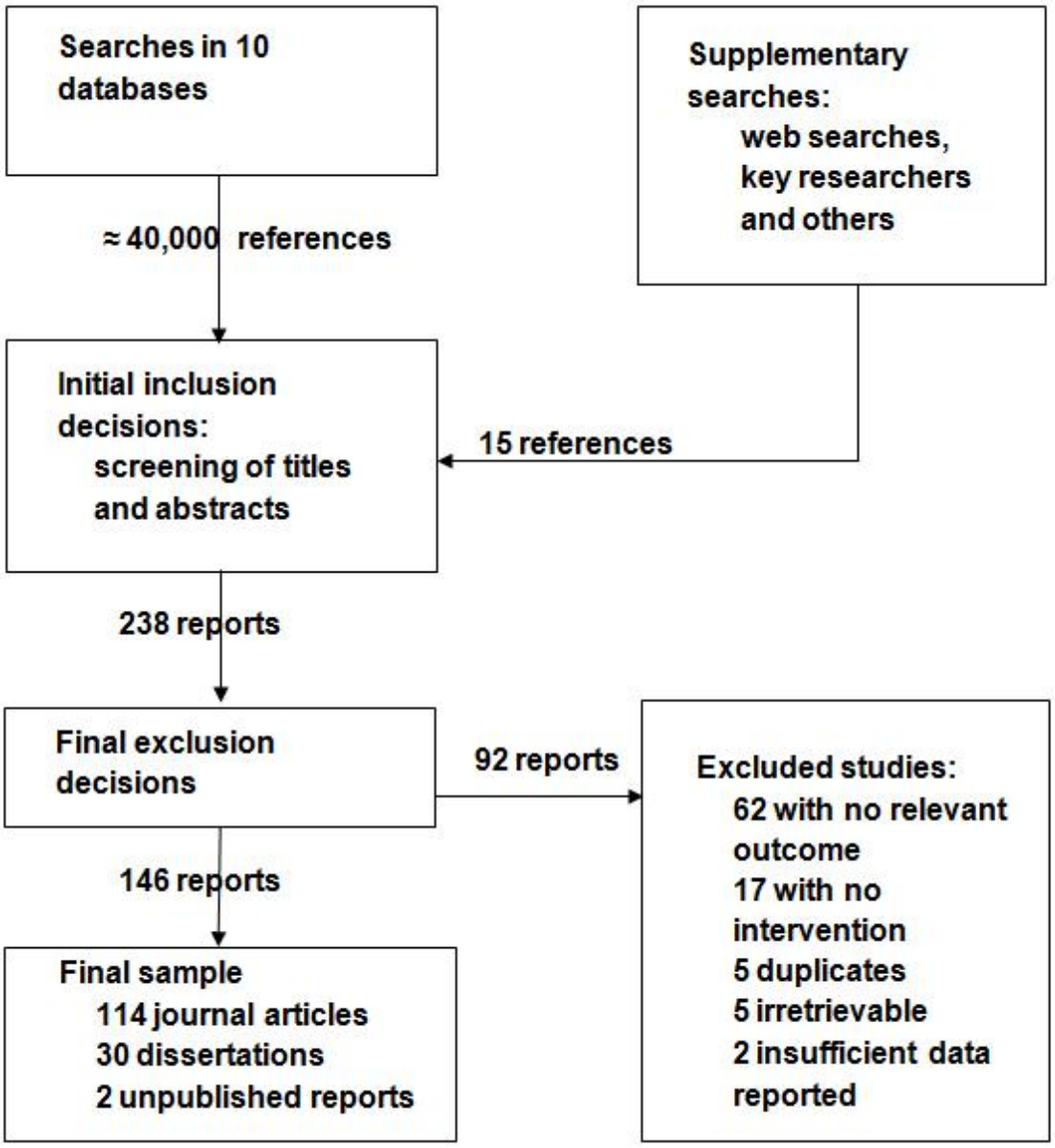


Figure 2. Effect of education on attitudes.

