CREATE

Canterbury Research and Theses Environment

Canterbury Christ Church University's repository of research outputs

http://create.canterbury.ac.uk

PETER BRENNAN BSc Hons MSc

AN INVESTIGATION INTO GENDER NONCONFORMITY AND DEPRESSION.

Section A: An investigation of the association between gender nonconformity and depression
Word Count
7958 (40)

Section B: A study of the association between childhood gender nonconformity, self-esteem and childhood depression
Word Count
7139 (334)

Overall Word Count
15097 (374)

A thesis submitted in partial fulfilment of the requirements of Canterbury Christ Church University for the degree of Doctor of Clinical Psychology

APRIL 2017

SALOMONS
CANTERBURY CHRIST CHURCH UNIVERSITY

Summary of the Major Research Project

Section A

This review investigated the association between gender nonconformity (GN) and depression, including the mechanisms that might explain this association. Several papers found an association, with one finding that abuse and bullying victimisation mediates the association. Other papers found a stronger association in males in comparison to females, with more recent papers also comparing rates of GN and depression in heterosexual and sexual minority populations. Methodological strengths and limitations of each paper were discussed in addition to implications for future research and clinical practice.

Section B

This study investigated the association between childhood gender nonconformity (CGN) and childhood depression, including whether any association was mediated by self-esteem. The dataset was obtained from the Avon Longitudinal Study of Parents and Children (ALSPAC). In females, self-esteem was found to mediate the association between CGN and childhood depression. This mediation was not apparent in males and the effect size in females was very small. Limitations include the high levels of missing data and research implications include the need for more UK studies in addition to studies in different countries and cultures. Interventions such as community psychology, public health and school interventions may help increase awareness of the stigma experienced by GN individuals. Effective clinical interventions should be made available.

Table of Contents

Section A: An investigation of the association between gender nonconformity and depression.

Section B: A study of the association between childhood gender nonconformity, self-esteem and childhood depression

List of Tables and figures

List of Appendices

Section A

Section B

MAJOR RESEARCH PROJECT: SECTION A

An investigation into the association between gender nonconformity and

depression

A literature review

Accurate Word Count: 7958

Abstract

As there is a growing body of evidence of an association between gender nonconformity (GN) and depression, this review aimed to investigate and evaluate the evidence for this association. Literature searches were performed on three on-line databases, using terms such as gender atypicality / nonconformity / variance, in combination with depression. Thirteen papers were identified which investigated these associations; five related to childhood GN (CGN) and eight to adolescent or adult GN. Several papers found an association between GN and depression with some finding biological sex differences, where stronger associations were found in males and sexual minority groups. Eleven papers included a cross sectional design and it was therefore not possible to infer causal relationships between study variables in these studies. Other methodological issues, such as limited sampling, and issues with measures of key variables, have been discussed.

Many studies also investigated mediating and moderating variables, such as bullying victimisation, abuse and quality of parental relationships and several significant associations with GN and depression were identified,. Research implications include the requirement for more longitudinal studies, UK based studies, studies in non-westernised societies and those containing robust, reliable and valid measures. Only one study investigated CGN and childhood depression. Community, public health and school-based interventions will help improve awareness and tackle GN stigmatisation. Tailored emotional support to GN individuals and parental support are necessary.

Introduction

Gender nonconformity and depression

Gender nonconformity (GN) may be defined as the expression of actions or traits that fall outside of society's expectations of an individual's biological sex i.e. being more like the opposite sex in terms of interests, gestures, behaviour, speech and play preferences (Collaer & Hines, 1995; Lippa, 2000; 2002;). When assessing GN in children we consider gender atypical play, such as boys playing with dolls, or activities such as girls climbing. Traits such as risk taking or exploring new surroundings would be considered atypical among young girls.

Various studies have demonstrated that childhood and adolescent gender nonconformity (GN) are associated with increased rates of depression, anxiety and psychological distress (Collier, Boss & Sandfort, 2013; Roberts, Rosario, Slopen, Calzo & Austin, 2013; Smith & Leaper, 2006). However, the findings are not always very clear and more research is needed to increase confidence in the study findings.

Several studies which investigated the relationship between GN and mental health issues, in addition to their mediation by peer victimisation, have reported a biological sex difference, whereby GN males reported higher levels of mental distress in comparison to GN females (D'Augelli, Grossman & Starks, 2006; Young & Sweeting, 2004). For example, Young & Sweeting (2004) found that 15-year-old gender nonconforming boys reported more loneliness, peer victimisation and reduced psychological well-being, in comparison to GN girls.

In order to understand the effect of GN on an individual it is first necessary to consider the psychology of acquiring gender roles. Social learning theories such as those of Mischel (1970) and Bandura (1977) hypothesised that gender roles were learnt at an early age through observation of, and identification with, individuals of the same gender. These roles were reinforced through a series of rewards and punishments (Skinner, 1953).

Attributes such as independence, agency and activity were considered masculine, whereas characteristics such as emotionality, cooperation and superior communication skills were considered feminine (Bem & Bem, 1974; Spence & Helmreich, 1978). Male and female gender roles are usually focused on notions of masculinity and femininity respectively.

Discussions on the degree to which gender roles are socially constructed and the extent to which they are biologically determined are ongoing. Biological sex is assigned at birth and is based on the individual's internal reproductive anatomy, external genitalia, the existence or absence of a Y-chromosome, the type of gonads and sex hormones (e.g. Geary, 2009). Geary (2009) argues that the best way to understand the difference in male and female gender roles is through the evolutionary process. However the World Health Organization (2014) considers gender roles to be socially constructed, meaning that gender roles are created through societal and cultural expectations.

Cognitive developmental theories such as gender schema theory hypothesised that children developed theoretical constructs of femininity and masculinity, known as schemata. These schemata influenced their personality, temperament, self-perception, perception of others and behaviours (Trew & Kramer, 1998). Bem (1985) proposed that the societal expectations placed on different genders and the subsequent roles taken on by different genders, influenced the development of the above schemata. As a cultural phenomenon,

gender roles are subject to change over time (Bem, 1993), thereby influencing the development of gender specific schemata (Trew & Kramer, 1998).

Masculinity and femininity have been conceptualised as polar opposites on a one-dimensional continuum of gender related characteristics and behaviours (Foushee, Helmreich & Spence, 1979). This type of construct carries the inherent assumption that masculinity and femininity are polar opposites and mutually exclusive. Therefore, the more feminine an individual, the less masculine they are judged to be. However, theorists (e.g. Bem & Bem, 1974) who proposed that it was possible for an individual to possess masculine and feminine characteristics simultaneously have challenged this position.

Tajfel's social identity theory (1982) proposes that the groups to which an individual belongs plays a crucial role in their self-esteem and confidence. Tajfel and Turner (2004) proposed three sequential psychological mechanisms of social categorisation, identification and comparison. For example, a GC male may categorise other men as either GC (in-group) or GN (out-group). This helps him to understand himself, others and the social setting in which he operates.

Social identification then follows, whereby the individual embraces the group identity. For example, the individual may take on the stereotypical characteristics of GC males such as assertiveness, agency and independence. Social comparison then follows, where the GC male may compare himself to the GN out-group. In order to preserve his self-esteem he will perceive his GC in-group to be preferable to the out-group. This may help to explain why GN men experience discrimination and reduced self-esteem at the hands of the bigger, more powerful and socially acceptable GC in-group.

More recent theorists have introduced the concept of gender fluidity. For instance, Renold (2004) highlighted the move towards a more sophisticated conceptualisation of gender roles, whereby gender expression may fluctuate in different social situations, settings and at different times (Butler, 1990).

GN is occasionally confused with gender dysphoria. However, whereas GN refers to disparities in an individual's gender related behaviour in comparison to societal norms, gender dysphoria (GD) relates to an individual's anguish with respect to their dissonant gender identity. While most GN individuals do not experience GD, many GD individuals do also experience GN. Additionally, both are associated with sexual minority status.

GD in childhood is apparent in the child's choice of opposite sex playmates, play, clothing and characteristics considered more typical of the opposite sex. In the majority of cases, childhood GD does not remain in adulthood, with many individuals identifying with a sexual minority status in adulthood (Clark et al., 2014; Conron, Scott, Stowell & Landers, 2012). GD that persists throughout adolescence tends to remain in adulthood (Zucker & Bradley, 1995).

Gender nonconformity is associated with homosexuality (e.g. Steensma, Van Der Ende, Verhulst & Cohen-Kettenis, 2013). One study, reported that the majority of gay men who responded to a study of gender nonconformity (GN) (Taywaditep, 2001) stated that they were gender nonconforming in childhood. A significant proportion resorted to gender conforming behaviours in adolescence and adulthood, possibly because of societal expectations including those of parents and friendship groups (Whitam, 1977). Additionally, GN individuals, especially males, experience mockery, taunting, exclusion and isolation (Harry, 1983; Pilkington and D'Augelli, 1995). From the perspective of social identity

theory, these individuals may have been attempting to join the in-group of GC males, in an attempt to improve their self-esteem.

However, many GN individuals identify as heterosexual and this highlights the importance of implementing studies that sample the entire population, regardless of sexual orientation. These studies demonstrate that irrespective of sexual orientation, higher levels of GN result in increased instances of homophobic name-calling (van Beusekom, Baams, Bos, Overbeek, Sandfort, 2016) and reduced levels of well-being (Rieger & Savin-Williams, 2012).

The situation is often complicated by the fact that people who are gender nonconforming are assumed to belong to a sexual minority group, whereas individuals who conform to gender norms are assumed to be heterosexual (Johnson & Ghavami, 2011; Valentova, Rieger, Havlicek, Linsenmeier & Bailey, 2011). This partially explains why GN individuals experience increased levels of victimisation among their contemporaries, as a proportion of the victimisation is homophobic in nature (Pascoe, 2007; Plummer, 2001). Research also demonstrates that sexual minority teens and young adults report that there is a positive correlation between GN and verbal homophobic abuse (D'Augelli, Grossman & Starks, 2006). Abuse related to perceived LGBT status has been found to mediate the relationship between GN and mental health issues in LGBT teens and young adults (Baams, Beek, Hille, Zevenbergen, & Bos, 2013; Sandfort, Melendez & Diaz, 2007; Toomey, Ryan, Diaz, Card & Russell, 2010).

Additionally, beliefs around gender roles are constantly evolving and vary across cultures. Therefore, it may not be possible to assume that studies from different cultures are generalisable to other cultures. Furthermore, studies implemented several years ago on one cohort of individuals may not be generalisable to other more recent cohorts because of

shifting cultural attitudes. These facts highlight the need for a review of the literature, in addition to the requirement for further studies in this area.

In summary, gender roles and expectations are established in early childhood. Most societies hold very strong and rigid beliefs about gender appropriate behaviour and deviations from this behaviour often results in societal disapproval, which may manifest in the form of abuse, bullying and victimisation from family members, peers and other adults. Additionally, gender nonconforming children may pick up more subtle, indirect forms of societal disapproval and stigmatisation, resulting in reduced self-esteem. It is hypothesised that these factors, amongst others, may mediate the relationship between GN and depression.

Aim of the review

With these ideas in mind, this paper aims to review the literature on the associations between childhood, adolescent and adult GN and depression in childhood, adolescence and adulthood. It will also investigate the hypothesised mechanisms that attempt to explain these associations. It is hoped that by producing a comprehensive analysis and review of the evidence for a relationship between GN and depression, in addition to the proposed mechanisms that may explain or influence this relationship, it will be possible to identify gaps in the literature. This will lead to recommendations for further studies in order to gain a more complete picture of the above relationship and its mechanisms. This in turn could result in the development and implementation of targeted interventions that seek to minimise the impact of these mechanisms, where practicable. The interventions could be directed at individuals, families or wider community interventions.

Method

Literature search

An electronic literature search was performed on PsychInfo, Medline and the Applied Social Sciences Index and Abstracts (ASSIA) databases, *on* 02/09/2016. The search items applied to each database were ("Gender nonconform*" or "Gender non-conform*" or "Gender atypical*" or "Gender varia*") AND ("Depress*").

The '*' symbol ensured that different terms were included in the search, such as variation or variance, for example. The 'or' operator ensured that the various relevant terms would be include, while the 'and' operator combined the areas of interest and excluded studies which did not investigate both areas of interest.

A flowchart of the search is provided on page 8 and shows that when all searches were combined, 168 papers were identified, 35 of which were excluded as duplicates. After reading the titles of the 133 remaining papers and subsequently the abstracts, and excluding those that were irrelevant, 24 papers remained. The inclusion criteria are indicated below. Upon reading these articles, a further seven were excluded as they were also deemed irrelevant. After reading these articles, four were excluded as they conceptualised GN as a mediator as opposed to an input variable. This is because this review is specifically interested in papers in which GN is the input variable and depressive symptoms or other measures which include depression, are the output variable. This allows the consideration of the role of various mediating variables in the association between GN and depression. The 13 remaining studies are included in the review (section 4) and are summarised in Table 1.
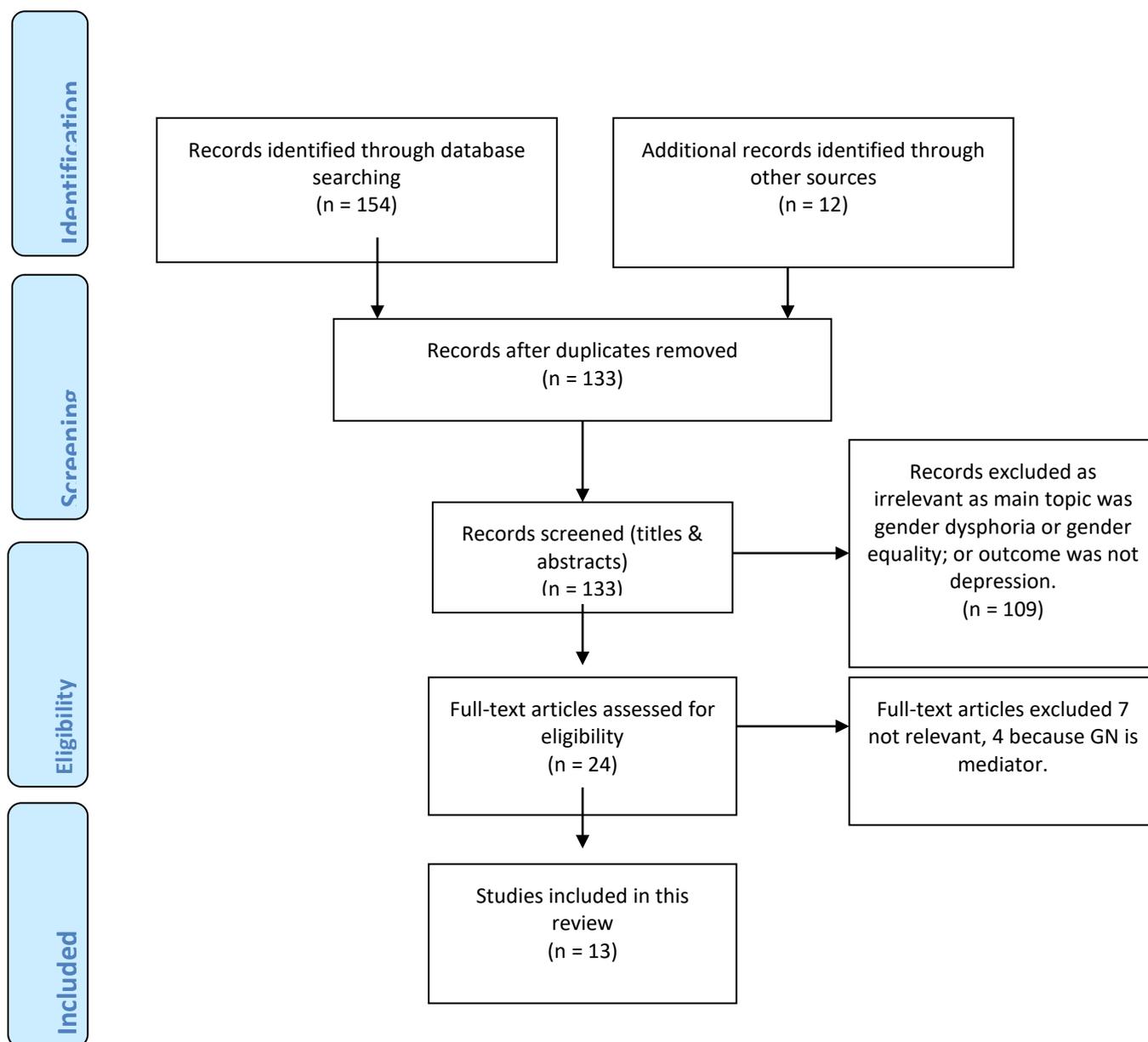
Figure1. Flowchart of Literature Searches.

Inclusion and exclusion criteria

Articles utilising a quantitative methodology and published in English language, peer reviewed academic journals were included in the review. Studies were included if they investigated the relationship between gender nonconformity / atypicality / variance and depression. Additionally, studies with an outcome of psychological distress or mental distress were included if they included depression items in their measure. Furthermore, one study, which referred to psychological well-being, was included. However, these studies received less weighting in comparison to the studies with an outcome of depression. There was no restriction on the age of participants.

Studies in which GNC is a mediator variable, as opposed to an input variable, have been excluded.

Review structure

This review will provide a summary of each study's aims, design, sample features, intervention, measures, findings and limitations. Papers will be grouped together either as childhood GN studies or as adolescent / adulthood GN. They will then be critiqued as a group as many of the criticisms will apply to many studies. This will provide an understanding of the general pattern of findings.

Study evaluation

The Effective Public Health Practise Project tool (EPHPP, 2009) has been used as an aid to the systematic evaluation of the methodological quality of the studies.

Review

Thirteen papers were reviewed, 11 of which were cross sectional, with the remainder being longitudinal. Studies were based primarily in Europe and North America, the exception being two South African studies. There was a significant variance in sample sizes across studies, ranging from 94 to 10665.

All papers reviewed the association between GN and depression, while some additionally explored the mechanisms that may explain this relationship. Five papers focused on childhood GN, and eight focused on adolescent or adult GN. The former are reviewed in section 4.1, while the latter are reviewed in section 4.2. Moreover, longitudinal studies have been explored in more depth. While the majority of studies used depression as their outcome variable, four studies used psychological distress or psychological well-being. The studies measuring psychological distress used combinations of depression and anxiety questionnaires (e.g. Skidmore et al., 2007), or measures which incorporate items on depression and anxiety (e.g. Sandfort et al., 2007).

Information on sample size / composition and study measures are provided in Table 1. As most measures were self-report please assume that this is the case, unless otherwise stated.

The associations between childhood gender nonconformity and depression.

Findings

Of the five studies which investigated the associations between childhood GN (CGN) and later depression or distress using retrospective measures of CGN, one was longitudinal (Roberts, Rosario, Slopen, Calzo & Austin, 2013[1]) and four were cross sectional (Alanko et

al., 2009[4]; Feinstein, Goldfried & Davila, 2012[3]; Petterson, VanderLaan & Vasey, 2016[5]; Skidmore, Linsenmeier & Bailey, 2006[2]).

Roberts et al., (2013) investigated the relationship between CGN, bullying victimisation and depression. The data were obtained from a national cohort study and consisted of 10,655 records. CGN was assessed retrospectively in 2005, using four questions from the Recalled Childhood Gender Identity / Gender Role Questionnaire (RCGIGR; Zucker et al., 2006). Participants were divided into three categories depending on their CGN score: those participants with a CGN score less than or equal to the sample median (gender conforming – GC); those with CGN score above median but below the top decile (gender nonconforming – GN); and those in the top decile (highly gender nonconforming).

The results demonstrated that, between 12 and 30 years, those participants in the top decile of CGN had significantly higher levels of depression in comparison to the gender conforming group. Further analyses demonstrated that approximately 50% of the increased rates of depression could be explained by abuse and bullying victimisation. The increased rates of depression were particularly apparent in heterosexuals and males. Hence, the key finding in this study was that early CGN contributed to elevated rates of depression beginning in adolescence.

Skidmore et al., (2006) found an association between retrospectively recalled CGN and psychological distress in adulthood (a composite indicator of depression and anxiety measures) in gay men, but not in lesbian participants. However Alanko et al., (2009) found that retrospectively recalled CGN was associated with later psychological distress in adulthood (again a composite indicator of depression and anxiety symptoms) regardless of gender or sexual orientation. They also found a negative correlation between CGN and recalled parental relationships in childhood, especially in males. While rates of parental

coldness and over-control were associated with psychological distress, the effects of parenting on distress were the same for GC and GN participants.