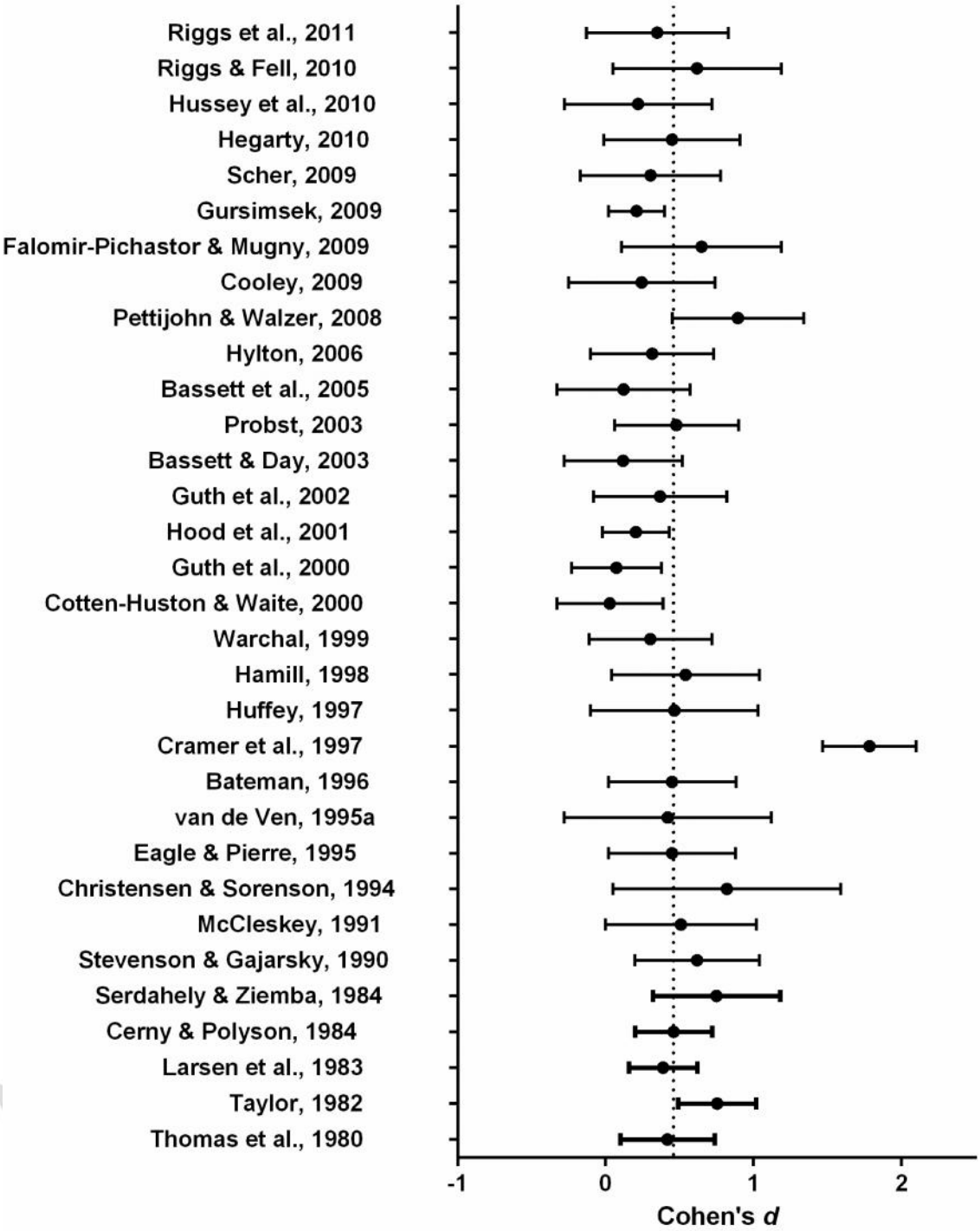


Figure 3. Effect of education on knowledge.