Feinstein et al., (2012) utilised path analysis to investigate the mediation roles of internalised homonegativity and rejection sensitivity in the associations between retrospectively recalled CGN, discrimination and adulthood depression in a cross sectional study sample of 467 gay men and lesbians. Internalised homonegativity refers to the internalisation of negative societal attitudes towards homosexual identity, behaviours and relationships, while rejection sensitivity involves the belief that an individual will experience rejection based on their sexual minority status (Meyer, 2003).

They found a good fit in terms of the data and model, with significant associations between CGN, discrimination and depression / anxiety being mediated by internalised homonegativity / rejection sensitivity, respectively. Petterson et al., (2016) noted that retrospectively recalled CGN was associated with increased rates of retrospectively recalled depression in childhood. This is the only study in the review that investigated childhood depression and this highlights a pressing need for further studies in childhood depression. The other studies investigated current depression in adulthood (Skidmore et al., 2006; Alanko et al., 2009 & Feinstein et al., 2011), whereas Roberts et al., (2013) studied depression from early adolescence to early adulthood.

Detailed information on sample size and composition in all reviewed studies, in addition to measures used in each study, are provided in Table 1. Larger sample sizes permit increased statistical power and confidence in findings, while representative samples, such as those that sample heterosexual and sexual minority participants, allow more generalisability of the findings.

With the exception of Petterson et al., (2016) all of these measures were self-report. They were all current measures of depression and therefore the mean age of respondents can be gained from the above information on mean sample age.

Analyses

Roberts et al., (2013) determined whether GN children were at increased risk of depression in comparison to GC children, by comparing the means of their depression scores across several assessments from 1999 to 2010.They also utilised the CESD-10 (Radloff, 1977) in 2007 and 2010 in order to establish the number of cases of mild or moderate depression in GN children. Additionally, they calculated age-specific z-scores of depression by GN.

The next stage of their analysis involved assessing depression across participant's teens and early twenties by modelling GN as the input variable and depression as the outcome variable in a repeated-measures longitudinal model. In order to investigate whether the relationship between GN and depression differed by age, they tested age-by-nonconformity interaction terms. They also evaluated whether this relationship differed by sexual orientation, using a sexual orientation by GN term in a longitudinal model in which depression was the outcome variable, where sexual orientation was categorised as sexual minority or heterosexual. Additionally, depression was modelled as a function of GN and sexual orientation (SO), in order to ascertain whether GN predicted depression independently of SO.

The mediation variables of bullying victimisation and abuse (emotional, physical and sexual) were tested by adding them to the model as independent variables. The mediation effects were then calculated using the SAS Mediate macro. The authors ensured that the bullying victimisation measures preceded the depression measures in the model.

Furthermore, as the risk of victimisation due to GN differs by sex, as does the risk of developing depression, the authors assessed for the child's sex-by-non conformity term in every model tested. Moreover, all models adjusted for age and sex.

In order to avoid confounding between the effects of GN and sexual orientation on depression, matched pairs of heterosexual and sexual minority males and females with similar levels of GN were created. This was achieved by matching up pairs according to their RCGIGR questionnaire scores. The next stage of the analysis involved creating groups of GC and GN individuals for males and females, based on median levels of GC / GN scores. This then allowed the authors to analyse parental style differences for the GC and GN groups of participants. Group comparisons were performed in SPSS, in addition to multiple linear regression analyses.

Feinstein at al., (2012) tested their hypothesised model using path analysis with SPSS Amos, which allowed for the inclusion of multiple independent variables, such as CGN and dependent variables such as depression. Model fit was tested through use of the comparative fit index (CFI), the Tucker-Lewis index (TLI) and the root-mean-square error of approximation (RMSEA). In order to test experiences of discrimination, through internalised homonegativity, the authors employed bootstrapping techniques with the SPSS macro (Preacher & Hayes, 2008).

Skidmore et al., (2006) performed correlations between the main study variables in addition to group comparisons between gay men and lesbian participants. Using educational level as a proxy measure, the authors attempted to control for socioeconomic status and still found that GN predicted depression. Petterson et al., (2016) employed one-way analysis of covariance (ANCOVAs) in order to compare childhood GN, sexual orientation and gender.

They also incorporated covariates which differed by group, namely: age, education, religiosity and area of residence.

Strengths and Limitations

In terms of the Effective Public Health Practise Project (2009), several issues should be considered. For instance, the cross sectional study design of four of the studies means that it is not possible to infer causation from their results, as any proposed causal relationships could actually occur in reverse or be bidirectional. The vast majority of input, mediator and output variables were self-report, resulting in potentially biased results. All studies included retrospective self-report accounts of CGN, with the measure being taken several years after childhood, leading to potentially biased accounts from participants. For instance, participants experiencing significant mental health problems may apply a negative bias to their childhood memories. Recollection of previous experiences may also be influenced by current attitudes and behaviour as well providing a socially acceptable response, as perceived by the participant (EPHPP, 2009). The same principle applies to self-reported and retrospective accounts of negative childhood experiences and other potential mediators, which may also be affected by current levels of depression or other mood problems. Two studies only included lesbians and gay men in their samples (Skidmore et al., 2006; Feinstein et al., 2012) thereby limiting their generalisability to other sexual minority groups and to the heterosexual population (EPHPP, 2009).

Strengths of the Roberts et al., (2013) study included the fact that there was simultaneous assessment of bullying and depression between the ages of 12 and 30. During this period depression was assessed at several time points. Additionally, with such a large sample it was possible to consider gender and sexual orientation in evaluating the relationship between CGN and depression (EPHPP, 2009). This allowed the authors to demonstrate the

impact of CGN in the heterosexual population, in addition to the lesbian, gay and bisexual (LGB) population. Additionally, the longitudinal study design allows for the inference of causation between CGN, bullying / victimisation and depression.

All studies utilised reliable measures of independent, mediator and dependent variables. For example, reliable measures of CGN include the RCGIGR (Zucker et al., 2006); the CGNS (Bailey & Oberschneider, 1997) and the CGIS (Johnson et al., 2004). They also provided detailed descriptions of their analyses (EPHPP, 2009).

The associations between adolescent and adulthood GN and depression

Findings

Of the eight studies that investigated the associations between adolescent and adulthood gender nonconformity (GN) and depression or psychological distress, only one had a longitudinal study design (Li, Pollitt & Russell, 2016), with the remainder having a cross sectional design (e.g. Cook, Sandfort, Nel & Rich, 2013[6]; Baams, Beek, Hille, Zevenbergen & Bos, 2013[8]; van Beusenkom, Baams, Bos, Overbeek & Sandfort, 2016[9]). The study findings will now be discussed.

In their sample of black South African gay and bisexual men, Cook et al., (2013) found no significant differences in rates of depression between adult GN and GC males, in spite of higher rates of victimisation in GN participants in comparison to GC participants. This surprising finding led to a follow up South African study by Sandfort, Bos, Knox & Reddy[7] (2016) who reported similar findings to Cook et al., (2013) in that GN black South African gay and bisexual men experienced higher levels of discrimination, yet reported similar levels of depression in comparison to their gender conforming (GC) counterparts.

Baams et al., (2013) found that GNC participants as assessed in early adulthood, experienced significantly lower psychological well-being as compared to GC participants. This association was mediated by experiences of perceived stigmatisation. Psychological well-being was evaluated with 15 questions from the personal and social well-being module (Huppert et al., 2009).

Van Beusenkom et al., (2016[9]) found that adolescent GN was associated with psychological distress and this association was partially mediated by homophobic name-calling, which was moderated by levels of same sex attraction (SSA). The mediation effects were only significant at mean and high levels of SSA, and were more prominent in males than females. In another study, van Beusenkom, Bos, Overbeek & Sandfort (2015[10]) investigated the mediating role of parental acceptance and found that in males with high levels of father acceptance, adolescent GN was not significantly associated with psychological distress. These results contrasted with those of males who had low levels of father acceptance.

In their longitudinal study Li, Pollitt & Russell (2016[11]) found that young adult GN (18 – 26 years) was positively related to concurrent depression. It was also related to prospective depression (24 – 32 years) in heterosexual participants. Toomey, Ryan, Diaz, Card & Russell (2013[12]) demonstrated that young adult GN lesbian, gay and bisexual (LGB) participants (ages 21 – 25) experienced higher levels of depression in comparison to their GC counterparts and that this association was mediated by school victimisation. Sandfort, Melendez and Diaz, (2007[13]) reported similar findings, where GN gay and bisexual Latino men reported significantly higher levels of mental distress in comparison to their GC counterparts. This association was mediated by experiences of homophobia.

Analyses

In their longitudinal study, Li et al., (2016) developed multiple regression models to determine the impact of Wave 3 GN and SO on Wave 3 concurrent depression and Wave 4 prospective depression, while also controlling for demographic variables such as: age, ethnicity, parental education and financial problems.

As Cook et al., (2013) did not find a significant association between GNC and depression, a mediation analysis was not performed as the necessary conditions were not achieved (Baron & Kenny, 1986). Linear regression analyses assessed whether discrimination, gay community involvement or level of outness moderated GNC and depression. Initially, Baams et al., (2013) performed independent sample t-tests to evaluate sex differences in study variables such as GN. Subsequently correlations of key variables were performed by gender. In stage three, the Hayes (2012) Process macro was utilised to perform a bootstrapped mediation analysis, investigating the association between GN and psychological wellbeing, incorporating the potential mediator of experiences of stigmatisation and controlling for sexual minority status (SMS). In the final stage of the analysis a further bootstrapped mediation analysis was performed to evaluate the association between GN and psychological wellbeing, incorporating the potential mediator of experiences of stigmatisation and controlling for same sex attraction (SSA), was moderated biological sex.

Van Beusenkom et al., (2016) also utilised the Hayes (2012) Process macro in order to investigate whether the association between GNC and psychological distress was mediated by homophobic name-calling, by performing bootstrapped mediation analyses. The effect size for the mediation effects was indicated by the kappa-squared measure (Preacher & Kelly, 2011). Additionally, a bootstrapped moderation mediation analysis was performed in order to

determine whether homophobic name-calling was moderated by biological sex and same sex attraction (SSA). One analysis was performed for each potential moderator variable.

In contrast, van Beusenkom et al., (2015 b) performed multiple regression analyses separately by biological sex, in order to evaluate the potential moderating variable of parental acceptance on the association between GNC, SSA and psychological distress. Four interaction terms were entered simultaneously; GNC by mother / father acceptance and SSA by mother / father acceptance. Simple slopes analysis was utilised in the interpretation of significant interactions (Jose, 2013).

A different analytical process was utilised by Toomey et al., (2013) who implemented structural equation modelling (SEM) to evaluate the potential mediator of school victimisation due to perceived or actual sexual minority status (SMS), in the association between adult GN and depression. Standard measures of SEM model fit were utilised, including nonnormed fit index (NFI), CFI and RMSEA.

Following mediation analysis guidance (Baron & Kenny, 1986), Sandfort et al., (2007) performed their analysis across four stages. In the first stage, they investigated the association between GN and mental distress with multiple linear regression, controlling for age, recruitment location and acculturation. Using linear and logistic regression, they then investigated the associations between GN, potential mediators such as social support, experiences of homophobia, emotional, physical and / or sexual abuse / rape by partners or relatives; and mental distress. Those mediators which were associated with GN and mental distress at a level of $p <= .10$ were included in the final stage of the analysis, where each potential mediator was added individually to the regression model in order to evaluate their impact on the regression coefficient. Those that reduced the regression coefficient by 10% or

rendered it insignificant were included in the model. It is important to note that potential

confounders were controlled for, in all of the analyses.

Strengths and limitations

The cross sectional study design of eleven of the studies means that it is not possible

to infer causation from their results, as any proposed causal relationships could actually occur

in reverse or be bidirectional. The vast majority of input, mediator and output variables were

self-report, resulting in potentially biased results. All CGN studies included retrospective self-

report accounts of CGN, with the measure being taken several years after childhood, leading

to potentially biased accounts from participants. For instance, participants experiencing

significant mental health problems may apply a negative bias to their childhood memories.

Recollection of previous experiences may also be influenced by current attitudes and

behaviour as well as providing a socially acceptable response, as perceived by the participant.

The same principle applies to self-reported and retrospective accounts of negative childhood

experiences and other potential mediators, which may also be affected by current levels of

depression or other mood problems. Two CGN studies only included lesbians and gay men in

their samples (Feinstein et al., 2012; Skidmore et al., 2006) thereby limiting their

generalisability to other sexual minority groups and to the heterosexual population. This was

also the case in five adolescent or adult GN studies (Baams et al., 2013; Cook et al., 2013;

Sandfort et al., 2007; 2016; Toomey et al., 2013). Furthermore, studies that do not incorporate

ethnic and socioeconomic diversity into their study design may be less representative of the

general population.

Strengths of the Roberts et al., (2013) study included the fact that there was simultaneous assessment of bullying and depression between the ages of 12 and 30. During this period depression was assessed at several time points. Additionally, with such a large sample it was possible to consider gender and sexual orientation in evaluating the relationship between CGN and depression. This allowed the authors to demonstrate the impact of CGN in the heterosexual population, in addition to the lesbian, gay and bisexual (LGB) population. Additionally, the longitudinal study design allows for the inference of causation between CGN, bullying / victimisation and depression.

All CGN studies utilised reliable measures of independent, mediator and dependent variables. For example, reliable measures of CGN include the RCGIGR (Zucker et al., 2006; CGNS; Bailey & Oberschneider, 1977; CGIS; Johnson et al., 2004). Some adolescent or adult GN (AAGN) studies also used established, reliable measures of GN (Baams et al., 2013; van Beusenkom et al., 2016; 2015). *However,* others only used a single measure (e.g. Cook et al, 2013; Sandfort et al., 2007; Toomey et al.), which may not be satisfactory for such a complex, multi-dimensional construct.

The studies reviewed, utilised a variety of depression measures such as the Centre for Epidemiologic Studies Depression Scale (CES-D; e.g. Feinstein et al., 2012; Li et al., 2016) and the Depression and Anxiety Stress Scales (Sandfort et al., 2016). This may have an adverse effect on direct comparison of results. Moreover, several studies measured psychological distress and psychological wellbeing as opposed to depression, therefore calling into question the comparability of the different measures. Furthermore, the studies investigating psychological distress utilised different measures of varying specificity which again calls into question the validity of their comparison. For instance, the study by Sandfort et al., (2007) measured mental distress using only a 5-item measure covering anxiety,

depression and suicidality. It is questionable as to whether such complex multi-dimensional variables can be measured (represented) by such a small scale.

Another strength of most studies reviewed was that they provided detailed information on the statistical analyses performed.

Review Synthesis

Childhood GN

In comparison to the other studies, the relatively small sample size of Skidmore et al., (2006) leads to a lower level of confidence in their findings. Additionally, they only sample gay men and lesbians, as do Feinstein et al., (2012) and this limits the generalisability of their findings. The findings of Alanko et al., (2009) and Petterson et al., (2016) are more generalisable as they included the heterosexual population in their samples and their sample sizes were relatively large, thereby permitting increased confidence in their findings. Finally, the longitudinal design, large sample and detailed analysis of Roberts et al., (2013) permits a good level of confidence in their findings that abuse and bullying victimisation mediated the relationship between CGN and depression in adolescence and adulthood (EPHPP, 2009).

Adolescent and adult GN

Sample sizes were generally large enough to ensure sufficient power and therefore confidence in study findings. However, most studies sampled sexual minority participants or a subset thereof, meaning that they were not generalisable to the heterosexual population (EPHPP, 2009). Interestingly, in their study of heterosexual and homosexual participants, van Beusenkom et al., (2015) found that father acceptance in GN boys led to significantly lower

levels of psychological distress, in comparison to those with low levels of father acceptance. Another interesting finding from van Beusenkom et al., (2016) was that homophobic name-calling mediated the relationship between adolescent GN and psychological distress, at mean and high rates of same sex attraction. Both studies contained heterosexual and homosexual participants and are therefore more representative of the general population (EPHPP, 2009).

Methodological limitations aside, this review has identified an association between gender nonconformity and depression. Various mediators have been hypothesised and associations have been established with both gender nonconformity and depression.

Table 1

Summary of Review Articles

No	Authors	Study type and participants	Measures	Analysis	Findings	Outcome measures
1	Roberts, Rosario, Slopen, Calzo, Austin, 2013.	US national longitudinal cohort (N = 16,882 children). Data analysed from the 1999, 2001, 2003, 2005, 2007 & 2010 waves (when children were between 11 and 30 years old). Those providing a response to the CGN question in 2005 or 2007 were included	CGN: 4 items from the Recalled Childhood Gender Identity / Gender Role Questionnaire (RCGIGR). A nonconformity score for each gender was split into 3 groups: Below median CGN (GC), above median CGN (below top decile - GN) and top decile CGN (highly GN). Sexual Orientation: 2 Qs in 2007 regarding sexual feelings and contact. Depression: Depression Symptoms Scale of the McKnight Risk Factor Survey in 1999, 2001 & 2003. The CESD-10 in 2007 & 2010. Bullying victimization: A) measured in 2001 with a question from the WHO Health Behaviour of School-aged children survey. B) Emotional / Physical bullying victimization by an adult non-family member (prior to age 11): 2 Qs from the	Differences in depressive symptoms evaluated from ages 12 to 30 years by CGN before age 11 years. They analysed levels of bullying victimization by CGN. Then analysed whether abuse and bullying partially mediated the relationship between CGN and depressive symptoms. Results split by sexual orientation.	In comparison to gender conforming (GC) participants, highly GN participants reported significantly higher rates of depression through adolescence and early adulthood. In comparison to their GC participants, CGN males / females recorded higher rates of depression by factors of 0.34 / 0.19 standard deviations respectively. In the 2010 wave, 26% of highly CGN individuals reported depression, in comparison to 18% of GC individuals (p < .001). Approximately 50% of this variance was mediated by abuse and bullying victimisation. CGN heterosexuals and gay men reported significantly higher rates of depression	Depression.

		(N = 10,655).	Child Trauma Questionnaire (CTQ) / Conflict Tactics Scales (CTS) respectively. C) Sexual abuse victimization before age 11: 2 Qs about adult or older child touching, forcing or attempting to force sexual activity, by hurting, threatening or restraining the individual. D) Physical and emotional abuse prior to age 11 from a family member: 4 Qs from the CTQ. E) Physical aggression from a family member: 2 Qs from the CTS.		than GC participants.
2	Skidmore, Linsenmeier, Bailey, 2006.	A cross sectional study of 44 lesbians (M age of 31, SD = 9.3) and 50 gay men (M age of 35, SD = 9).	Recalled CGN: Childhood Gender Nonconformity Scale (CGNS). Current adult GN: Continuous Gender Identity Scale (CGIS). Observer ratings of GN. Psychological distress: A composite of the Beck Depression Inventory (BDI-II) and the State Trait Anxiety Inventory (STAI). Attitudes towards GNC: 50 items. Attitudes towards GNC Scale: 4 feeling thermometer items.		GN in gay men was associated with higher levels of psychological distress, in comparison to GC gay men. This result was not found in lesbian participants.	psychological distress

| 3 | Feinstein, Goldfried & Davila, (2012). | Cross sectional study. 218 lesbians and 249 gay men (total N = 467). Mean age = 31.24 (SD = 11.67, range = 18 - 72). 76% Caucasian, 6% Hispanic, 6% African American, 5% Asian, 6% Other. | CGN: Childhood Gender Nonconformity Scale (CGNS). Discrimination: LGB-inclusive version of the Heterosexist Harassment, Rejection and Discrimination Scale (HHRDS). Depression: Centre for Epidemiologic Studies Depression Scale (CES-D). BFNE-II Scale. Internalized homonegativity: subscale from the Lesbian, Gay and Bisexual Identity Scale (LSBIS) and the Gay-Related Rejection Sensitivity Scale (GRRSS). | Hypothesised model tested with Path analysis, with CGN and discrimination as DVs and including 2 potential mediators- internalized homonegativity and rejection sensitivity. | CGN significantly associated with discrimination. The association was significantly higher for gay men than lesbian women. Discrimination was significantly associated with internalised homonegativity & rejection sensitivity, which in turn were significantly associated with depression and anxiety respectively. | depression and social anxiety |
| 4 | Alanko, Santtila, Witting, Varjonen, Jern, Johansson, von Der Pahlen & Sandnabba (2009). | Cross sectional Finnish study, using Retrospective questionnaires. Matched pairs, 79 men (n=158) and 148 women (n=296), with equal levels of | Recalled CGN: The RCGIGR Q. Sexual Orientation: Sell Assessment of Sexual Orientation (2 items). Measure of Parenting Style which assessed aspects of care, abuse and over-control in people under 16. Psychological distress: Brief Symptom Inventory (BSI-18) (depression and anxiety subscales). | Childhood GAB was correlated with negative ratings of parental relationships. | CGN was associated with higher levels of psychological distress. Homosexual participants reported higher levels of maternal over-control, in comparison to heterosexual participants. Heterosexual males reported significantly higher levels of parental coldness in comparison to homosexual males. Heterosexual females reported significantly lower levels of parental | Psychological distress (composite of depression and anxiety scores) |

GAB, one with same-sex sexual attraction or behavior and one without.

coldness in comparison to homosexual females.

| 5 | Petterson, VanderLaan, Vasey, 2016. | Canadian cross sectional study. Sample: heterosexual men (n = 98; M age = 37.4, SD = 16.61), heterosexual women (n = 142; M age = 35.3, SD = 16.5), gay men (n = 289; M age = 35.8, SD = 14.6), and lesbian women (n = 69; M age = 27.9, SD = 12.7); Bisexual Participants were excluded. | CGN: Childhood Gender Identity Scale (CGIS). Adult GN: Gender Diagnosticity Scale (GDS). Sexual orientation: Kinsey scale enquiring about sexual feelings over the past 12 months. Participants described their sexual feelings over the past 12 months. Separation anxiety: A measure based on DSM-IV-TR & DSM-5. Depression and anxiety: Measures based on DSM-IV-TR for GAD, major depression, agoraphobia, specific phobias, OCD, panic disorder and social anxiety disorder. | Factor analyses produced three factors: (1) indicators of childhood separation anxiety, (2) indicators of childhood depression and anxiety, and (3) indicators of adulthood depression and anxiety. | Lesbian women reported significantly higher levels of childhood separation anxiety, than other groups. Additionally, gay men had significantly higher levels than heterosexual men. Heterosexual men reported significantly lower childhood and adulthood depression as compared to the other groups. Gender atypical behaviour (GAB) was associated with higher rates of depression in men, but not in women. | Depression and anxiety. |

| 6 | Cook, Sandfort, Rich, 2013. | South African Cross-sectional community-based survey. 353 Black gay-identified or bisexual men. | All measures created by authors. GN: A single item: "How do you present yourself?" Depression: A 5-item likert scale (alpha = 0.83) (included questions on sleep & suicidality). School / General Discrimination: A 3-item (alpha = 0.73) / 4-item (alpha = 0.77) likert scale respectively. A 5-item (alpha = 0.79) / 3-item (alpha = 0.73) likert scale for Gay Community Involvement / Level of outness respectively | No significant association between GN and depression. Regression analysis assessed whether discrimination, and / or gay community involvement and / or level of outness, moderated association between GN and depression. | No significant difference in rates of depression in GN gay black men in comparison to GC gay black men. Outness and involvement in the gay community does not mediate the association between GNC and depression. | depression |
| 7 | Sandfort, Bos, Knox, Reddy. (2016). | 199 Black South African gay men (ages between 18 and 40; M age, 26.65 years) | GN: 2 items on how participants compared to other men in terms of masculinity / femininity. SO: 3 items on sexual attraction, fantasies and experiences with M/F. Sexual minority stressors: Items on being emotionally or physically abused or victimised for their GN. Discrimination in past year using an adapted, validated scale (Herek & Berrill, 1992). Sexual identity confusion measured with 4 items from Mohr & Fassinger (2000). Internalized homophobia; 10-item validated measure (Mohr & Fassinger, 2006). Resilience factors: Openness about sexuality – 2 items | Hierarchical regression analysis run to assess potential mediators of stressors and resilience factors on potential association between GN and depression. Step 1: GN was included; Step 2: Risk and resilience factors added. Step 3: Interaction between GN and each stress & resilience factor added to regression model. | GNC males experienced more discrimination in comparison to the GC counterparts. However they did not experience higher rates of depression. | Depression and anxiety. |

			on who is aware of their SO; gay community involvement – scale developed by Vanable, McKirnan & Stokes, (1998); social support – 5-item scale (Dandona et al., 2005). Depression and anxiety measured with the Depression Anxiety and Stress Scales (Lovibond & Lovibond 1995).			
8	Baams, Beek, Hille, Zevenbergen, Bos, (2013).	A Dutch cross sectional study of sexual minority youth and young adults (106 females and 86 males, 16-24 years old). Males M = 18.9, SD= 2.3; Females M = 19.6, SD = 2.8.	Current GN: Adapted version of Continuous Gender Identity Scale (CGIS). Perceived experiences of stigmatisation: Experienced rejection scale - 7 items. Psychological Well-Being: Personal and social well-being module - 15 items. Same sex attraction (SSA): 1 Q "How would you describe your sexual orientation?"	The Hayes Process Macro (2013) was used to test whether perceived stigmatisation mediated the association between GN and psychological well-being. The analysis controlled for SSA.	Perceived experiences of stigmatization mediated the negative association between current GN and psychological well-being. This mediation was not moderated by biological sex.	Psychological wellbeing.
9	van Beusekom, Baams, Bos, Overbeek,	A Dutch cross sectional study of 1,026 adolescents	GN: Adapted version of the CGNS with separate scales for males & females. Homophobic peer victimization: Homophobic content target subscale	The Hayes Process Macro (2013) was used to investigate whether homophobic name calling mediates the relationship	Homophobic name-calling partially mediated the association between GN and social anxiety and psychological distress in males and females (stronger	social anxiety and psychological

	Sandfort, (2016).	(females - 509, males - 517) ages 11 to 16 (M age = 13.4).	(HCTS) modified. Social anxiety: Social Interaction Anxiety Scale (SIAS). Psychological distress (depressive and somatic symptoms): Brief symptom inventory (BSI). Sexual minority status (SMS): 1 item "Have you ever had romantic and / or sexual feelings for someone of the same sex?"	between GN and social anxiety / psychological distress.	in males). This mediating role varied as a function of levels of same-sex attraction (SSA) and biological sex. The mediation effects increased when levels of SSA increased. They were significant for teens with mean and high levels of SSA.	distress
10	van Beusekom, Bos, Overbeek, Sandfort, (2015).	A Dutch cross sectional study of 1,121 secondary school students (582 females and 539 males; Mage = 16.6).	Current GN: Adapted version of the CGNS. Same sex attraction (SSA): 1 item, "Do you feel sexually attracted to someone of your own sex?" Parental acceptance: Inventory of the Parent and Peer Attachment subscale (Cronbach's alpha for males was .79 & .76 for maternal & paternal acceptance respectively; and for females was .81 & .79 for maternal & paternal acceptance respectively. Psychological distress: Brief Symptom Inventory (BSI). 24 items covering depressive and anxiety symptoms. Social anxiety: SIAS - shortened version.	The hypothesised moderator of parental acceptance in the association between GN & SSA on psychological distress was investigated using multiple regression analyses for males and females.	Parental acceptance moderated the relationship between GNC and psychological distress.GN females with high levels of maternal acceptance did not experience significantly higher psychological distress in comparison to GC females. Similarly, GN males with high levels of paternal acceptance did not experience significantly higher psychological distress in comparison to GC males.	psychological distress and social anxiety

11	Li, Pollitt, Russell, (2016).	US National Longitudinal Study of Adolescent Health (N = 9421). Add Health Study Waves 3 & 4.	Young adult GN: New measure created by authors, where GN is determined by the individual's activities. Depression: Short version of Center for Epidemiological Studies Depression Scale (CES-D). Sexual orientation: Kinsey-type scale of 100% heterosexual to 100% homosexual.	A comparison of concurrent (18 – 26 years) and prospective (24 – 32 years) depression was performed between GN and GC heterosexual and sexual minority sub-groups where GN/GC and SO were measured during young adulthood (18 – 26 years)	Young adult GN was associated with higher concurrent depression, regardless of SO. Also associated with prospective depression in heterosexual males.	depression
12	Toomey, Ryan, Diaz, Card, Russell, (2013).	A US cross sectional study of 245 LGBT young adults ranging in age from 21 to 25 years.	Adolescent and Young Adult GN: 1 item, consisting of a feminine - masculine scale. Past school LGBT related victimization: Authors used a 10-item scale, including questions on physical violence. Depression: CESD-20 scale. Life satisfaction: Authors used an 8-item scale.	Structural Equation Modelling (SEM) tested the model that depression and life satisfaction in GN LGBT adolescents are mediated by school victimization.	LGBT GN participants reported higher rates of depression and reduced life satisfaction in comparison to the LGBT GC participants. This association was mediated by school victimization.	life satisfaction and depression

| 13 | Sandfort, Melendez, Diaz (2007). | A US cross sectional study of 912 Latino gay and bisexual (GB) men in three major U.S. cities. The data were analysed from face-to-face interviews. | All measures were created, based on a large qualitative study of 300 gay and bisexual men. Adult GN: 1 item, "Do you consider yourself to be effeminate?"(Y/N/don't know response). Childhood Sexual Abuse (CSA): 2 items on underage sex with someone at least 5 years older, against their will. Abuse by a relative / romantic partner: 2 items on emotional and physical abuse. Rape or sexual abuse by a relative: 1 item. Experiences of homophobia: 11 item questionnaire re: SMS and GN. Social support: Adapted UCLA Loneliness Scale 7-item scale. Mental distress: 5-item measure covering anxiety, depression and suicidal feelings over the previous 6 months. | Data were analysed in four steps as per mediation analysis guidance of Baron & Kenny (1986). Step 1: Multiple linear regression tested association between GN and mental distress, controlling for location, degree of acculturation and age. Step 2: Regression analysis to test association between GN and mediators - abuse / rape, social support, homophobic experiences – and mental distress. Step 3: Tested the role of each mediator on GN. Step 4: Included all mediators that reduced GN coefficient by 10% or more. | In comparison to adult GC gay and bisexual (GB) Latino men, GN GB Latino men reported higher levels of mental distress and negative experiences, which were mediated by increased homophobic experiences. | mental distress |

Note. The full details of each measure listed in the table are:

Recalled Childhood Gender Identity / Gender Role Questionnaire (RCGIGR, Zucker et al., 2006); Childhood Gender Nonconformity Scale (CGNS; Bailey & Oberschneider, 1997); Childhood Gender Identity Scale (CGIS; Johnson et al., 2004); Depression Symptoms Scale of the McKnight Risk Factor Survey (Shisslack et al., 1999); Center for Epidemiological Studies Depression Scale, shortened version (CESD-10; Radloff, 1977); Brief Symptom Inventory-18

(BSI-18; Derogatis); Centre for Epidemiologic Studies Depression Scale (CESD; Radloff, 1997); Beck Depression Inventory II (BDI-II; Beck & Steer, 1987); State Trait Anxiety Inventory (STAI: Spielberger, 1983); WHO Health Behaviour of School aged children survey, Child Trauma Questionnaire (CTQ; Bernstein et al., 2003) & Conflict Tactics Scales (CTS; Straus, Hamby, Boney-McCoy, & Sugarman, 1996); Sell Assessment of Sexual Orientation (Sell, 1996); Measure of Parenting Style (Parker et al., 1997); Heterosexist Harassment, Rejection and Discrimination Scale (HHRDS; Szymanski, 2006) (LGB-inclusive version); Internalised homonegativity subscale from the Lesbian, Gay and Bisexual Identity Scale (LSBIS; Mohr & Fassinger 2003); Gay-Related Rejection Sensitivity Scale (Pachankis, Goldfried & Ramrattan, 2008); Sexual Orientation: Kinsey scale (Kinsey, Pomeroy & Martin, 1948); Adapted Continuous Gender Identity Scale (CGIS; Bailey, Finkel, Blackwelder & Bailey, 1998), adapted version; Adapted Childhood Gender Nonconformity Scale (CGNS; Collier, Bos & Sandfort, 2013); Personal and Social Wellbeing Module; Brief Symptom Inventory (BSI; Derogatis, 2001); Experienced Rejection Scale (Bos, Van Balen, Van den Boom & Sandfort, 2004); Homophobic Content Target Subscale (Poteat & Espelage, 2005) as translated and modified by Collier et al., (2013); UCLA Loneliness Scale (Russell, 1996); Parent and Peer Attachment (Armsden & Greenberg, 1987).

Discussion

General findings.

This review has considered the evidence for associations between gender nonconformity and depression, in addition to the role of potential mediators and moderators of this association. Papers have been categorised into those that studied CGN versus those focusing on adolescent or adult GN. Of the former, only Roberts et al., (2013) utilised a longitudinal study design. As the others were cross sectional, it is not possible to infer causation from their results.

As already discussed, Roberts et al., (2013) found significant associations between CGN and depression which were mediated by abuse and bullying victimisation. Their findings applied to male and female participants, although they were stronger in males. This included higher rates of bullying victimisation in CGN boys in comparison to CGN girls. Considering the results of this study, it would be reasonable to assume that the increased likelihood of depression is related to their experiences of bullying, violence and harassment.

Their study had a robust and detailed methodology and while acknowledging some of the limitations such as self-report, retrospective accounts of CGN, childhood abuse and bullying by adult nonfamily members, they also acknowledged study strengths (EPHPP, 2009). These included the regular assessment of bullying and levels of depression across adolescence and early adulthood. This allowed them to assess experiences of bullying and depression levels over time. Additionally, the large sample size permitted the evaluation of the relationship between CGN and depression by biological sex and sexual orientation.

The other studies in this category reported varied findings. For instance, Skidmore's findings of a positive association between CGN and psychological distress in gay men, but not in lesbian participants, may be related to the fact that CGN is more socially acceptable in females, in comparison to males. This may lead to different treatment of CGN girls and boys, potentially resulting in differing levels of distress (Maccoby, 1998). Evidence exists of CGN boys being rejected by their peers and this may lead to distress. This rejection is not always evident in CGN girls (Friedman & Downey, 1999).

Petterson et al., (2016) reported similar findings, in that CGN was associated with increased levels of depression in sexual minority and heterosexual males, but not in females. They also proposed that these differences were related to relationship difficulties between CGN boys and their peers, parents and other adults. Increased levels of bullying in CGN boys are associated with higher rates of depression (e.g. Young & Sweeting, 2004). They may also experience parental over-protectiveness and this is associated with higher levels of depression (e.g. Alanko et al., 2009).

There were no gender differences reported in the findings of Alanko et al., (2009) and it could be argued that although CGN boys may experience more discrimination in comparison to CGN girls, their negative perception of their CGN is similar and therefore leads to similar levels of distress. It may also be the case that the gender differences in maltreatment of CGN individuals may not be as evident in Finnish society as it is in other cultures (EPHPP, 2009). Similarly, Feinstein et al., (2012) found significant associations between CGN and depression in gay men and lesbian women, although the findings were stronger for gay men. A similar rationale to those posited by Skidmore et al., (2006) and Petterson et al., (2016) was provided in order to explain the stronger association in males.

Alanko et al., (2009) found that CGN individuals recalled poorer quality parental relationships in comparison to their CGC counterparts. However, it was noted that the effects of parenting on distress were the same for CGN and CGC participants. Feinstein (2012) noted a good model fit with CGN being significantly associated with discrimination, which in turn was significantly associated with internalised homonegativity and depression.

The findings of these CGN studies support gender schema theory (Trew & Kramer, 1998), in which children develop masculinity and femininity schemata. Children who do not conform to these schemata may experience rejection, ridicule and stigmatisation from other children. Social identity theory posits the psychological mechanisms of social categorisation, identification and comparison (Tajfel & Turner, 1979). It could be argued that the bigger and more socially acceptable group of GC individuals rejects the smaller group of GN individuals and considers them as inferior and treating them accordingly.

In terms of adolescent and adult GN (AAGN), Li et al., (2016) found that young adult GN was associated with increased rates of depression in comparison to GC individuals. However this association was not apparent when measuring prospective depression six years later, indicating that the effect of adult GN on depression declines over time. The remaining studies have a cross sectional design.

Cook et al., (2013) sought to explain their surprising findings that although gay and bisexual GN males experienced more discrimination in comparison to their GC counterparts, they did not experience increased rates of depression. This finding may be explained by cultural differences between South African gay men and those living in more westernised, high-income countries (EPHPP, 2009). Although the potentially protective factors of outness and gay community involvement were not found to mediate the association between GN and depression, other unidentified confounding variables may have masked the difference, such as

higher levels of resilience in GN males, perhaps developed as a method of coping with increased abuse and victimisation. They may also have higher levels of self-esteem perhaps as a result of being protected by family and friends, and being permitted to participate in traditionally female activities (Rabie & Lesch, 2009). This elevated self-esteem may counteract the discrimination. Study findings would have been more robust if an established, validated scale had been used to measure GNC, as opposed to a single question on levels of masculinity or femininity.

Baams et al., (2013) found that higher levels of early adult GNC resulted in reduced levels of psychological wellbeing and this association was mediated by perceived stigmatisation. Although no significant differences were found between males and females, males experienced higher rates of stigmatisation and this difference approached significance. Further analysis demonstrated that the GN, as opposed to sexual orientation, predicted the level of stigmatisation.

The findings of van Beusenkom et al., (2016) were that GNC was associated with psychological distress and that this association was partly mediated by homophobic name calling, especially in males and those participants with medium or high levels of SSA.

In another study, van Beusenkom et al., (2015) found that father acceptance resulted in lower levels of psychological distress in GNC males, in comparison to their counterparts who experienced low levels of father acceptance. This may be because males who are highly GNC experience high levels of discrimination and stigmatisation in comparison to sexual minority males who are gender conforming, therefore they have more to gain from a positive, supportive relationship with their father. In contrast, mother acceptance did not moderate the association between GNC females and psychological distress. As levels of stigmatisation and

victimisation are lower in females, as GN is more socially acceptable, they may have less to gain from a supportive maternal relationship.

Toomey et al., (2013) reported that adolescent GN was associated with depression in early adulthood and this relationship was mediated by victimisation as a result of assumed or actual sexual minority status. Although this association was found in males and females, young males experienced significantly higher levels of depression in comparison to their female counterparts. One surprising finding was that young GN males did not experience more school LGBT victimisation than GN females. GN is a multidimensional phenomenon and this study may have been improved by utilising a validated GN scale instead of a single question. It would also have been strengthened by incorporating other experiences of victimisation, outside of the school context (EPHPP, 2009).

Sandfort et al., (2007) found that GN gay and bisexual males experienced higher levels of mental distress in comparison to their GC counterparts. This association was mediated by experiences of homophobic abuse and harassment. Additionally, this study was the first to demonstrate that GN gay and bisexual males experienced increased levels of childhood emotional, physical and sexual abuse in comparison to their GC counterparts. Although most homophobic experiences result from interactions with heterosexual individuals, some also result from interactions with individuals who identify as sexual minority. This negative attitude towards GN in the gay community has been referred to as sissyphobia (e.g. Taywaditep, 2001).

As with the previous study, GN was assessed through a single question, therefore this study could have been strengthened by using a validated questionnaire on GN. The use of the term feminine, as opposed to effeminacy, may have been considered less offensive by many GN males and may have resulted in more open and honest responses (EPHPP, 2009).

As with the CGN studies, the adolescent and adult GN studies lend support to the gender schema theory and the social identity theory, as evidenced by the mechanisms under investigation such as perceived stigmatisation, school victimisation and homophobic name-calling.

Research implications

There were only two longitudinal studies in this review and more are needed to establish casual relationships between the study variables. Additionally, they were based in the US and therefore may have limited generalisability to the UK, highlighting the urgent need for UK based longitudinal studies. Additional UK cross sectional studies, although not as desirable as longitudinal studies, would provide a valuable addition to the evidence base. Furthermore, studies that investigate the mechanisms that seek to explain the relationship between GNC and depression should be prioritised.   Broad, representative study samples, including heterosexuals, ethnic and socioeconomic diversity, will help to make the results more representative of the general population. It would also be interesting and informative to incorporate more cultural diversity and implement more studies in non-westernised societies.

More specifically, all five CGN studies in the review involved retrospective, recalled accounts of CGN, leading to the possibility of recall bias. Therefore, it would be beneficial to implement studies that measure CGN prospectively. Additionally, only one CGN study investigated depression in childhood (Petterson et al., 2016) and this was a retrospective account of childhood depression in adulthood, again leading to the possibility of recall bias. Therefore, there is a pressing need for further studies that investigate CGN and depression in childhood, measured prospectively.

In terms of specific mechanisms, it would be interesting to investigate the mediating role of self-esteem in the association between CGN and depression. In their review, Sowislo and Orth (2013) found evidence to support the vulnerability model, which proposes that low self-esteem results in higher levels of depression. Additionally, the increased negative interpersonal experiences of CGN children, particularly boys may lead to lower self-esteem (e.g. Baams et al., 2013; Meyer, 2003). Indeed lower self-esteem may result from the mere observation of societal attitudes towards GN (e.g. Meyer & Dean, 1998). Furthermore, self-esteem is recognised as a key factor in theoretical models of psychological distress (e.g. Fonagy, Target, Cottrell, Phillips & Kurtz, 2002) and more specifically to depression (e.g. Beck 1976).