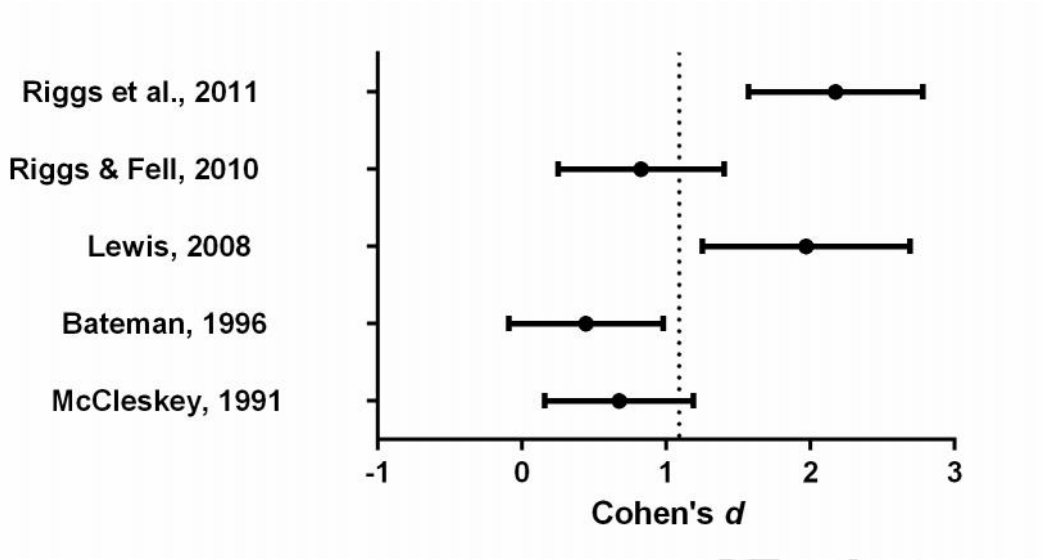


Figure 4. Effect of education on emotions.

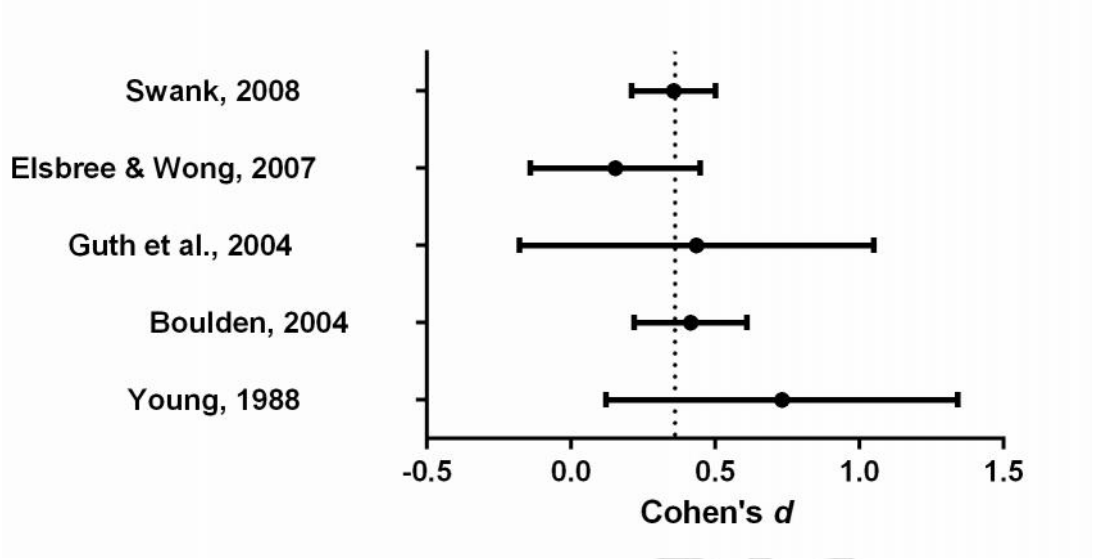


Figure 5. Effect of contact on attitudes.