Moreover, it would be interesting to explore potential mechanisms that may explain the relationship between self-esteem and depression, such as rumination or reassurance seeking (Kuster, Orth & Meier, 2012; Sowislo & Orth, 2013).

Further studies investigating other mechanisms discussed in this review, such as the quality of the relationship with parents, experiences of bullying, abuse and other victimisation and perceived stigmatisation, would add to the evidence base. Longitudinal studies using a causal model would help to infer causation and add weight to the findings.

Clinical implications

Community psychology and public health initiatives to improve peoples' awareness of GN and of the stigma that GN individuals experience may help to improve societal attitudes. For example, Stonewall produce detailed LGBT anti-bullying resources which refer to gender nonconformity. These are disseminated among primary and high schools and include questionnaires to ascertain the prevalence of homophobic and transphobic bullying, in

addition to advice on management of homophobic bullying. Anti-bullying and discrimination initiatives in schools may also help to improve the quality of life of GN children. Early life initiatives are preferable as they may help to prevent mental health problems in childhood and adulthood.

Additionally, the increased rates of depression and other mental health problems reported in GN individuals should prompt mental health services to identify and monitor the mental health of these individuals and prioritise their treatment, within reason. For example children suffering from depression and /or anxiety could be supported with a suitably modified version of cognitive behavioural therapy (CBT). As parental support is an important mediator, services should provide support to parents of GN children. Moreover, commissioners and policy makers should also be aware of the increased need for mental health support within this population and allocate resources accordingly.

Conclusion

All of the papers in this review have found associations between childhood, adolescent or adult GN and depression (or psychological distress) and some noted this association only in males (e.g. Skidmore et al., 2006). However, Alanko et al., (2009) found an association between CGN and adulthood psychological distress in males and females, which could be partially explained by parental coldness and over-control.

Both studies were cross-sectional and of the thirteen papers reviewed, only two were longitudinal. Roberts et al., (2013) found that the relationship between CGN and depression in adolescence and early adulthood was mediated by abuse and bullying victimisation. Li et al., (2016) found a positive relationship between young adult GN and concurrent depression and later depression in heterosexual GN participants.

The other cross-sectional papers found that the associations between the key study variables were partially mediated by, for example, internalised homophobia (Feinstein et al., 2012), perceived stigmatisation (Baams et al., 2013), homophobic name-calling (van Beusenkom et al., 2016) and parental acceptance (van Beusenkom et al., 2015).

Study limitations include the fact that only two papers had a longitudinal study design, with the others being cross sectional; or the location of the studies, with many being based in the US and the Netherlands. Other limitations relate to the measures utilised, such as GN being measured with one question, rather than a validated questionnaire. The majority of studies had sufficiently large samples to ensure sufficient statistical power and confidence in their findings. However, many only sampled the sexual minority population, thus limiting their generalisability. Future research should include longitudinal studies with large, representative samples. However

Research implications have been discussed, such as the urgent requirement for more UK based studies and more longitudinal studies in general. There was only one study of CGN and childhood depression, highlighting a need for further studies in this area. Additionally, none of the studies reviewed investigated self-esteem as a mediator. Clinical implications include the requirement for community psychology, public health and school interventions to tackle the stigmatisation of GN individuals. Early interventions may help to prevent future mental health problems and parental support are both important.

REFERENCES

Alanko, K., Santtila, P., Witting, K., Varjonen, M., Jern, P., Johansson, A., ... & Kenneth

Sandnabba, N. (2009). Psychiatric symptoms and same-sex sexual attraction and

behavior in light of childhood gender atypical behavior and parental

relationships. *Journal of Sex Research*, *46*, 494-504. doi:

10.1080/00224490902846487.

Armsden, G. C., & Greenberg, M. T. (1987). The inventory of parent and peer attachment:

Individual differences and their relationship to psychological well-being in

adolescence. *Journal of Youth and Adolescence, 16,* 427–454. doi:

10.1007/BF02202939

Aspenlieder, L., Buchanan, C. M., McDougall, P., & Sippola, L. K. (2009). Gender

nonconformity and peer victimization in pre-and early adolescence. *International

Journal of Developmental Science*, *3*, 3-16. doi: 10.3233/DEV-2009-3103

Baams, L., Beek, T., Hille, H., Zevenbergen, F. C., & Bos, H. M. (2013). Gender

nonconformity, perceived stigmatization, and psychological well-being in Dutch

sexual minority youth and young adults: A mediation analysis. *Archives of Sexual

Behavior*, *42*, 765-773. doi: 10.1007/s10508-012-0055-

Bailey, J. M., Finkel, E., Blackwelder, K., & Bailey, T. (1998). *Masculinity, femininity, and

sexual orientation*. Unpublished manuscript, North-western University, Evanston, IL.

Bailey, J. M., & Oberschneider, M. (1997). Sexual orientation and professional dance.

Archives of Sexual Behavior, 26, 433–444. doi:10.1023/A:1024595419744

Bailey, J. M., & Zucker, K. J. (1995). Childhood sex-typed behavior and sexual orientation: A conceptual analysis and quantitative review. *Developmental Psychology*, *31*, 43. doi: 10.1037/0012-1649.31.1.43

Bandura, A. (1977). *Social learning theory*. Eaglewood Cliffs, New Jersey: Prentice Hall.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, *51*, 1173. Retrieved from http://www.mcole-psy.com/mcole/Cole-publications/Baron-Kenny1986_The%20moderator-mediator%20variable%20distinction.pdf

Beck, A.T. (1976). *Cognitive therapy and the emotional disorders*. New York: New American Library.

Beck, A. T., & Steer, R. A. (1987). *Manual for the revised Beck depression inventory*. San Antonio, TX: Psychological Corporation.

Bem, S. L., & Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155-162. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.472.525&rep=rep1&type=pdf

Bem, S. (1985). Androgyny and gender schema theory: A conceptual and empirical integration. In T. N. Sonderegger (Ed.) *Nebraska symposium on motivation 1984: Psychology and gender* (pp. 179-226). Lincoln, NE: University of Nebraska Press.

Bem, S. L. (1993). *The lenses of gender: Transforming the debate on sexual inequality*. Yale University Press.

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., ... & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse & Neglect*, *27*(2), 169-190. doi: 10.1016/S0145-2134(02)00541-0

Blakemore, J. E. O. (2003). Children's beliefs about violating gender norms: Boys shouldn't look like girls, and girls shouldn't act like boys. *Sex roles*, *48*, 411-419. doi: 10.1023/A:1023574427720

Bos, H.M., VanBalen, F., VandenBoom, D. C., & Sandfort, T. G. (2004). Minority stress, experience of parenthood and child adjustment in lesbian families. *Journal of Reproductive and Infant Psychology, 22*, 291–304. doi:10.1080/02646830412331298350.

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity.* New York: Routledge.

Clark, T. C., Lucassen, M. F., Bullen, P., Denny, S. J., Fleming, T. M., Robinson, E. M., & Rossen, F. V. (2014). The health and well-being of transgender high school students: results from the New Zealand adolescent health survey (Youth'12). *Journal of Adolescent Health*, *55*(1), 93-99. doi: 10.1016/j.jadohealth.2013.11.008

Cohen-Kettenis, P. T., Owen, A., Kaijser, V. G., Bradley, S. J., & Zucker, K. J. (2003). Demographic characteristics, social competence, and behavior problems in children with gender identity disorder: A cross-national, cross-clinic comparative analysis. *Journal of Abnormal Child Psychology*, *31*(1), 41-53. doi: 10.1023/A:1021769215342

Collaer, M. L., & Hines, M. (1995). Human behavioral sex differences: a role for gonadal

hormones during early development? *Psychological Bulletin*, *118*(1), 55. doi:

10.1037/0033-2909.118.1.55

Collier, K. L., Bos, H. M., & Sandfort, T. G. (2013). Homophobic name-calling among

secondary school students and its implications for mental health. *Journal of Youth and

Adolescence*, *42*, 363-375. doi: 10.1007=s10964-012-9823-2

Conron, K. J., Scott, G., Stowell, G. S., & Landers, S. J. (2012). Transgender health in

Massachusetts: results from a household probability sample of adults. *American

Journal of Public Health*, *102*(1), 118-122. doi: 10.2105/AJPH.2011.300315

Cook, S. H., Sandfort, T. G., Nel, J. A., & Rich, E. P. (2013). Exploring the relationship

between gender nonconformity and mental health among black South African gay and

bisexual men. *Archives of Sexual Behavior*, *42*, 327. doi: 10.1007/s10508-013-0087-z

D'Augelli, A. R., Grossman, A. H., & Starks, M. T. (2006). Childhood gender atypicality,

victimization, and PTSD among lesbian, gay, and bisexual youth. *Journal of

Interpersonal Violence*, *21*, 1462-1482. doi: 10.1177/0886260506293482

Dandona, L., Dandona, R., Gutierrez, J. P., Kumar, G. A., McPherson, S., Bertozzi, S. M., &

Asci FPP Study Team. (2005). Sex behaviour of men who have sex with men and risk

of HIV in Andhra Pradesh, India. *Aids*, *19*(6), 611-619.

doi: 10.1097/01.aids.0000163938.01188.e4

Diaz, R. M., Ayala, G., Bein, E., Henne, J., & Marin, B. V. (2001). The impact of

homophobia, poverty, and racism on the mental health of gay and bisexual Latino

men: findings from 3 US cities. *American Journal of Public Health*, *91*, 927. doi:

   10.2105/AJPH.91.6.927

Derogatis, L. R. (2001). Brief Symptom Inventory–18 (BSI–18); Administration, scoring, and

   procedures manual. Minneapolis: NCS Pearson. doi:10.1037/t07502-000

Effective Public Health Practice Project (EPHPP). (2009). Quality Assessment Tool for

   Quantitative Studies. Retrieved March 17, 2016 from

   http://www.ephpp.ca/tools.html

Feinstein, B. A., Goldfried, M. R., & Davila, J. (2012). The relationship between experiences

   of discrimination and mental health among lesbians and gay men: An examination of

   internalized homonegativity and rejection sensitivity as potential

   mechanisms. *Journal of Consulting and Clinical Psychology*, *80*, 917. doi:

   10.1037/a0029425

Fonagy, P., Target, M., Cottrell, D., Phillips, J., & Kurtz, Z. (2002). What works for whom. *A

   Critical Review of Treatments for Children and Adolescents*, 1-41.

   doi:10.1080/07317107.2016.1203157

Foushee, H. C., Helmreich, R. L., & Spence, J. T. (1979). Implicit theories of masculinity

   and femininity: Dualistic or bipolar? *Psychology of Women Quarterly, 3*, 259-269.

   doi: 10.1111/j.1471-6402.1979.tb00544.x

Friedman, R., & Downey, J. (2002). *Sexual orientation and psychodynamic psychotherapy:

   Sexual science and clinical practice*. Columbia University Press.

Harry, J. (1983). Defeminization and adult psychological well-being among male

   homosexuals. *Archives of Sexual Behavior, 12*(1), 1-19. doi: 10.1007/BF01542112

Hatzenbuehler, M. L. (2009). How does sexual minority stigma "get under the skin"? A

psychological mediation framework. *Psychological Bulletin*, *135*, 707. doi:

10.1037/a0016441

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable

mediation, moderation, and conditional process modeling [White paper]. Retrieved

from http://www.afhayes.com/ public/process2012.pdf

Herek, G. M., & Berrill, K. T. (1990). Documenting the victimization of lesbians and gay

men: Methodological issues. *Journal of Interpersonal Violence*, *5*(3), 301-315.

doi: 10.1177/088626090005003005

Huppert, F. A., Marks, N., Clark, A.E., Siegrist, J., Stutzer, A., Vitterso, J., et al. (2009).

Measuring well-being across Europe: Description of the ESS well-being module and

preliminary findings. Social Indicators Research, 91, 301–315. doi:10.1007/s11205-

008-9346-0.

Johnson, L. L., Bradley, S. J., Birkenfeld-Adams, A. S., Kuksis, M. A. R., Maing, D. M.,

Mitchell, J. N., & Zucker, K. J. (2004). A parent-report gender identity questionnaire

for children. *Archives of Sexual Behavior*, *33*(2), 105-116.

doi: 10.1177/1359104506059135

Johnson, K. L., & Ghavami, N. (2011). At the crossroads of conspicuous and concealable:

What race categories communicate about sexual orientation. *PLoS One*, *6*(3). doi:

10.1371/journal.pone.0018025.

Kerr, D. L., Santurri, L., & Peters, P. (2013). A comparison of lesbian, bisexual, and

heterosexual college undergraduate women on selected mental health issues. *Journal

of American College Health*, *61*(4), 185-194. doi: 10.1080/07448481.2013.787619

Kinsey, A. C., Pomeroy, W. B., Martin, C. E., & Sloan, S. (1948). *Sexual behavior in the

human male.* Indiana University Press

Kuster, F., Orth, U., & Meier, L. L. (2012). Rumination mediates the prospective effect of

low self-esteem on depression: A five-wave longitudinal study. *Personality and

Social Psychology Bulletin*, *38*(6), 747-759. doi: 10.1177/0146167212437250

Lee, E. A. E., & Troop-Gordon, W. (2011). Peer processes and gender role development:

Changes in gender atypicality related to negative peer treatment and children's

friendships. *Sex Roles*, *64*(1-2), 90-102. doi: 10.1007/s11199-010-9883-2

Li, G., Pollitt, A. M., & Russell, S. T. (2016). Depression and sexual orientation during

young adulthood: Diversity among sexual minority subgroups and the role of gender

nonconformity. *Archives of Sexual Behavior*, *45*, 697-711. doi: 10.1007/s10508-015-

0515-3

Lippa, R. A. (2000). Gender-related traits in gay men, lesbian women, and heterosexual

men and women: The virtual identity of homosexual-heterosexual diagnosticity and

gender diagnosticity. *Journal of Personality*, *68*, 899-926. doi: 10.1111/1467-

6494.00120

Lippa, R. A. (2002). Gender-related traits of heterosexual and homosexual men and

women. *Archives of Sexual Behavior*, *31*, 83-98. doi: 10.1023/A:1014035302843

Lippa, R. A. (2008). The relation between childhood gender nonconformity and adult masculinity–femininity and anxiety in heterosexual and homosexual men and women. *Sex Roles*, *59*, 684-693. doi: 10.1007/s11199-008-9476-5

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*, 335-343. doi: 10.1016/0005-7967(94)00075-u

Maccoby, E. E. (1998). *The two sexes: Growing up apart, coming together* (Vol. 4). Harvard University Press.

Martin-Storey, A., & August, E. G. (2016). Harassment due to gender nonconformity mediates the association between sexual minority identity and depressive symptoms. *The Journal of Sex Research*, *53*, 85-97. doi: 10.1080/00224499.2014.980497

Menon, M. (2011). Does felt gender compatibility mediate influences of self-perceived gender nonconformity on early adolescents' psychosocial adjustment? *Child Development*, *82*, 1152-1162. doi: 10.1111/j.1467-8624.2011.01601.x

Meyer, I. H. (1995). Minority stress and mental health in gay men. *Journal of Health and Social Behavior*, 38-56. doi: 10.1037/0033-2909.129.5.674

Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. *Psychological Bulletin*, *129*, 674. doi: 10.1037/0033-2909.129.5.674

Meyer, I. H., & Dean, L. (1998). Internalized homophobia, intimacy, and sexual behavior among gay and bisexual men. *Psychological Perspectives on Lesbian and Gay issues*, *4*, 160-186. doi: 10.4135/9781452243818.n8

Mischel, W. (1970). Sex-typing and socialization. In P. H. Mussen (Ed.), *Manual of Child Psychology* (pp. 3-72) New York: Wiley.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D.G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine, 6.* doi:10.1371/journal.pmed1000097

Mohr, J., & Fassinger, R. (2000). Measuring dimensions of lesbian and gay male experience. *Measurement and Evaluation in Counseling and Development*, *33*(2), 66-66.

Mohr, J. J., & Fassinger, R. E. (2003*). Information sheet: Lesbian, Gay, and Bisexual Identity Scale*. Unpublished manuscript.

Pachankis, J. E., Goldfried, M. R., & Ramrattan, M. E. (2008). Extension of the rejection sensitivity construct to the interpersonal functioning of gay men. *Journal of Consulting and Clinical Psychology, 76,* 306–317. doi:10.1037/0022-006X.76.2.306

Parker, G., Roussos, J., Hadzi-Pavlovic, D., Mitchell, P., Wilhelm, K., & Austin, M.-P. (1997). The development of a refined measure of dysfunctional parenting and assessment of its relevance in patients with affective disorders. Psychological Medicine, 27, 1193–1203. doi:10.1017/s003329179700545x

Pascoe, C. J. (2007). Dude, you're a fag. Berkeley. doi: 10.1177/0092055x14533906

Petterson, L. J., VanderLaan, D. P., & Vasey, P. L. (2016). Sex, Sexual Orientation, Gender Atypicality, and Indicators of Depression and Anxiety in Childhood and Adulthood. *Archives of Sexual Behavior*, *45*, 1-10. doi: 10.1007/s10508-016-0690-x

Pilkington, N. W., & D'Augelli, A. R. (1995). Victimization of lesbian, gay, and bisexual

youth in community settings. *Journal of Community Psychology, 23*(1), 34-56. doi:

10.1002/1520-6629(199501)

Plummer, D. C. (2001). The quest for modern manhood: Masculine stereotypes, peer culture

and the social significance of homophobia. *Journal of Adolescence*, *24*, 15-23. doi:

10.1006/jado.2000.0370

Poteat, V. P., & Espelage, D. L. (2005). Exploring the relation between bullying and

homophobic verbal content: The Homophobic Content Agent Target (HCAT) Scale.

Violence and Victims, 20, 513–528. doi: 10.1891=0886-6708.2005.20.5.513

Rabie, F., & Lesch, E. (2009). 'I am like a woman': constructions of sexuality among gay

men in a low-income South African community. *Culture, Health & Sexuality*, *11*(7),

717-729. doi: 10.1080/13691050902890344

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the

general population. *Applied Psychological Measurement*, *1*(3), 385-401. doi:

10.1177/014662167700100306

Renold, E. (2004). 'Other' boys: Negotiating non-hegemonic masculinities in the primary

school. *Gender and Education, 16*, 247-265. doi: 10.1080/09540250310001690609

Riggle, E. D., Rostosky, S. S., & Reedy, C. S. (2005). Online surveys for BGLT research:

Issues and techniques. *Journal of Homosexuality*, *49*(2), 1-21. doi:

10.1300/J082v49n02_01

Roberts, A. L., Rosario, M., Slopen, N., Calzo, J. P., & Austin, S. B. (2013). Childhood

gender nonconformity, bullying victimization, and depressive symptoms across

adolescence and early adulthood: An 11-year longitudinal study. *Journal of the*

American Academy of Child & Adolescent Psychiatry, *52*, 143-152. doi:
10.1016/j.jaac.2012.11.006

Rieger, G., Linsenmeier, J. A., Gygax, L., & Bailey, J. M. (2008). Sexual orientation and
childhood gender nonconformity: evidence from home videos. *Developmental
Psychology*, *44*, 46 - 58. doi: 10.1037/0012-1649.44.1.46

Rieger, G., & Savin-Williams, R. C. (2012). Gender nonconformity, sexual orientation, and
psychological well-being. *Archives of Sexual Behavior*, *41*, 611-621. doi:
10.1007/s10508-011-9738-0

Rosario, M., Reisner, S. L., Corliss, H. L., Wypij, D., Frazier, A. L., & Austin, S. B. (2014).
Disparities in depressive distress by sexual orientation in emerging adults: The roles
of attachment and stress paradigms. *Archives of Sexual Behavior*, *43*, 901-916. doi:
10.1007/s10508-013-0129-6

Russell, D. W. (1996). The UCLA Loneliness Scale (Version 3): Reliability, validity, and
factor structure. *Journal of Personality Assessment, 66*, 20–40. doi:
10.1207/s15326934crj1701_4

Sandfort, T., Bos, H., Knox, J., & Reddy, V. (2016). Gender nonconformity, discrimination,
and mental health among Black South African men who have sex with men: A further
exploration of unexpected findings. *Archives of Sexual Behavior*, *45*(3), 661-670.
doi: 10.1007/s10508-015-0565-6

Sandfort, T. G., Melendez, R. M., & Diaz, R. M. (2007). Gender nonconformity,
homophobia, and mental distress in Latino gay and bisexual men. *Journal of Sex
Research*, *44*, 181-189. doi: 10.1080/00224490701263819

Shisslak, C. M., Renger, R., Sharpe, T., Crago, M., McKnight, K. M., Gray, N., ... & Taylor, C. B. (1999). Development and evaluation of the McKnight Risk Factor Survey for assessing potential risk and protective factors for disordered eating in preadolescent and adolescent girls. *International Journal of Eating Disorders*, *25*(2), 195-214. doi: 10.1002/(SICI)1098-108X(199903)25:2<195::AID-EAT9>3.0.CO;2-B

Skidmore, W. C., Linsenmeier, J. A., & Bailey, J. M. (2006). Gender nonconformity and psychological distress in lesbians and gay men. *Archives of Sexual Behavior*, *35*, 685-697. doi: 10.1007/s10508-006-9108-5

Skinner, B. F. (1953). *Science and human behavior*. Simon and Schuster.

Smith, T. E., & Leaper, C. (2006). Self-perceived gender typicality and the peer context during adolescence. *Journal of Research on Adolescence*, *16*(1), 91-104. doi: 10.1111/j.1532-7795.2006.00123.

Sowislo, J.F. & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. Psychological Bulletin, 139 (1), 213 – 240. DOI: 10.1037/a0028931.

Spence, J. T., & Helmreich, R. L. (1978). *Masculinity & femininity: Their psychological dimensions, correlates, and antecedents*. Austin: University of Texas Press. doi: 10.1017/s0033291700053782

Spielberger, C. D. (1983). Manual for the State-Trait Anxiety Inventory STAI (form Y)(" self-evaluation questionnaire"). doi: 10.1037/t06496-000

Steensma, T. D., van der Ende, J., Verhulst, F. C., & Cohen-Kettenis, P. T. (2013). Gender

    variance in childhood and sexual orientation in adulthood: A prospective study. *The*

    Journal of Sexual Medicine, *10*, 2723-2733. doi: 10.1111/j.1743-6109.2012.02701.x

Straus, M. A., Hamby, S. L., Boney-McCoy, S., & Sugarman, D. B. (1996). The revised

    conflict tactics scales (CTS2) development and preliminary psychometric

    data. *Journal of Family Issues*, *17*(3), 283-316. doi: 10.1177/019251396017003001

Szymanski, D. M. (2006). Does internalized heterosexism moderate the link between

    heterosexist events and lesbians' psychological distress? *Sex Roles, 54,* 227–234.

    doi:10.1007/s11199-006-9340-4

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology,*

    33(1), 1-39. doi: 10.1146/annurev.ps.33.020182.000245

Tajfel, H., & Turner, J. C. (2004). The social identity theory of intergroup behavior. In J. T.

    Jost and J. Sidanius (Eds.) *Political psychology: Key readings* (pp. 276-293). New

    York: Psychology Press.

Taywaditep, K. J. (2001). Marginalization among the marginalized. *Journal of*

    Homosexuality, 42(1) 1-28. doi: 10.1300/J082v42n01_01

Toomey, R. B., Ryan, C., Diaz, R. M., Card, N. A., & Russell, S. T. (2013). Gender-

    nonconforming lesbian, gay, bisexual, and transgender youth: school victimization

    and young adult psychosocial adjustment. *Developmental Psychology*, *46*, 1580. doi:

    10.1037/a0020705

Trew, K., & Kremer, J. (1998). *Gender & Psychology.* London: Arnold Publishers.

Valentova, J., Rieger, G., Havlicek, J., Linsenmeier, J. A., & Bailey, J. M. (2011). Judgments

    of sexual orientation and masculinity–femininity based on thin slices of behavior: A

cross-cultural comparison. *Archives of Sexual Behavior*, *40*(6), 1145-1152.

doi: 10.1007/s10508-011-9818-1

Vanable, P. A., McKirnan, D. J., & Stokes, J. P. (1998). Identification and involvement with the gay community scale. *Handbook of sexuality-related measures*, 407-409.

van Beusekom, G., Bos, H. M., Overbeek, G., & Sandfort, T. G. (2015). Same-sex attraction, gender nonconformity, and mental health: The protective role of parental acceptance. *Psychology of Sexual Orientation and Gender Diversity*, *2*(3), 307. doi: 10.1037/sgd0000118

van Beusekom, G., Baams, L., Bos, H. M., Overbeek, G., & Sandfort, T. G. (2016). Gender nonconformity, homophobic peer victimization, and mental health: How same-sex attraction and biological sex matter. *The Journal of Sex Research*, *53*, 98-108. doi: 10.1080/00224499.2014.993462

Whitam, F. L. (1977). Childhood indicators of male homosexuality. *Archives of Sexual Behavior, 6*(2), 89-96. doi: 10.1007/BF01541701

Young, R., & Sweeting, H. (2004). Adolescent bullying, relationships, psychological well-being, and gender-atypical behavior: A gender diagnosticity approach. *Sex Roles*, *50*(7-8), 525-537. doi: 10.1023/B:SERS.0000023072.53886.86

Zucker, K. J., & Bradley, S. J. (1995). *Gender identity disorder and psychosexual problems in children and adolescents*. Guilford Press.

Zucker, K. J., Mitchell, J. N., Bradley, S. J., Tkachuk, J., Cantor, J. M., & Allin, S. M. (2006). The recalled childhood gender identity/gender role questionnaire:

Psychometric properties. *Sex Roles*, *54*(7-8), 469-483. doi: 10.1007/s11199-006-9019-

x

MAJOR RESEARCH PROJECT: SECTION B

A study of the association between gender nonconformity and depression: Does

self-esteem mediate the association?

For submission to the journal Archives of Sexual Behaviour

Accurate Word Count: 7139

Abstract

Several studies have reported associations between childhood gender nonconformity (CGN) and depression or other mental health problems and potential mechanisms that seek to explain this association. However, methodological issues limit the power of, and therefore the confidence in, their findings. Additionally, only one previous study has investigated childhood depression as an outcome variable. The current study accesses longitudinal data to investigate whether self-esteem mediates the association between early childhood gender nonconformity and childhood depression. The dataset was obtained from the Avon Longitudinal Study of Parents and Children and consisted of 15,247 records. A mediation analysis (Hayes, 2012) showed that self-esteem mediated the association between CGN and childhood depressive symptoms in females. However, the effect size was very small. In males the mediation model demonstrated an association between CGN and depressive symptoms which approached significance, but this was not mediated by self-esteem. Again, the effect size was very small. Study limitations were discussed, such as high levels of missing data. Research implications include the need for further UK based, preferably longitudinal studies exploring the mechanisms which explain the association between CGN and childhood depression. This should help to develop effective, tailored clinical interventions for this client group. Additionally, community psychology and public health interventions have been discussed.

Introduction

Childhood Gender Nonconformity (CGN)

From early childhood, individuals are expected to conform to societal norms (Gecas, 1992). A notable example is gender, and gender nonconformity can be defined as being more like the opposite sex in terms of behaviour, interests, way of speaking, gestures and play preferences (Roberts, Rosario, Corliss, Koenen & Austin, 2012). This may elicit negative reactions and victimisation from parents, school teachers, other authority figures and peers (e.g. Roberts, Rosario, Slopen, Calzo & Austin, 2013). Such responses are likely to influence the development of identity and self-esteem (e.g. Baams, Beek, Hille, Zevenbergen & Bos, 2013; Meyer, 2003; Roberts et al., 2013). Additionally, gender non-conforming children will experience wider negative societal views and stigmatisation processes which are also likely to affect their self-image (e.g. Meyer & Dean, 1998; Baams et al, 2013).

Childhood depression

Before puberty, approximately one to three percent of children experience depression and it is one of the most common mental health problems affecting children (Costello, Foley & Angold, 2006). Hence it is a very important area to investigate. At this life stage rates of depression are similar in both sexes. After puberty, depression is more common in females than in males and this trend continues into adulthood (Mills & Baker, 2016). Depression is also associated with low self-esteem (Mills & Baker, 2016).

Although effect sizes are modest, research consistently demonstrates that self-esteem is lower in females than in males. This is evident in studies of children, adolescents and adults (Kling, Shibley Hyde, Showers & Buswell, 1999). In their comprehensive review of

longitudinal childhood, adolescent and adult studies, Sowislo and Orth (2013) found support

for the vulnerability model of depression and self-esteem, which proposes that self-esteem

leads to depression. Kuster, Orth & Meier (2012), found that rumination mediated the

relationship between self-esteem and depression in adolescents and adults. Other

hypothesised mechanisms include reassurance seeking and negative feedback seeking

(Sowislo & Orth, 2013). The evidence for the scar model, which hypothesises that depression

leads to reduced self-esteem was less convincing than the evidence for the vulnerability

model. This may be because the hypothesised mechanisms that seek to explain the

relationship between depression and self-esteem, such as the negative impact of depression

on close relationships or social networks, do not have the same impact as the proposed

mechanisms which attempt to explain the effect of self-esteem on depression (Sowislo &

Orth, 2013).

Several studies have investigated associations between childhood gender

nonconformity (CGN) and mental health problems, including depression. While their

findings vary, many report an association which is only significant in males (e.g. Skidmore,

Linsenmeier & Bailey, 2006; Lippa, 2008; Rieger & Savin-Williams; 2011). However,

numerous methodological issues result in lower levels of confidence in their findings, which

will be discussed later in the introduction. Additionally, there is limited evidence about the

mechanisms which may explain this association, such as low self-esteem.

There are many other potential factors that may affect depression in childhood such as

bullying, emotional, physical or sexual abuse, neglect, separation from carers or social

isolation (e.g. Heim, Newport, Mletzko, Miller, & Nemeroff, 2008; Lindert, von Ehrenstein,

Grashow, Gal, Braehler, & Weisskopf 2014; Pillemer, Suitor, Pardo, & Henderson 2010).

Additionally, many medical illnesses such as hypothyroidism, chronic pain, diabetes, stroke,

cancer and Lyme disease are associated with increased rates of depression (e.g. Li, Fitzgerald & Rodin, 2012; Rustad, Musselman & Nemeroff, 2011; Saravane et al., 2009). Medications such as anticonvulsants, antimigraine drugs and hormonal agents may also exacerbate depression). Moreover, the use of substances such as alcohol, sedatives, opioids, stimulants, hallucinogens and inhalants are also associated with depression (e.g. Boden & Fergusson, 2011).

Childhood Gender Nonconformity (CGN) and depression

In their study of 94 adult gay and lesbian participants, Skidmore et al., (2006) found that higher levels of retrospectively self-reported CGN were associated with increased current psychological distress, a composite indicator of depression and anxiety measures, in gay men but not in lesbian women. In their study of 467 adult gay and lesbian participants, Feinstein, Goldfried and Davila, (2012) found that retrospectively recalled CGN was associated with current levels of depression. In comparison to Skidmore et al., (2006), this study's larger sample increases the power of and confidence in, their findings. However both studies only sample the gay and lesbian population which limits the generalisability of their findings.

Alanko et al., (2009) overcame this issue by sampling 454 gay and heterosexual adult participants. They found that higher levels of retrospectively self-reported CGN were associated with current elevated psychological distress, again a composite measure of depression and anxiety, regardless of biological sex or sexual minority status. None of the above studies investigated childhood depression, whereas Petterson, VanderLaan & Vasey, (2016) found that retrospectively self-reported CGN was associated with increased rates of retrospectively self-reported depression, between the ages of 6 and 12, irrespective of

biological sex or sexual orientation. Their sample consisted of 598 gay and heterosexual participants.

The aforementioned studies are cross sectional, whereas a study by Roberts et al., (2013) used a longitudinal approach. Its aim was to assess the potentially mediating role of bullying victimisation in the relationship between CGN before age 11 and depression across adolescence and early adulthood. However, CGN was measured retrospectively in early adulthood. They found that between 12 and 30 years, participants who were highly gender nonconforming before age 11, recorded significantly higher depression rates in comparison to the gender conforming group. This discrepancy was less among the lesbian group.

Mechanisms

Alanko et al., (2009) investigated the quality of the relationship between parent and child and found that it mediated the association between CGN and psychological distress in adulthood. Similarly, Feinstein et al., (2011) found that internalised homonegativity and rejection sensitivity mediated the association between CGN and depression and anxiety, respectively. Internalised homonegativity involves the internalisation of negative societal attitudes towards homosexual identity, behaviours and relationships. Rejection sensitivity refers to the belief that an individual will experience rejection based on their sexual minority status (Meyer, 2003). Roberts et al., (2013) found that child abuse and bullying victimisation mediated the relationship between CGN and depression. Together they accounted for 50% of the increased depression in the CGN group.

None of the studies have investigated the potentially mediating role of self-esteem. This would be interesting to study as it plays a key role in the theoretical models of the

development of psychological problems, such as in the cognitive behavioural model of depression (e.g. Beck 1976) and psychodynamic approaches to distress (e.g. Fonagy, Target, Cottrell, Phillips & Kurtz, 2002). Self-esteem has been identified as a key predisposing and perpetuating factor in a number of psychological problems such as depression, as mentioned above, and eating disorders (e.g. Fairburn & Harrison 2003; Sowislo & Orth, 2013). Therefore the role of self-esteem in the development of depression associated with CGN needs to be investigated.

Of the papers discussed in this section, only one investigated childhood depression (Petterson et al., 2016) and one investigated depression in adolescence (Roberts et al., 2013). There is clearly a need for further research into the association between CGN and childhood depression. Preferably this would consist of prospective reports of both, as retrospective accounts are open to recall bias.

No other literature exists with respect to CGN and childhood depression, although one cross-sectional study of adolescent GN reported current rates of current psychological distress (van Beusenkom, Baams, Bos, Overbeek & Sandfort, 2016). This consisted of 1026 participants (517 males), mean age 13.4 years. Significant associations between adolescent GN and increased psychological distress, which were higher in boys, were partially mediated by homophobic name calling and were higher in sexual minority participants.

Additionally, several studies report a significant association between childhood gender dysphoria and childhood emotional and behavioural problems, (e.g. Steensma et al., 2104; Wallien, Goozen & Cohen-Kettenis, 2007; Zucker & Bradley, 1995; Zucker, Owen, Bradley & Ameeriar, 2002). These are worthy of mention as those children who experience gender dysphoria may be considered to be at the extreme end of the gender nonconformity

spectrum. However, these studies focus on behavioural outcomes (e.g. Zucker et al., 2002), psychosexual outcomes (e.g. Zucker & Bradley, 1995) and anxiety (Wallien et al., 2007). Gender dysphoria refers to the distress and discomfort that an individual feels regarding their assigned gender, which does not match their gender identity (e.g. Zucker & Bradley, 1995)

Steensma et al. (2014) studied the rates of internalising (e.g. anxiety and depression) and externalising (behavioural) problems in GD children. However, they analyse the differences between GD children and adolescents in addition to differences between GD children in Toronto and Amsterdam. Although they report percentage differences between internalising problems in GD children and the norm population of children (22% in GD children v 9% in the norm population), they do not statistically analyse or discuss this difference.

In their study of GD children and adolescents, Di Ceglie, Freedman, McPherson and Richardson (2002) found that rates of depression were reported in 26% and 52% of the child and adolescent participants respectively. Although these percentages are higher than the population norms for children and adolescents (e.g. Mills & Baker, 2016), a statistical analysis was not performed and therefore it is not clear if the differences are statistically significant.

Holt, Skagerberg and Durnsford (2016) found that 7.3% of GD children reported low mood and depression, whereas 47.1% of adolescents reported the same. Again, these percentages were not compared to population norms and it was therefore not possible to determine whether the differences were statistically significant.

Methodological issues with existing papers

It is clear that the findings from these studies vary a great deal, highlighting the need for further studies. There are some methodological concerns such as samples which are limited to the gay and lesbian population and are therefore not generalisable to the overall population (Skidmore et al., 2006; Feinstein et al., 2012) or smaller sample sizes, resulting in limited statistical power of the findings, which reduces confidence in the results (Skidmore et al., 2006). With the exception of Roberts et al., (2013) all studies are cross sectional and it is therefore not possible to infer causality from their findings. Additionally, they all depend on retrospective accounts of CGN and people's recollection may be subject to bias. Two studies utilised an outcome of psychological distress which combined both anxiety and depression so it is not clear whether their findings would be replicated in depression alone (Alanko et al., 2009; Skidmore et al., 2006).

Furthermore, four of the studies were based in North America and one in Finland, highlighting the need for studies to be implemented in other countries, including the UK, to investigate whether findings are replicated in different sociocultural settings. More longitudinal studies which permit the measurement of indicators throughout the individual's life and limit the need for retrospective accounts of key study variables such as CGN, would allow researchers to test temporal relationships between CGN and subsequent depression, as well as potential mediators. This highlights the need for further studies in CGN and childhood depression.

The current study

In consideration of the above issues, it is clear that there is a need for further research into the associations between CGN and childhood depression. In the current study, the hypothesised association between childhood gender nonconformity (CGN) at 30 months and depressive symptomatology at 10 years will be investigated by considering the mediator variable of low self-esteem at 8 years of age.

The measure of CGN at 30 months has been selected as this was the earliest age at which CGN was measured, thereby allowing the study to minimise the influence of other people's potentially negative attitudes towards gender nonconformity on the individual's behaviour. Self-esteem at eight years old was the only measure available in the dataset. Although several measures of depressive symptomatology were available, the measure at age ten was selected as this was the closest measure to self-esteem at age eight.

Previous studies (e.g. Roberts et al., 2013) have utilised retrospective designs. In the current study, gender nonconformity, self-esteem, and depressive symptomatology have been measured prospectively at different points in time. Understanding the mechanisms that may explain the association between CGN and childhood depressive symptomatology could help inform prevention and treatment interventions.

Study hypothesis

With this in mind the current study hypothesis is that early childhood gender nonconformity (CGN) results in an increased likelihood of developing depressive symptomatology and that this relationship is mediated by low self-esteem. The mediation model is provided below.
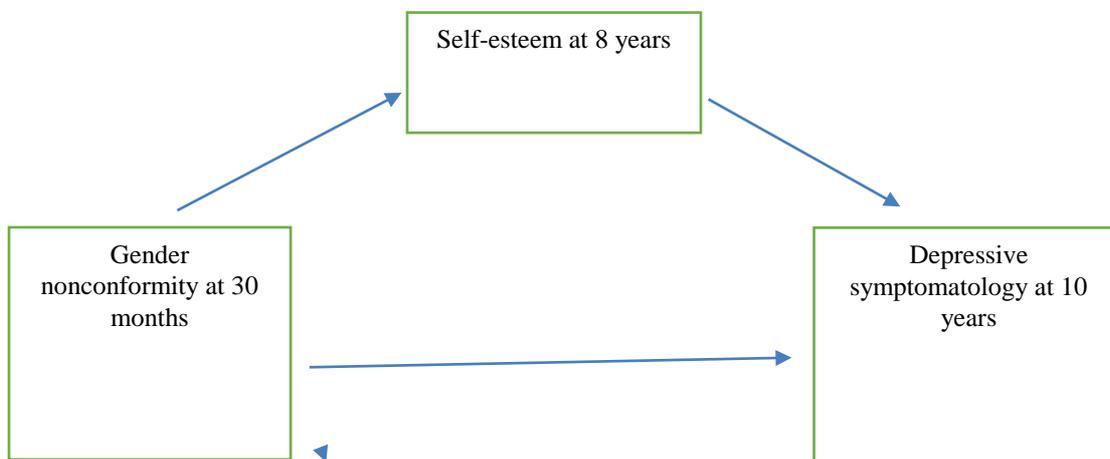
Figure 1. Hypothesised mediation model for gender nonconformity, self-esteem and Depressive symptomatology.

The model will be run separately for boys and girls, as they often have different experiences in relation to their CGN, with boys generally having more negative experiences than girls. The selected measure for CGN in childhood, the Preschool Activities Inventory (PSAI; Golombok & Rust, 1993) is standardised around different scores for boys and girls. Additionally, as already discussed, self-esteem is lower in females in comparison to males (Kling et al., 1999).

Method

Design

This is a correlational study which has made use of a longitudinal dataset obtained from the Avon Longitudinal Study of Parents and Children (ALSPAC), a cohort study from birth, spanning different generations which examines the factors that impact an individual's health and development (Golding, 2010). In addition to biological, psychological, social and environmental indicators, the dataset also includes genetic and epigenetic information (Boyd et al., 2012).

Participants

The ALSPAC dataset consists of pregnancies where the estimated delivery date fell between 1st April 1991 and 31st December 1992 and covered the three District Health Authority areas of Bristol & Weston, Frenchay and Southmead. Of the 20248 eligible pregnancies, 16,734 were known to ALSPAC and were asked to participate. Of those, 15717 were interested in being involved and of those 14,541 signed up to ALSPAC.