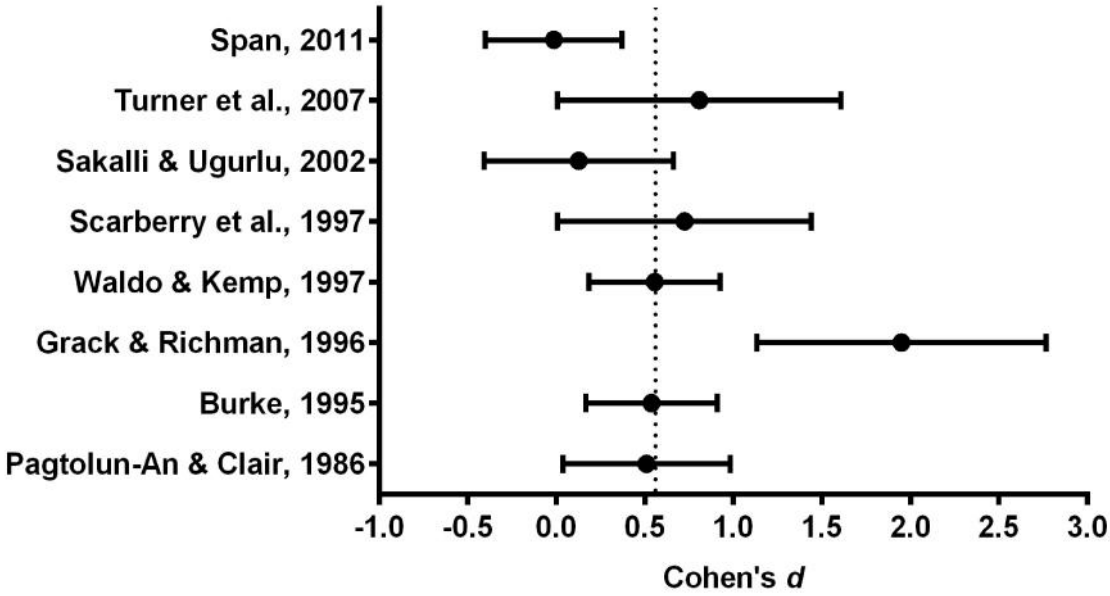


Figure 6. Effect of contact-plus-education on attitudes.

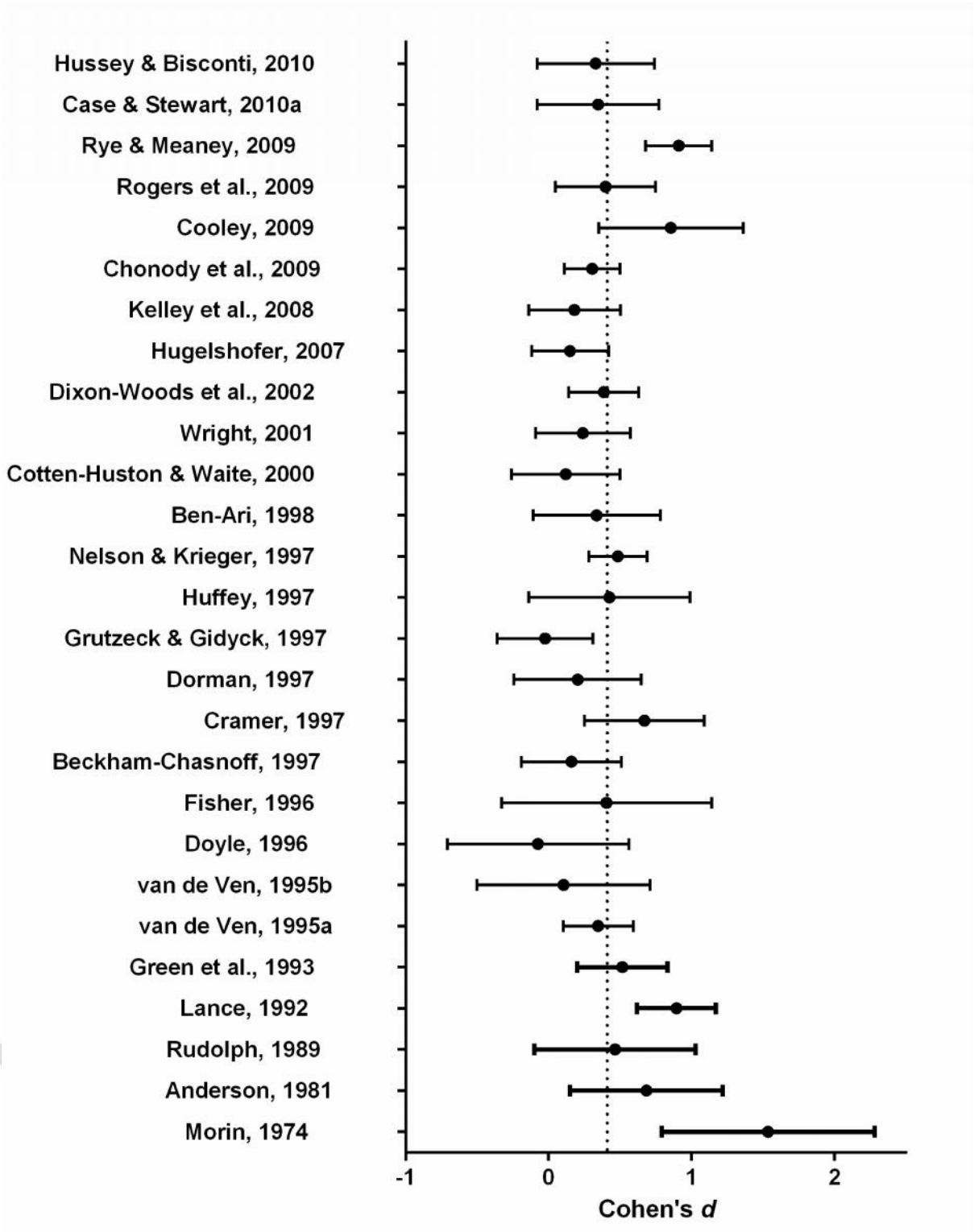


Figure 7. Effect of contact-plus-education on emotions.

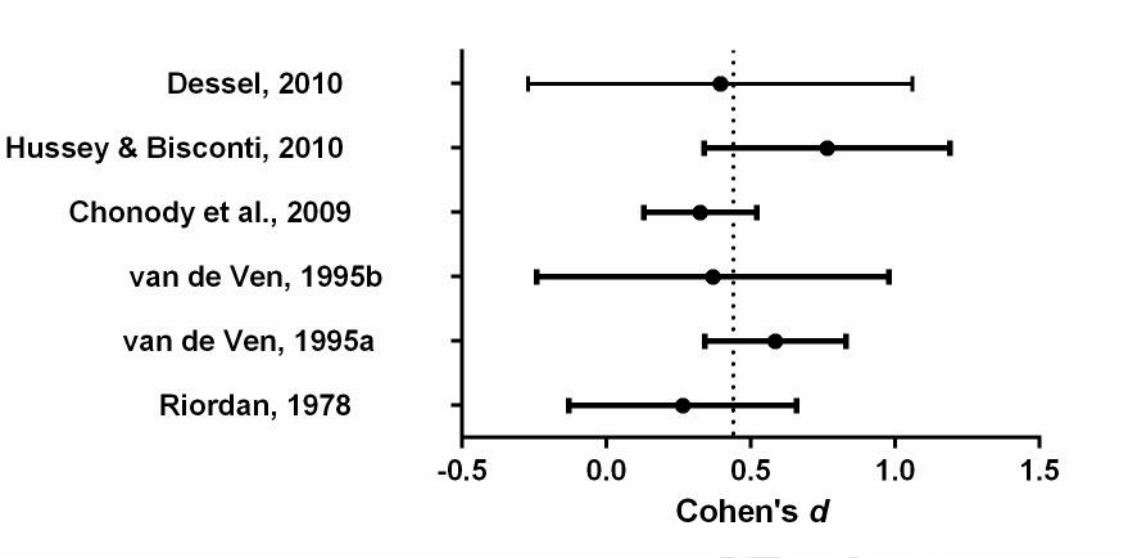


Figure 8. Effect of contact-plus-education on behavioural intentions.