Two further recruitment phases were implemented at ages 7 & 8, referred to as phases 2 & 3 respectively. Of the 5,707 potential participants who were not involved in ALSPAC, 452 agreed to take part in the project, from phase 2 and 254 from phase 3. This resulted in a total sample of 15,247 participants from a population of 20,248 individuals who were eligible to take part (75.3% of total potential participants).

Materials

The Preschool Activities Inventory

Gender nonconformity is measured using the Preschool Activities Inventory (PSAI; Golombok & Rust, 1993), a continuous measure which assesses gender role play behaviours in children under 5 years old. This was completed three times by the mothers, when the child was aged 30, 42 and 57 months. The PSAI consists of 24 items, 12 masculine and 12 feminine and is divided into three sections which focus on the frequency of: playing with different toys, such as guns or dolls (7 items); engaging in specific activities such as climbing or playing at taking care of babies (11 items); displaying certain traits such as exploring new surroundings or avoiding taking risks (6 items). Hence each section measures typically masculine and feminine behaviours. Question responses range from "never" to "very often" on a 5-point Likert scale and have a score of 1 to 5. For example, a response of "never" gains a score of 1 and a response of "always" gains a score of 5. The PSAI score is calculated using the following formula (Golombok & Rust, 1993):

PSAI score = 1.1 * (sum of masculine items – sum of feminine items) + 48.25.

The more masculine the behaviour, the higher the score; the more feminine the behaviour, the lower the score.

Previous research indicates that the measure has adequate test-retest reliability and validity (Golombok & Rust, 1993). Internal consistency in the current study was satisfactory, with Cronbach's alpha of .73, .75 and .77 for PSAI scores at 30, 42 and 57 months respectively. Previous studies also report acceptable levels of internal consistency for

example, Golombok & Rust, (1993b) reported split-half reliability of 0.88 (N = 2330) and test-retest reliability over 1 year of 0.64 (N = 33).

Age standardised (AS) PSAI scores were calculated using the following equations (Golombok & Rust, 1993):

In boys, B (AS PSAI score) = A (Initial PSAI score) * 1.075 – Age (months) * .177 + 0.88

In girls, B (AS PSAI score) = A (Initial PSAI score) * 1.026 + Age (months) * .218 – 8.33.

The rationale for selecting the PSAI score at 30 months as input variable was because it was the youngest age at which the child could potentially be expressing their gender nonconforming preferences with the least social influence against gender nonconforming behaviour.

Child Activities Inventory (CAI)

The Children's Activity Inventory (CAI) is an adapted form of the Preschool Activities Inventory (PSAI; Golombok & Rust, 1993), which consists of 16 items; eight of which are masculine and eight are feminine. Feminine items include playing with jewellery, dolls and tea sets; while masculine items include playing with computer games, toy guns or wrestling with friends. This measure was scored in the same way as the PSAI items.

The Harter Self-Perception Profile for Children (SPPC)

Self-esteem was measured using the Global Self Worth subscale from the Harter SPPC (Harter, 1985), which was completed by the child and consisted of six items, such as "some kids are often unhappy with themselves, BUT other kids are pretty pleased with themselves". Item responses ranged from really true of me (1 point) or sort of true of me (2

points) for the negative statement, and sort of true of me (3 points) or really true of me (4

points) for the positive statement. Reverse coding is necessary for three items on the scale and

higher scores indicate higher levels of self-esteem in the child. The test-retest reliability is

satisfactory (Donnellan, Trzesniewski, Robins, Moffitt, & Caspi, 2005). Previous studies

reported satisfactory Cronbach's alpha of .71 (e.g. Lerner et al., 2005).

The Short Mood and Feelings Questionnaire (SMFQ)

Depression at age 10, as completed by the child, was measured using the SMFQ

(Angold et al., 1995), a questionnaire that consisted of 13 items, including feeling miserable,

restless and lonely. The SMFQ is a subset of the 34-item Mood and Feelings Questionnaire

(MFQ; Costello & Angold, 1988). Responses of not true, sometimes true and true attracted

scores of 0, 1 & 2 respectively. Higher scores indicated higher rates of depression. In terms of

reliability, factor loading correlations were high with coefficient alphas being .81, .84 and .87

for cohorts 1, 2 and 3 (Messer, 1995).

Maternal occupation

Mother's occupation is an established measure of socio-economic status generated by

the Office of National Statistics (ONS). It is included as a potential confounding variable in

this study as it has been associated with depression in other studies (e.g. Skidmore et al.,

2006). The categories of classification are: I Professional, II Managerial and Technical, IIINM

Skilled non-manual, IIIM Skilled manual, IV Partly skilled, V Unskilled.

Dummy codes were set up to provide dichotomous variables for each level of

classification and these were named professional, managerial, skilled non-manual etc. For

example, the dummy code named "Professional" contained a score of '1' for cases where

classification was professional and a score of '0' for cases that contained any other

classification. This process was repeated for each dichotomous variable and these were added to the linear regression equation, the results of which are provided in the results section.

Procedure

The data were obtained by Dr Katharine Rimes, Institute of Psychiatry, Psychology and Neuroscience, King's College London. Prior to this a detailed project proposal was submitted to, and accepted by, the ALSPAC research committee and the Institute of Psychiatry, Psychology and Neuroscience (IOPPN) Ethics committees.

With respect to the ALSPAC data collection process, parents completed surveys about themselves and their study child's health and development from the early stages of pregnancy onwards. Additionally, from the age of 7, the study children attended annual physical and psychological assessments. Further information on the wealth of information available from ALSPAC can be found in their data dictionary (HTTP://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary).

The sample for the current study consists of a substantial subset of data from the full dataset of participants. This incorporated selected demographic, psychological and social indicators to investigate mental health in relation to sexual orientation and gender nonconformity. It included Preschool Activities Inventory (PSAI; Golombok & Rust, 1993) data collected at 30, 42 and 57 months, plus Children's Activity Inventory (CAI, adapted from the PSAI) data collected at 8 years of age. Data on self-esteem was obtained from the Harter Self Perception Profile for Children (Harter, 1985) and data on depression and other mental health issues were obtained from the Short Mood and Feelings Questionnaire (SMFQ; Angold et al., 1995) and the Development and Well-Being Assessment (DAWBA; Meltzer, Gatward, Goodman & Ford, 2000).

Data preparation

The PSAI and CAI unadjusted raw scores were calculated as described in the Materials section above. Subsequently, Age standardised PSAI scores were calculated for boys and girls separately, using the equations detailed in the Materials section. Analyses were conducted using SPSS v23.

Multiple Imputation of missing variables

A common problem with longitudinal studies is loss to follow up resulting in significant levels of missing data and a missing data analysis is provided in the results section.

An analysis of the pattern of missing values was performed in SPSS v23 using the Estimation Maximisation (EM) algorithm (Little & Rubin, 2014). Firstly, a missing values analysis function was selected from the Multiple Imputation menu option in SPSS, followed by an Impute Missing Values function. The former provided a numerical and graphical analysis of missing values, while the latter produced multiple copies (five) of the original dataset with imputed values of previously missing data. This analysis was performed on the full dataset and therefore includes the cases that were excluded from the main analyses as they did not include sex.

Key study variables auxiliary variables were added into the multiple imputation (MI) model. The auxiliary variables included DAWBA depressive disorder diagnoses at ages 7, 10 and 13, and socioeconomic status.

Statistical analysis

Descriptive statistics and percentages of missing data were produced for the key study variables. This was then followed by running the MI process as described above.

Associations between the main study variables were calculated used the Pearson Correlation Coefficient. This also provided an opportunity to assess the stability of the PSAI score over time. Mediation analyses were then performed which tested whether the association between CGN at 30 months and depression at 10 years being was mediated by self-esteem (Hayes, 2012). This was done separately on data from males and females.

<div align="center">**Results**</div>

An analysis was performed using SPSS version 23. Associations between gender nonconformity at 30, 42 & 57 months, self-esteem at 8 years and depression at age 10 were measured using the Pearson correlation statistic. Mediation between gender nonconformity at 30 months (Independent variable), self-esteem at 8 years (Mediator variable) and depression at age 10 (Dependent Variable) was performed using the Hayes Process macro for SPSS (Hayes, 2012).

Descriptive and missing data

A common problem with longitudinal studies is loss to follow up resulting in significant levels of missing data. Descriptive statistics and percentage missing data are provided in tables 1 & 2 below for females and males respectively. Although CGN at 42 months, 57 months and 8 years have not been included in the mediation analysis, they are reported in tables 1 & 2 as later in the results section, they are compared to CGN at 30 months in order to obtain an indication of the stability of the score over time. Similarly, social class based on mother's occupation (hereafter referred to as mother's occupation) has been

included as it was investigated as a potential confounding variable (tables 7 & 8). The

descriptive statistics refer to the valid data i.e. to those fields where the data is present.

Table 1

*Number and percentage of missing data and descriptive data including mean, median, standard
deviation in females.*

	CGN at 30 months	CGN at 42 months	CGN at 57 months	CGN at 8 years	Self-esteem at 8	Depression at 10	Social class (mother)	Socio economic group (mother)
N valid	4925	4834	4537	3627	3489	3734	5304	5304
N missing	2294	2385	2682	3592	3730	3485	1915	1915
% missing	32	33	37	50	52	48	27	27
Mean	40.08	38.77	40.11	40.04	19.33	3.91	3.15	6.31
Median	39.81	39.04	40.06	39.55	20.00	3.00	3.00	6.00
Std. Deviation	8.77	9.55	9.62	12.44	3.36	3.56	1.22	2.22

Table 2

Number and percentage of missing data and descriptive data including mean, median, standard deviation in males:

	CGN at 30 months	CGN at 42 months	CGN at 57 months	CGN at 8 years	Self-esteem at 8	Depression at 10	Social class (mother)	Socio economic group (mother)
N valid	5303	5198	4871	3530	3445	3629	5763	5763
N missing	2332	2437	2764	4105	4190	4006	1872	1872
% missing	31	32	36	54	55	52	25	25
Mean	60.91	60.56	59.88	59.85	19.12	4.17	3.17	6.39
Median	60.45	60.69	60.40	62.35	20.00	3.00	3.00	6.00
Std. Deviation	8.98	9.29	9.43	11.48	3.49	3.45	1.23	2.19

From tables 1 and 2 it is clear that there are large amounts of missing data for the key study variables. For example in females, CGN varies between 32% missing data at 30 months to 50% at 8 years. Additionally there is 48% and 52% missing data for self-esteem and depression respectively.

As expected, the mean CGN scores for females and males are approximately 40 and 60 respectively. The scales were constructed to provide means around these figures. Please refer to the measures sub-section of the method section for details on the calculation of CGN in females and males (Golombok & Rust, 1993).

Missing data analysis

Following on from this, the high levels of missing data necessitated an analysis of the pattern of missing values was performed in SPSS v23 using the Estimation Maximisation (EM) algorithm. Firstly, a missing values analysis function was selected from the Multiple Imputation menu option, followed by an Impute Missing Values function. The former provided a numerical and graphical analysis of missing values, while the latter produced multiple copies (five) of the original dataset with imputed values of previously missing data. This analysis was performed on the full dataset and therefore includes the 592 cases that were excluded from Tables 1 & 2, as they did not include sex.

Figure 1 below, demonstrates that all variables contained some missing data, while almost 80% of cases contained missing variables. In all approximately 43% of data was missing. The percentage of missing data for the key study variables plus additional related variables are displayed in table 3. In multiple imputation, additional variables that are related to the study variable, such as further gender nonconformity and depression variables, increase the accuracy of the of the MI process and output.
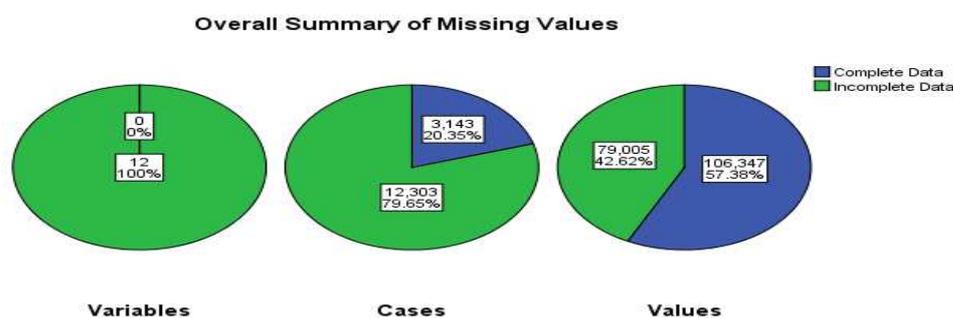
Figure 2. Missing values

Table 3

A summary of the variables utilised in the multiple imputation process

Variable Name	No. missing	Percentage missing	No. of valid entries	Mean	Standard Deviation
Self-esteem at 8 years old	8505	55.1	6941	19.2	3.43
Depressive disorder at 13 years – Y/N	8486	54.9	6960	n/a	n/a
CGN at 8 years	8282	53.6	7164	49.81	15.54
Depression score at 10 years old	8077	52.3	7369	4.04	3.51
Depressive disorder at 10 years – Y/N	7789	50.4	7657	n/a	n/a
Depressive disorder at 7 years – Y/N	7357	47.6	8089	n/a	n/a
CGN at 57 months	6038	39.1	9408	50.34	13.72
CGN at 42 months	5414	35.1	10032	50.06	14.4
CGN at 30 months	5218	33.8	10228	50.88	13.68
Socio economic group (mother)	4319	28	11127	6.35	2.21
Social class based on occupation (mother)	4319	28	11127	n/a	n/a

A comparison of correlations from the original dataset and the imputed dataset

A comparison of the correlations between the main study variables from the original dataset, and the pooled dataset, which consisted of the mean of the fifty imputed datasets (i.e. combining the results for the fifty imputed datasets), was performed (tables 4 and 5). This comparison demonstrated negligible differences between the original and pooled data (table

6). Therefore, any subsequent analysis was performed on the original data. The rationale for

using the original dataset as opposed to the imputed data set was partly based on the

negligible differences between original and pooled data. Additionally, the data were not

missing at random, thereby negatively influencing the validity of the MI process.

Furthermore, percentages of missing data varied from 28% to 55%, whereas the

recommended maximum levels of missing data are 10 to 15% to ensure the validity of the MI

process (Von Hippel, 2004). These factors outweighed the larger sample size provided by

multiple imputation. In tables 4, 5 & 6, female correlations are provided above the diagonal

line (i.e. above and to the right of the diagonal) and male correlations are provided below the

diagonal line (i.e. below and to the left of the diagonal).

Assessing whether data were normally distributed

As this is a large dataset, it was not appropriate to perform statistical tests of normality such

as the Shapiro-Wilks test. Histograms were produced to visually assess whether the data were

approximately normally distributed. This was the case for the gender nonconformity measures

in males and females, but not for self-esteem and depression in males and females (see

Appendix H). However, all variables were included in the analysis, due to the use of the

bootstrapping sampling technique (MacKinnon, Lockwood & Williams, 2004), as

recommended by Hayes (2013).

Table 4

Female and male correlations[1] between gender nonconformity (CGN), self-esteem at 8 years and depression[1]. Original dataset.

Pooled data analysis		CGN at 30 months	CGN at 42 months	CGN at 57 months	CGN at 8 years	Self-esteem at 8 years	Depression at 10 years
CGN at 30 months	Correlation	-	0.634**	0.524**	0.138**	-0.052**	0.046**
	Sig. (2-tailed)		0.000	.000	.000	.005	.009
	N	5303	4447	4203	3110	2992	3154
CGN at 42 months	Correlation	.630**	-	0.685**	0.193**	-0.052**	0.052**
	Sig. (2-tailed)	0.000		0.000	.000	.005	.004
	N	4782	5198	4248	3088	2980	3134
CGN at 57 months	Correlation	.509**	.663**	-	0.263**	-0.059**	.026
	Sig. (2-tailed)	.000	0.000		.000	.001	.156
	N	4505	4547	4871	3530	2934	3079
CGN at 8 years	Correlation	.086**	.125**	.154**	-	-0.094**	.006
	Sig. (2-tailed)	.000	0.000			.000	.735
	N	3075	3066	2985	3039	3417	3148
Self-esteem at 8 years	Correlation	-0.0004	0.013	0.009	.072**	-	-0.248**
	Sig. (2-tailed)	.982	.489	.636	.000		.000
	N	3002	2992	2920	3327	3445	3033
Depression at 10 years	Correlation	.044*	0.009	-.003	-.046*	-.211**	-
	Sig. (2-tailed)	.015	.622	.888	.013	.000	
	N	3114	3118	3042	2990	2923	3629

[1] Female correlations are presented above the diagonal line, male correlations below.

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 5

Female and male correlations1 between gender nonconformity (CGN), self-esteem at 8 years and depression[1]. Pooled dataset.

Pooled data analysis		CGN at 30 months	CGN at 42 months	CGN at 57 months	CGN at 8 years	Self-esteem at 8 years	Depression at 10 years
CGN at 30 months	Correlation	-	0.636**	0.521**	0.139**	-0.063**	0.048**
	Sig. (2-tailed)		0.000	0.000	.000	.000	.006
	N	6283	5953	5953	5953	5953	5768
CGN at 42 months	Correlation	.634**	-	0.686**	0.19**	-0.054**	0.056**
	Sig. (2-tailed)	0.000		0.000	0.000	.003	.001
	N	6283	6283	5953	5953	5953	5768
CGN at 57 months	Correlation	.508**	.662**	-	0.263**	-0.066**	.030
	Sig. (2-tailed)	0.000	0.000		0.000	.000	.091
	N	6283	6283	6283	5953	5953	5768
CGN at 8 years	Correlation	.079**	.118**	.147**	-	-0.092**	-.007
	Sig. (2-tailed)	.000	0.000	0.000		.000	.672
	N	6283	6283	6283	6283	5953	5768
Self-esteem at 8 years	Correlation	-0.002	0.005	0.019	.072**	-	-0.252**
	Sig. (2-tailed)	.933	.785	.275	.000		0.000
	N	6283	6283	6283	6283	6283	5768
Depression at 10 years	Correlation	.051**	0.018	.009	-.043*	-.204**	-
	Sig. (2-tailed)	.003	.245	.566	.023	0.000	
	N	6283	6283	6283	6283	6283	6283

[1] Female correlations are presented above the diagonal line, male correlations below.

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 6

Differences between correlation coefficients for original and pooled data[1]

Actual differences	CGN at 30 months	CGN at 42 months	CGN at 57 months	CGN at 8 years	Self-esteem at 8 years	Depression at 10 years
CGN at 30 months	-	-0.002	0.003	-0.001	0.011	-0.002
CGN at 42 months	-0.004	-	-0.001	0.003	0.002	-0.004
CGN at 57 months	0.001	0.001	-	0.000	0.007	-0.004
CGN at 8 years	0.007	0.007	0.007	-	-0.002	0.001
Self Worth at 8 years	0.001	0.008	-0.011	0.000	-	0.004
Depression at 10 years	-0.007	-0.009	-0.012	-0.003	-0.007	-
Percentage differences						
CGN at 30 months (male)	-	-0.315	0.573	-0.725	-21.154	-4.348
CGN at 42 months (male)	-0.635	-	-0.146	1.554	-3.846	-7.692
CGN at 57 months (male)	0.196	0.151	-	0.000	-11.864	-16.762
CGN at 8 years (male)	8.140	5.600	4.545	-	2.128	-19.008
Self-esteem at 8 years (male)	-275.239	62.365	-121.041	0.000	-	-1.613
Depression at 10 years (male)	-15.909	-107.029	459.143	6.522	3.318	-

[1] Female correlations are presented above the diagonal line, male correlations below.

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Although the majority of percentage differences are negligible, it is evident that large percentage differences exist in the self-esteem and depression scores, especially in males. However, these large percentage differences are based on very small correlation coefficients and are not considered sufficiently significant to warrant the use of the Multiple Imputation (MI) datasets, especially in view of the large quantities of missing data, which would invalidate the MI process.

Consideration of the confounding variable of maternal occupation

A multiple linear regression was used to test whether the subject's maternal occupation, acted as a confounding variable. As this is a categorical variable, dummy variables were set up for each category, with the exception of Professional, which was selected as the reference variable. Firstly, the multiple linear regression was run with CGN at 30 months and self-esteem as the input variables and depression at age 10 as the output variable. For males, this provided an Adjusted R Square value of .047. When maternal occupation, was added to the model, the Adjusted R Square increased to .050. For females, the original Adjusted R Square value was .059, and with the addition of maternal occupation it reduced to .057. The results are presented in tables 7 and 8 below. Because of the minimal effects of maternal occupation, it was not included in the mediation analysis as a confounder.

Table 7

Multiple linear regression results with input variables of CGN at 30 months and Self-esteem at 8 years in females. Maternal occupation variables are added in Model 2.

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	SE	Beta	t	Sig.
1	(Constant)	8.410	.506		16.620	.000
	CGN at 30 months	.008	.008	.019	1.008	.314
	Self-esteem at 8 years	-.252	.020	-.243	-12.829	.000
2	(Constant)	8.517	.530		16.075	.000
	CGN at 30 months	.007	.008	.018	.949	.343
	Self-esteem at 8 years	-.251	.020	-.242	-12.763	.000
	Professional v Manager	-.140	.202	-.018	-.691	.490
	Professional v Skilled Non-manual	-.228	.195	-.032	-1.165	.244
	Professional v Skilled Manual	.171	.420	.008	.408	.683
	Professional v Part skilled	.049	.253	.005	.194	.846
	Professional v Unskilled	.206	.465	.009	.443	.658

a. Dependent Variable: Depression Score: F10+

Model 1, Adjusted R Square = .059. Model 2 Adjusted R Square = .057.

Table 8

Multiple linear regression results with input variables of CGN at 30 months and Self-esteem at 8 years in males. Maternal occupation variables are added in Model 2.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	SE	Beta		
1	(Constant)	7.330	.576		12.722	.000
	CGN at 30 months	.014	.007	.037	1.915	.056
	Self-esteem at 8 years	-.214	.019	-.214	-11.151	.000
2	(Constant)	7.304	.591		12.366	.000
	CGN at 30 months	.013	.007	.034	1.779	.075
	Self-esteem at 8 years	-.214	.019	-.214	-11.118	.000
	Professional v Manager	-.094	.210	-.012	-.449	.654
	Professional v Skilled Non-manual	.122	.202	.017	.604	.546
	Professional v Skilled Manual	.168	.408	.009	.411	.681
	Professional v Part skilled	.316	.252	.031	1.254	.210
	Professional v Unskilled	.426	.491	.018	.867	.386

b. Dependent Variable: Depression Score: F10+

Model 1 Adjusted R Square = .047. Model 2 Adjusted R Square = .05.

Correlations between childhood gender nonconformity, self-esteem and depression.

Correlations between the main study variables are provided in Table 9, with results above the diagonal line representing females and those below the diagonal line representing males. In females, the correlations between GN at 30 months, self-esteem at 8 years and depression at age 10, are significant if very weak [r = -.052, p = .005] and [r = .046, p = .009],

respectively. The correlation between self-esteem and depression was relatively stronger in comparison to the preceding correlations [r = -.248, p < .001].

In males, the correlations between GN at 30 months, self-esteem at 8 years and depression at age 10 were very weak [r = .006, p = .759] and [r = .044, p = .015], respectively, with only the latter being significant. The correlation between self-esteem and depression was relatively stronger in comparison to the preceding correlations [r = -.211, p < .001]. As in previous correlations, the female results are provided above and to the right of the diagonal and the male results are present below and to the left of the diagonal. All valid entries were analysed for each variable in order to maximise the sample sizes for each correlation.

Table 9

Correlations between gender nonconformity, self-esteem *and* depression *in females and males[1]*

		Gender non-conformity	Self-esteem	Depression
Gender non-conformity	Pearson r	-	-.052**	.046**
	Significance		.005	.009
	N	4925	2992	3154
Self-esteem	Pearson r	.006	-	-.248**
	Significance	.759		.000
	N	2636	3489	3033
Depression	Pearson r	.044*	-.211**	-
	Significance	.015	.000	
	N	3114	2923	3734

**. Correlation is significant at the 0.01 level (2-tailed).
* . Correlation is significant at the 0.05 level (2-tailed).
[1] Results above the diagonal line refer to females and results below the diagonal line refer to males.

The results for females support hypothesis 1, in that higher rates of gender non-conformity, are associated with lower self-esteem and higher levels of depression. The hypothesis is partially supported by the results for males, as the correlation between gender non-conformity and self-esteem is very weak and not significant.

Mediation analyses

Mediation analyses were performed to determine whether self-esteem performed a mediating role in the association between gender nonconformity and depression. In females, although there was not a significant direct effect observed between GN and depression, there was a significant indirect effect of gender nonconformity on depression, B = .006, 95% CI [0.002, 0.0095]. This represents a very small effect, kappa squared = .0138, 95% CI [.0045, .235]. The mediation model for females is provided below:
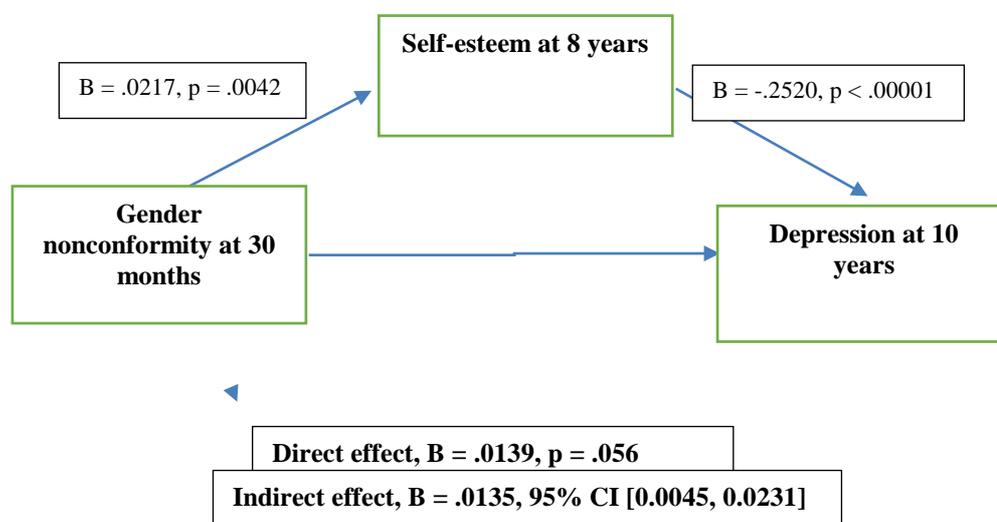
Figure 3: Mediation analysis for females

In males, there was no significant direct or indirect effect of gender nonconformity on depression, B = .0139, p = .056 and B = -.0003, 95% CI [-0.0080, 0.0089] respectively. Therefore, the hypothesis that self-esteem mediates the association between CGN and depression could not be supported. Hence the mediation diagram for males has not been provided.

Discussion

Overview

The aim of this study was to investigate the associations between early childhood gender nonconformity (CGN) and childhood depression and to ascertain if this association was mediated by low self-esteem. The study hypothesis stated that higher levels of early CGN would result in higher levels of depressive symptoms. It was also hypothesised that there would be different findings in females and males, as previous studies have demonstrated that distress related to CGN is higher in males (e.g. Skidmore et al., 2006). However rates of depression are generally higher in females from adolescence onwards (Mills & Baker, 2016).

Before commencing the analysis, the potential use of multiple imputation of missing variables (MI) was explored. However, levels of missing data were too high and non-random to justify the use of the multiple imputation model. Additionally, as already discussed in the results section, an analysis between the original and imputed datasets demonstrated minimal variability. Therefore the decision was made to analyse the original dataset in the mediation analysis.

Previous studies have considered the potentially confounding variable of social class when investigating the role of various factors in the development of depressive symptoms (e.g. Skidmore et al., 2006). Multiple linear regressions in the female and male samples demonstrated that an individual's social class, as based on their mother's occupation, had a very small and insignificant impact on the regression model. Therefore it was not included in the mediation analysis (Hayes, 2012).

The current study found that, in females, there were very small but significant correlations between childhood gender nonconformity (CGN) at 30 months, self-esteem at 8

years and depressive symptoms at 10 years. The mediation model demonstrated that this association was mediated by self-esteem. In males there was a very small, significant correlation between CGN at 30 months and depressive symptoms at 10 years, but no correlation was found with self-esteem. There was a small, significant negative correlation between self-esteem and depressive symptoms. The mediation model showed that there was no significant direct or indirect effect of CGN on depressive symptoms.

Discussion of study findings

Although the effect sizes are very small, the significant association between CGN and childhood depression in females and males partially supports the results of the Petterson et al., (2016) study, which only found this association in males. The difference between Petterson et al., (2016) and the current study includes the fact that the current study used prospective measurements of parent-reported CGN and self-reported childhood depressive symptoms whereas Petterson et al., (2016) relied on retrospective self-reports of both, leaving their results open to potential recall bias. Additionally, they used the CGIS which asks participants to recall childhood gender related behaviours before the age of 12, whereas the current study measures CGN at 30 months. This current finding also supports the results of Feinstein et al., (2012). However, they also relied on retrospective self-reports of CGN and current depressive symptoms in an adult sample, whereas the current study and Petterson et al., (2016) investigated childhood depression. Similarly, Alanko et al., (2009) noted associations between retrospectively recalled CGN and psychological distress in adulthood, a measure which combines depressive and anxiety symptoms, in females and males. Current findings are similar to Skidmore et al., (2006), who did not find an association between retrospectively recalled CGN and current depressive symptoms in adult females. However, as with other studies (e.g. Petterson et al., 2016) they rely on retrospective accounts of CGN

across childhood, rather than measuring it at a specific point in time. Additionally, their

study sample was a relatively small opportunity sample which was limited to lesbians and gay

males, thereby limiting the generalisability of their findings.

All of the above studies are cross sectional, and it is therefore not possible to infer

causation from the results. In their longitudinal study, Roberts et al., (2013) found that, in

comparison to childhood gender conforming (CGC) participants, CGN participants reported

significantly higher rates of depression from early adolescence through to early adulthood and

this was mediated by child abuse and bullying. This applied for females and males. However,

CGN was assessed retrospectively when the mean sample age was 20.

CGN is only one of many variables which may be contributing to childhood

depressive symptoms. Difficult early developmental experiences, genetic predisposition and

stressful life events such as abuse or bereavement are key factors which may lead to the

development of childhood depression (e.g. Mills & Baker, 2016). Additionally, parental

mental health problems, substance misuse and domestic violence are associated with

increased rates of childhood depression (e.g. Mills & Baker, 2016). When taking all of these

other factors into account, it is perhaps unsurprising that only weak associations were found

between CGN and depression. Additionally, the period between measuring CGN at 30

months and depression at 10 years, represents a long period of time in a child's life during

which many of the aforementioned risk factors may be precipitated and perpetuated.

This is the first study between CGN and depression that investigates self-esteem as a

mediator, whereas previous studies have investigated abuse and bullying victimisation

(Roberts et al., 2013), quality of relationship with parents (Alanko et al., 2009) and

homonegativity (Feinstein et al., 2011). They all reported mediation effects in both genders,

whereas the current study only found the mediation effect in females. The finding in females supports the review of Sowislo & Orth (2013), who found evidence to support the vulnerability model, in which low self-esteem contributes to depression. Additionally, in their review of studies on prevalence of self-esteem, Kling et al., (1999) found that low self-esteem was more common in females than males and this pattern was evident for children, adolescents and adults. Furthermore, the mediation effect in females supports the role of self-esteem in the development of psychological problems, as outlined in the cognitive model of depression (Beck, 1976) and the psychodynamic model of distress (Fonagy et al., 2002).

On the other hand, the absence of a mediation effect in males is surprising as boys high in CGN have more negative experiences than their female counterparts and this might be expected to have a more negative impact on their self-esteem in comparison to CGN girls (e.g. Meyer, 2003; Roberts et al., 2013). However there may be factors at play which are protective of the boys' self-esteem, such as positive parental support (e.g. Alanko et al., 2009). Additionally, there may be other factors affecting self-esteem that have not been investigated in this study. This is an interesting result that requires further exploration.

Limitations

As with the majority of other studies in this area, it is not possible to infer causation from the results. However, many of the studies in this area are cross sectional and therefore it adds important findings to the evidence base. In common with other longitudinal datasets, there were significantly high amounts of missing data which were too high to support the use of multiple imputation of missing variables (MI). This calls into question the generalisability of the findings, as does the fact that this is a geographically limited sample, as all participants were recruited from a relatively small part of south west England. Hence, it may not be

representative of other parts of the UK, or indeed other countries, which have a different ethnic or cultural mix. There is a large time difference between measurement of CGN and depression and this opens up the potential for other influences on childhood depression to impact findings. However, the intention of measuring CGN at 30 months was to ensure minimal impact of social influences on gender atypical behaviour (e.g. Gecas, 1992).

Research implications

The very small effect sizes in this study highlight the need for further research in this area. This is one of the few studies to investigate early CGN, self-esteem and depressive symptoms in childhood and there is a pressing need for further studies as this is an area which is gaining increasing attention. Longitudinal studies would be beneficial as they permit the inference of causation and it is important to be able to inform policy makers that early CGN results in increased depressive symptoms in addition to other mental health problems such as anxiety, if such a causal relationship can be established. However they tend to be very expensive and do suffer loss to follow up, therefore additional cross sectional studies will help to strengthen the evidence base in this area.

It is important that any new studies should sample the entire population, as previous earlier studies, obtained opportunity samples from LGBT communities and would even take a subset of this population, such as lesbians and gay males only (Skidmore et al., 2006).

It would also be beneficial to have studies from more diverse societies and cultures, in order to obtain a more rounded understanding of the associations between early CGN and mental health problems such as depression. Furthermore, it would be helpful to gain an improved understanding of the current situation in the UK, as there have been few studies

published in relation to CGN and depressive symptoms. These are needed to inform clinical practice.

Qualitative studies are also necessary in order to develop a deeper, more nuanced understanding of individuals' experiences of CGN. In combination with data on changing societal views, these could help researchers and policy makers to understand the issues in more detail and make recommendations on improving outcomes. These may be as broad as community psychology and public health initiatives to improve peoples' awareness of, and attitudes towards gender nonconforming individuals, perhaps by encouraging more flexible thinking around gender roles and challenging the gender-specific stereotypes. Discussion of the impact of stigma on these individuals, in terms of increased mental health problems, may help the general public to develop increased tolerance towards gender nonconforming individuals. It may also be validating for GN individuals to know that their unique experiences within society, resulting from other peoples' attitudes and behaviours, have been acknowledged.

Clinical implications

It is important that mental health professionals and other healthcare professionals' attitudes towards nonconforming individuals are positive and supportive. Beliefs around gender roles are developed at an early age and it may be difficult to change these beliefs and subsequent assumptions. Further, by developing a broader and more in depth understanding of CGN, depression and other mental health problems, in addition to the mechanisms which might explain the association, interventions may be developed in order to target these mechanisms. Examples of proposed mechanisms have been discussed previously and include child abuse and bullying (Roberts et al., 2013), quality of the individual's relationship with

their parents (Alanko et al., 2009), rejection sensitivity and homonegativity (Feinstein et al., 2011) and self-esteem as investigated in the current study. The evidence base for cognitive behavioural therapy for low self-esteem (e.g. Fennell, 2016) is increasing and based on current findings it may be useful to consider this intervention for females. In terms of abuse and bullying, Stonewall have comprehensive LGBT bullying initiatives and work closely with schools to ensure that these are implemented. Their materials include materials on GN and provide guidance and support to parents of GN children. Commissioners also need to be aware of the unique experiences of minority groups in order to ensure that services meet the needs of this population.

Conclusion

This study reported mixed findings in that the association between CGN and childhood depressive symptoms was mediated by low self-esteem in females, albeit with a very small effect size, whereas in males, the mediation analysis produced a direct effect with a very small effect size, which approached significance ($p = .056$). However, there was no indirect effect of CGN and self-esteem on childhood depressive symptoms. This surprising finding highlights the need for further research, especially as there are few studies investigating CGN and depression in childhood. Ideally there should be more UK based studies as there are few at present. Longitudinal studies are preferable as they allow the inference of causation. They also allow the use of prospective measures, which are preferable to the retrospective measures often used in cross sectional studies. This will help to make a stronger case for evidence based clinical interventions. However cross sectional studies are also desirable and perhaps more realistic as they often require less funding. Studies from diverse cultures are also welcome as they may help us to understand the cultural factors which improve the experiences of GN individuals, such as societal acceptance or more

flexible views on different forms of gender expression. An increased understanding of the mechanisms that explain and influence the association between CGN and childhood depression will facilitate the development and implementation of effective clinical interventions.

REFERENCES

Alanko, K., Santtila, P., Witting, K., Varjonen, M., Jern, P., Johansson, A., ... & Kenneth Sandnabba, N. (2009). Psychiatric symptoms and same-sex sexual attraction and behavior in light of childhood gender atypical behavior and parental relationships. *Journal of Sex Research*, *46*, 494-504. doi:10.1080/00224490902846487

Angold, A., Costello, E. J., Messer, S. C., & Pickles, A. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research*, 5, 237-249. Retrieved from http://devepi.duhs.duke.edu/AngoldMFQarticle.pdf

Baams, L., Beek, T., Hille, H., Zevenbergen, F. C., & Bos, H. M. (2013). Gender nonconformity, perceived stigmatization, and psychological well-being in Dutch sexual minority youth and young adults: A mediation analysis. *Archives of Sexual Behavior*, *42*, 765-773. doi: 10.1007/s10508-012-0055-z

Beck, A.T. (1976). *Cognitive therapy and the emotional disorders*. New York: New American Library.

Boden, J.M. & Fergusson, D.M. (2011). Alcohol and depression. Addiction, 106, 906 – 914. doi:10.1111/j.1360-0443.2010.03351.x

Boyd, A., Golding, J., Macleod, J., Lawlor, D. A., Fraser, A., Henderson, J., ... & Smith, G. D. (2012). Cohort profile: the 'children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, dys064. doi: 10.1093/ije/dys064

Costello, E. J., Foley, D. L., & Angold, A. (2006). 10-year research update review: the epidemiology of child and adolescent psychiatric disorders: II. Developmental epidemiology. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(1), 8-25. doi:10.1097/01.chi.0000184929.41423.c0

Costello, E. J., & Angold, A. (1988). Scales to assess child and adolescent depression: checklists, screens, and nets. *Journal of the American Academy of Child & Adolescent Psychiatry*, *27*(6), 726-737. doi: 10.1097/00004583-198811000-00011

Ceglie, D., Freedman, D., & McPherson, D. S. & Richardson, P. (2002) Children and adolescents referred to a specialist gender identity development service: Clinical features and demographic characteristics. *The International Journal of Transgenderism*, *6*(1).

Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., Moffitt, T. E., & Caspi, A. (2005). Low self-esteem is related to aggression, antisocial behavior, and delinquency. *Psychological Science*, *16*(4), 328-335.

Fairburn, C.G. & Harrison, P.J. (2003). Eating disorders. *The Lancet, 361*, 407 – 416. doi: 10.1016/S0140-6736(03)12378-1

Feinstein, B. A., Goldfried, M. R., & Davila, J. (2012). The relationship between experiences of discrimination and mental health among lesbians and gay men: An examination of internalized homonegativity and rejection sensitivity as potential mechanisms. *Journal of consulting and clinical psychology*, *80*, 917. doi:10.1037/a0029425

Fennell, M. (2016). *Overcoming low self-esteem: A self-help guide using cognitive behavioral techniques*. Hachette UK.

Fonagy, P., Target, M., Cottrell, D., Phillips, J., & Kurtz, Z. (2002). *What works for whom. A critical review of treatments for children and adolescents.* London: Guilford Press doi:10.1080/07317107.2016.1203157

Gecas, V. (1992). Socialization. *Encyclopaedia of Sociology, 4*, 1863 – 1872.

Golding, J. (2010). Determinants of child health and development: the contribution of ALSPAC—a personal view of the birth cohort study. *Archives of Disease in Childhood*, *95*, 319-322. doi: 10.1136/adc.2009.178954

Golombok, S., & Rust, J. (1993). The pre-school activities inventory: A standardized assessment of gender role in children. *Psychological Assessment, 5,* 131–136. doi: 10.1037/1040-3590.5.2.131

Harter, S. (1985). *Manual for the Self-perception Profile for Children: (Revision of the Perceived Competence Scale for Children).* University of Denver. doi: 10.1037/t05338-000

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]. Retrieved from http://www.afhayes.com/ public/process2012.pdf

Heim, C., Newport, D. J., Mletzko, T., Miller, A. H., & Nemeroff, C. B. (2008). The link between childhood trauma and depression: insights from HPA axis studies in humans. *Psychoneuroendocrinology*, *33*, 693-710. doi:10.1016/j.psyneuen.2008.03.008.