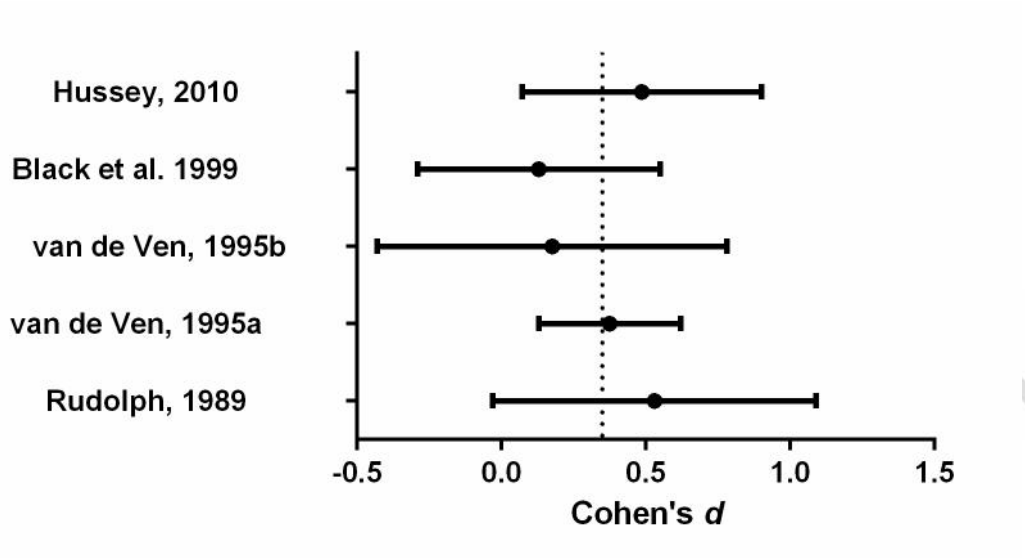


Figure 9. Effect of norm-and-expertise interventions on attitudes.

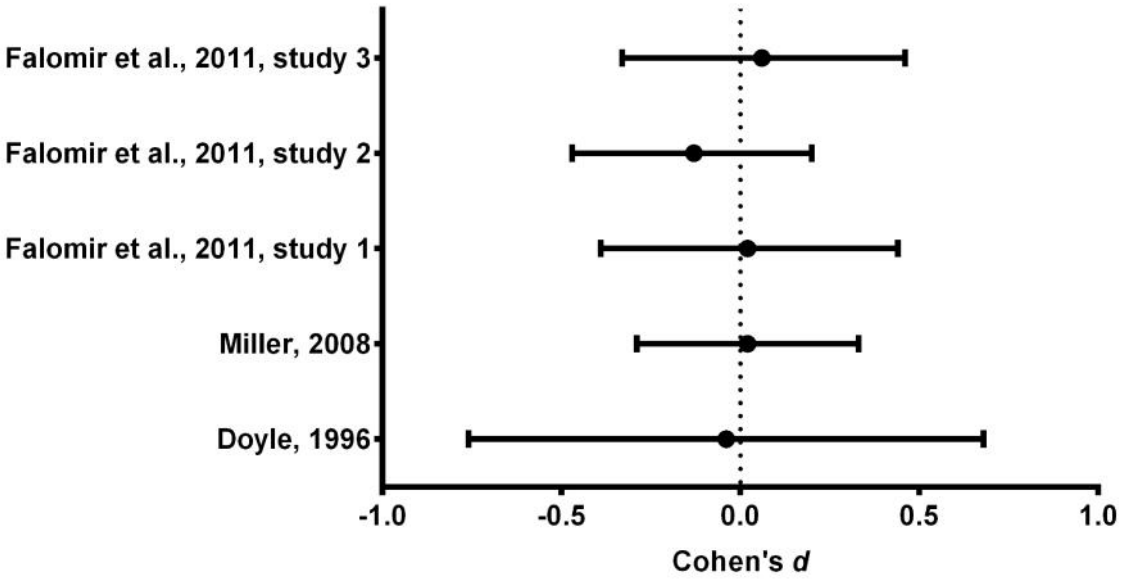
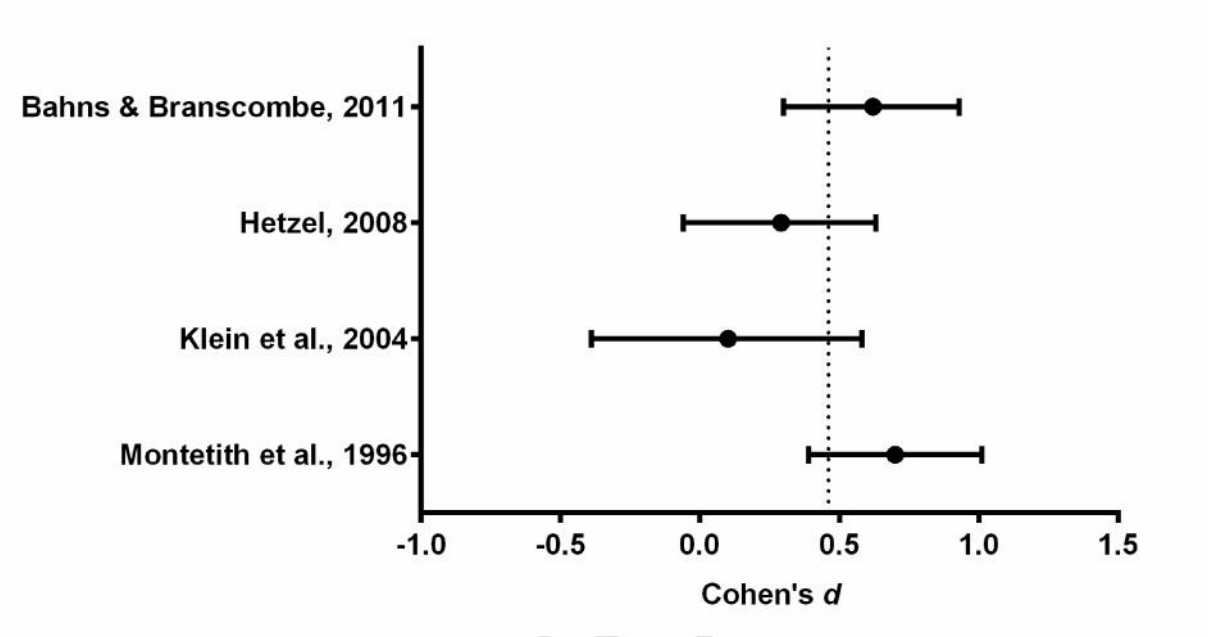


Figure 10. Effect of norms-and-expertise interventions on behaviour.



Footnotes

[1] The term *homophobia* is usually credited to George Weinberg, who used it in his 1972 book *Society and the Healthy Homosexual*. However, the term was in use earlier, a fact readily acknowledged by Weinberg himself (personal communication cited in Herek, 2004). The earliest academic paper using this word seems to be Kenneth Smith's (1971) "Homophobia: A Tentative Personality Profile."

[2] Note that *homosexuality* had previously enabled a similar focus on the individual psyche at the expense of other issues (Sell, 1997).

[3] Only the variables used in the final version of this paper are reported here. The initial coding scheme, which is slightly more detailed, is available from the authors upon request.

[4] We used the same symbol (*d*) both for the effect sizes of individual studies and for summary effect sizes; we appreciated that the context would always be clear enough to avoid confusions.

[5] Begg and Mazumdar's τ should be interpreted with caution when the number of studies is small; see Field and Gillet (2010) for details.

[6] Rosenthal's fail-safe number is meaningless in this case: since the mean effect size is naught, there is no need to consider the possibility of unpublished studies with nonsignificant results.

[7] One might question whether accountability induces a change in people's attitudes or merely a socially desirable behaviour. See Crandall, Eshleman, and O'Brien (2002) for a more sophisticated view on the matter.