Holt, V., Skagerberg, E., & Dunsford, M. (2016). Young people with features of gender dysphoria: Demographics and associated difficulties. *Clinical Child Psychology and Psychiatry*, *21*(1), 108-118. doi: 0.1177/1359104514558431

Kling, K. C., Hyde, J. S., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: a meta-analysis. *Psychological Bulletin*, *125*(4), 470. doi: 10.1037/0033-2909.125.4.470

Kuster, F., Orth, U., & Meier, L. L. (2012). Rumination mediates the prospective effect of low self-esteem on depression: A five-wave longitudinal study. *Personality and Social Psychology Bulletin*, *38*(6), 747-759. doi: 10.1177/0146167212437250

Lerner, R. M., Lerner, J. V., Almerigi, J. B., Theokas, C., Phelps, E., Gestsdottir, S., ... & Smith, L. M. (2005). Positive youth development, participation in community youth development programs, and community contributions of fifth-grade adolescents: Findings from the first wave of the 4-H study of positive youth development. *The Journal of Early Adolescence*, *25*(1), 17-71. doi: 10.1177/0272431604272461

Li, M, Fitzgerald, P, & Rodin, G (2012). Evidence-based treatment of depression in patients with cancer. *Journal of Clinical Oncology, 30,* 1187–96. doi:10.1200/JCO.2011.39.7372.

Lindert, J., von Ehrenstein, O. S., Grashow, R., Gal, G., Braehler, E., & Weisskopf, M. G. (2014). Sexual and physical abuse in childhood is associated with depression and anxiety over the life course: systematic review and meta-analysis. *International Journal of Public Health*, *59*, 359-372. doi:10.1007/s00038-013-0519-5

Lippa, R. A. (2008). The relation between childhood gender nonconformity and adult masculinity–femininity and anxiety in heterosexual and homosexual men and women. *Sex Roles*, *59*(9-10), 684. doi:10.1007/s11199-008-9476-5

Little, R. J., & Rubin, D. B. (2014). *Statistical analysis with missing data*. New York: Wiley.

Meltzer, H., Gatward, R., Goodman, R., & Ford, T. (2000). *The mental health of children and adolescents in Great Britain*. HM Stationery Office.

MESSER'PhD, S. C. (1995). Development of a short questionnaire for use in

    epidemiological studies of depression in children and adolescents: Factor composition

    and structure across development. Unpublished Thesis. North-western University.

Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual

    populations: Conceptual issues and research evidence. *Psychological Bulletin, 129(5),*

    674–697. doi:10.1037/0033-2909.129.5.674

Meyer, I. H., & Dean, L. (1998). Internalized homophobia, intimacy, and sexual behavior

    among gay and bisexual men. *Psychological Perspectives on Lesbian and Gay*

    Issues, *4*, 160-186. doi: 10.4135/9781452243818.n8

Mills, S. E. E., & Baker, L. D. (2016). Childhood depression. *InnovAiT: Education and*

    Inspiration for General Practice, 9(9), 524–530. doi:10.1177/1755738016658676

Petterson, L. J., VanderLaan, D. P., & Vasey, P. L. (2016). Sex, sexual orientation, gender

    atypicality, and indicators of depression and anxiety in childhood and

    adulthood. *Archives of Sexual Behavior*, *45*, 1-10.

    doi:10.1007/s10508-016-0690-x

Pillemer, K., Suitor, J. J., Pardo, S., & Henderson, C. (2010). Mothers' differentiation and

    depressive symptoms among adult children. *Journal of Marriage and Family*, *72*,

    333-345. doi:10.1111/j.1741-3737.2010.00703.x.

Rieger, G., & Savin-Williams, R. C. (2012). Gender nonconformity, sexual orientation, and

    psychological well-being. *Archives of Sexual Behavior*, *41*(3), 611-621.

    doi: 10.1007/s10508-011-9738-0

Roberts, A.L., Rosario, M., Corliss, H.L., Koenen, K.C. & Austin, S.B. (2012). Childhood

    gender nonconformity: A risk indicator for childhood abuse and posttraumatic stress

    in youth. *Pediatrics, 129(3),* 410 – 417. doi: 10.1542/peds.2011-1804.

Roberts, A.L., Rosario, M., Slopen, N., Calzo, J.P. & Austin, S.B. (2013). Childhood gender

    nonconformity, bullying victimization, and depressive symptoms across adolescence

    and early adulthood: An 11-year longitudinal study. *Journal of the American Academy*

    of Child & Adolescent Psychiatry, 52 (2), 143 – 152. doi: 10.1016/j.jaac.2012.11.006

Rustad, J. K., Musselman, D. L., & Nemeroff, C. B. (2011). The relationship of depression

    and diabetes: pathophysiological and treatment implications.

    Psychoneuroendocrinology, *36*, 1276-1286. doi:10.1016/j.psyneuen.2011.03.005.

Saravane, D., Feve, B., Frances, Y., Corruble, E., Lancon, C., Chanson, P., ... & Azorin, J. M.

    (2009). Drawing up guidelines for the attendance of physical health of patients with

    severe mental illness. *L'Encephale*, *35*, 330-339. doi:10.1016/j.encep.2008.10.014.

Skidmore, W. C., Linsenmeier, J. A., & Bailey, J. M. (2006). Gender nonconformity and

    psychological distress in lesbians and gay men. *Archives of Sexual Behavior*, *35*, 685-

    697.   doi:10.1007/s10508-006-9108-5

Sowislo, J.F. & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A

    meta-analysis of longitudinal studies. *Psychological Bulletin, 139 (1),* 213 – 240.

    doi: 10.1037/a0028931.

Steensma, T. D., Zucker, K. J., Kreukels, B. P., VanderLaan, D. P., Wood, H., Fuentes, A., &

    Cohen-Kettenis, P. T. (2014). Behavioral and emotional problems on the Teacher's

    Report Form: A cross-national, cross-clinic comparative analysis of gender dysphoric

    children and adolescents. *Journal of Abnormal Child Psychology*, *42* (4), 635-647.

    doi:10.1007/s10802-013-9804-2

van Beusekom, G., Baams, L., Bos, H. M., Overbeek, G., & Sandfort, T. G. (2016). Gender

    nonconformity, homophobic peer victimization, and mental health: How same-sex

attraction and biological sex matter. *The Journal of Sex Research*, *53*, 98-108. doi:

10.1080/00224499.2014.993462

Von Hippel, P. T. (2004). Biases in SPSS 12.0 missing value analysis. *The American*

Statistician, *58*, 160-164.

Wallien, M. S., Van Goozen, S. H., & Cohen-Kettenis, P. T. (2007). Physiological correlates

of anxiety in children with gender identity disorder. *European child & adolescent*

psychiatry, *16*(5), 309-315. doi: 10.1007/s00787-007-0602-7

Zucker, K. J., & Bradley, S. J. (1995). *Gender identity disorder and psychosexual problems*

in children and adolescents. Guilford Press.

Zucker, K. J., Owen, A., Bradley, S. J., & Ameeriar, L. (2002). Gender-dysphoric children

and adolescents: A comparative analysis of demographic characteristics and

behavioral problems. *Clinical Child Psychology and Psychiatry*, *7*(3), 398-411.

doi: 10.1177/1359104502007003007

MAJOR   RESEARCH PROJECT SECTION C

Appendix of Supporting Material

Appendix A – Quality assessment tool for quantitative studies component ratings (EPHPP)

This has been removed from the electronic copy

Appendix B Pre-school Activities Inventory

This has been removed from the electronic copy.

Appendix C – Harter Self Perception Profile for Children

This has been removed from the electronic copy.

Appendix D – Short Mood and Feelings Questionnaire

This has been removed from the electronic copy.

Appendix E – Ethics approval letter and email confirmation that this covers the MRP

This has been removed from the electronic copy.

Appendix F. Archives of Sexual Behaviour Journal Author Guidelines

 Archives of Sexual Behavior

IMPORTANT REMINDER

Be sure to read the sections on ethical requirements located at the end of these instructions. Be sure to submit all the required ethical statements in a separate section of your manuscript entitled "Compliance with Ethical Standards." This section will print immediately before the References section in accepted articles.

Manuscript Submission

Manuscripts, in English, should be submitted to Editor's Office via the journal's web-based online manuscript submission and peer-review system:

http://aseb.edmgr.com

Inquiries regarding journal policy, manuscript preparation, and other such general topics should be sent to the Editor:

Kenneth J. Zucker, Ph.D.

e−mail: kzucker.phd@gmail.com

The online system offers easy straightforward log-in and submission; supports a wide range of submission formats [such as Word, WordPerfect, RTF, TXT, and LaTeX for manuscripts; TIFF, GIF, JPEG, EPS, PPT, and Postscript for figures (artwork)]; eliminates the need to submit manuscripts as hard-copy printouts, disks, and/or e-mail attachments; enables real-time tracking of manuscript status by the author; and provides help should authors experience any submission difficulties.

For books for review, please contact Editor Zucker at the above e-mail address for details.

Submission is a representation that the manuscript has not been published previously and is not currently under consideration for publication elsewhere. A statement transferring copyright from the authors (or their employers, if they hold the copyright) to Springer Science+Business Media, LLC is required. Upon commencement of typesetting, the contact author will receive an e-mail directing him/her to a webpage where the transfer-of-copyright form can be signed online. Such a written transfer of copyright, which previously was assumed to be implicit in the act of submitting a manuscript, is necessary under the U.S.

Copyright Law in order for the publisher to carry through the dissemination of research results and reviews as widely and effectively as possible.

Manuscript Style

- Type double-spaced and left-justified in 12-point Times New Roman font using 1-inch margins on all sides. Number all pages (including table pages and figure-caption page), except the title page, consecutively with Arabic numerals placed in the upper right-hand corner. In order to facilitate masked (previously termed "double-blind") review, leave all identifying information off the manuscript, including the title page and the electronic file name. Appropriate identifying information is attached automatically to the electronic file. Upon initial submission the title page should include only the title of the article.

- An additional title page should be uploaded as a separate submission item and should include the title of the article, author's name (including highest degree received), and author's affiliation. Academic affiliations of all authors should be included. The affiliation should include the department, institution, city, and state (or nation) and should be typed as a numbered footnote to the author's name. The title page should also include the complete mailing address, telephone number, fax number, and e-mail address of the one author designated to review proofs.

- An abstract, preferably no longer than 250 words, is to be provided as the second page.

- A list of 4–5 key words is to be provided directly below the abstract. Key words should express the precise content of the manuscript, as they are used for indexing purposes.

Illustrations

Illustrations (photographs, drawings, diagrams, and charts) are to be numbered in one consecutive series of Arabic numerals and cited in numerical order in the text. Photographs should be high-contrast and drawings should be dark, sharp, and clear. Artwork for each figure should be provided on a separate page, placed at the end of the manuscript (i.e., after the References section). Each figure should have an accompanying caption. The captions for illustrations should be listed on a separate page.

Tables should be numbered consecutively with Arabic numerals and referred to by number in the text. Each table should be typed on a separate page, placed at the end of the manuscript (i.e., after the References section), and should have a descriptive title. Center the title above the table, and type explanatory footnotes (indicated by superscript lowercase letters) below the table.

References

List references alphabetically at the end of the paper and refer to them in the text by name and year in parentheses. References should include (in this order): last names and initials of all authors, year published, title of article, name of publication, volume number, and inclusive pages. The style and punctuation of the references should conform to strict APA style −− illustrated by the following examples:

Journal Article

Meston, C. M., & Frohlich, P. F. (2000). The neurobiology of sexual function. Archives of General Psychiatry, 57, 1012−1030.

Book

Dixson, A. F. (1998). Primate sexuality: Comparative studies of the prosimians, monkeys, apes, and human beings. New York: Oxford University Press.

Chapter in a Book

Bem, D. (2000). The exotic−becomes−erotic theory of sexual orientation. In J. Bancroft (Ed.), The role of theory in sex research (pp. 67−81). Bloomington, IN: Indiana University Press.

Footnotes

Footnotes should be avoided. When their use is absolutely necessary, footnotes should be numbered consecutively using Arabic numerals and should be typed at the bottom of the page to which they refer. Place a line above the footnote, so that it is set off from the text. Use the appropriate superscript numeral for citation in the text.

Style Guide

The 2010 Publication Manual of the American Psychological Association (Sixth Edition) should be used as the style guide for the preparation of manuscripts, particularly with respect to such matters as the citing of references and the use of abbreviations, numbers, and

symbols. Manuscripts departing significantly from the Sixth Edition style will not be reviewed until a corrected manuscript has been received.

RESEARCH DATA POLICY

No Page Charges

The journal makes no page charges. Reprints are available to authors. Upon the commencement of typesetting, the contact author will receive an e-mail directing him/her to a webpage that provides reprint ordering information, including the current price schedule, and where reprints can be ordered online.

ENGLISH LANGUAGE EDITING

Springer Open Choice

In addition to the normal publication process (whereby an article is submitted to the journal and access to that article is granted to customers who have purchased a subscription), Springer now provides an alternative publishing option: Springer Open Choice. A Springer Open Choice article receives all the benefits of a regular subscription-based article, but in addition is made available publicly through Springer's online platform SpringerLink. After acceptance and upon commencement of typesetting, the contact author will receive an e-mail directing him/her to a webpage that provides Springer Open Choice ordering information, including the current price schedule, and where Springer Open Choice can be ordered online. Payment must be received in full before publication or articles will publish as regular subscription-model articles. We regret that Springer Open Choice cannot be ordered for published articles. For more information, please visit

- www.springeronline.com/openchoice

ETHICAL RESPONSIBILITIES OF AUTHORS

This journal is committed to upholding the integrity of the scientific record. As a member of the Committee on Publication Ethics (COPE) the journal will follow the COPE guidelines on how to deal with potential acts of misconduct.

Authors should refrain from misrepresenting research results which could damage the trust in the journal, the professionalism of scientific authorship, and ultimately the entire scientific endeavour. Maintaining integrity of the research and its presentation can be achieved by following the rules of good scientific practice, which include:

- The manuscript has not been submitted to more than one journal for simultaneous consideration.

- The manuscript has not been published previously (partly or in full), unless the new work concerns an expansion of previous work (please provide transparency on the re-use of material to avoid the hint of text-recycling ("self-plagiarism")).

- A single study is not split up into several parts to increase the quantity of submissions and submitted to various journals or to one journal over time (e.g. "salami-publishing").

- No data have been fabricated or manipulated (including images) to support your conclusions

- No data, text, or theories by others are presented as if they were the author's own ("plagiarism"). Proper acknowledgements to other works must be given (this includes material that is closely copied (near verbatim), summarized and/or paraphrased), quotation marks are used for verbatim copying of material, and permissions are secured for material that is copyrighted.

  Important note: the journal may use software to screen for plagiarism.

- Consent to submit has been received explicitly from all co-authors, as well as from the responsible authorities - tacitly or explicitly - at the institute/organization where the work has been carried out, **before** the work is submitted.

- Authors whose names appear on the submission have contributed sufficiently to the scientific work and therefore share collective responsibility and accountability for the results.

- Authors are strongly advised to ensure the correct author group, corresponding author, and order of authors at submission. Changes of authorship or in the order of authors are **not** accepted **after** acceptance of a manuscript.

- Adding and/or deleting authors **at revision stage** may be justifiably warranted. A letter must accompany the revised manuscript to explain the role of the added and/or deleted author(s). Further documentation may be required to support your request.

- Requests for addition or removal of authors as a result of authorship disputes after acceptance are honored after formal notification by the institute or independent body and/or when there is agreement between all authors.

- Upon request authors should be prepared to send relevant documentation or data in order to verify the validity of the results. This could be in the form of raw data, samples, records, etc. Sensitive information in the form of confidential proprietary data is excluded.

If there is a suspicion of misconduct, the journal will carry out an investigation following the COPE guidelines. If, after investigation, the allegation seems to raise valid concerns, the accused author will be contacted and given an opportunity to address the issue. If misconduct has been established beyond reasonable doubt, this may result in the Editor-in-Chief's implementation of the following measures, including, but not limited to:

- If the article is still under consideration, it may be rejected and returned to the author.

- If the article has already been published online, depending on the nature and severity of the infraction, either an erratum will be placed with the article or in severe cases complete retraction of the article will occur. The reason must be given in the published erratum or retraction note. Please note that retraction means that the paper is **maintained on the platform**, watermarked "retracted" and explanation for the retraction is provided in a note linked to the watermarked article.

- The author's institution may be informed.

COMPLIANCE WITH ETHICAL STANDARDS

To ensure objectivity and transparency in research and to ensure that accepted principles of ethical and professional conduct have been followed, authors should include information regarding sources of funding, potential conflicts of interest (financial or non-financial), informed consent if the research involved human participants, and a statement on welfare of animals if the research involved animals.

Authors should include the following statements (if applicable) in a separate section entitled "Compliance with Ethical Standards" when submitting a paper:

- Disclosure of potential conflicts of interest

- Research involving Human Participants and/or Animals

- Informed consent

Please note that standards could vary slightly per journal dependent on their peer review policies (i.e. single or double blind peer review) as well as per journal subject discipline. Before submitting your article check the instructions following this section carefully.

The corresponding author should be prepared to collect documentation of compliance with ethical standards and send if requested during peer review or after publication.

The Editors reserve the right to reject manuscripts that do not comply with the above-mentioned guidelines. The author will be held responsible for false statements or failure to fulfill the above-mentioned guidelines.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

Authors must disclose all relationships or interests that could influence or bias the work. Although an author may not feel there are conflicts, disclosure of relationships and interests affords a more transparent process, leading to an accurate and objective assessment of the work. Awareness of real or perceived conflicts of interests is a perspective to which the readers are entitled and is not meant to imply that a financial relationship with an organization that sponsored the research or compensation for consultancy work is inappropriate. Examples of potential conflicts of interests **that are directly or indirectly related to the research** may include but are not limited to the following:

- Research grants from funding agencies (please give the research funder and the grant number)
- Honoraria for speaking at symposia
- Financial support for attending symposia
- Financial support for educational programs
- Employment or consultation
- Support from a project sponsor
- Position on advisory board or board of directors or other type of management relationships
- Multiple affiliations
- Financial relationships, for example equity ownership or investment interest
- Intellectual property rights (e.g. patents, copyrights and royalties from such rights)

- Holdings of spouse and/or children that may have financial interest in the work

In addition, interests that go beyond financial interests and compensation (non-financial interests) that may be important to readers should be disclosed. These may include but are not limited to personal relationships or competing interests directly or indirectly tied to this research, or professional interests or personal beliefs that may influence your research.

The corresponding author collects the conflict of interest disclosure forms from all authors. In author collaborations where formal agreements for representation allow it, it is sufficient for the corresponding author to sign the disclosure form on behalf of all authors. Examples of forms can be found

- here:

The corresponding author will include a summary statement **on the title page that is separate from their manuscript that reflects what is recorded in the potential conflict of interest disclosure form**.

See below examples of disclosures:

Funding: This study was funded by X (grant number X).

Conflict of Interest: Author A has received research grants from Company A. Author B has received a speaker honorarium from Company X and owns stock in Company Y. Author C is a member of committee Z.

If no conflict exists, the authors should state:

Conflict of Interest: The authors declare that they have no conflict of interest.

RESEARCH INVOLVING HUMAN PARTICIPANTS AND/OR ANIMALS

1) Statement of human rights

When reporting studies that involve human participants, authors should include a statement that the studies have been approved by the appropriate institutional and/or national research ethics committee and have been performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

If doubt exists whether the research was conducted in accordance with the 1964 Helsinki Declaration or comparable standards, the authors must explain the reasons for their approach, and demonstrate that the independent ethics committee or institutional review board explicitly approved the doubtful aspects of the study.

The following statements should be included in the text before the References section:

Ethical approval: "All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards."

For retrospective studies, please add the following sentence:

"For this type of study formal consent is not required."

2) Statement on the welfare of animals

The welfare of animals used for research must be respected. When reporting experiments on animals, authors should indicate whether the international, national, and/or institutional guidelines for the care and use of animals have been followed, and that the studies have been approved by a research ethics committee at the institution or practice at which the studies were conducted (where such a committee exists).

For studies with animals, the following statement should be included in the text before the References section:

Ethical approval: "All applicable international, national, and/or institutional guidelines for the care and use of animals were followed."

If applicable (where such a committee exists): "All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted."

If articles do not contain studies with human participants or animals by any of the authors, please select one of the following statements:

"This article does not contain any studies with human participants performed by any of the authors."

"This article does not contain any studies with animals performed by any of the authors."

"This article does not contain any studies with human participants or animals performed by any of the authors."

INFORMED CONSENT

All individuals have individual rights that are not to be infringed. Individual participants in studies have, for example, the right to decide what happens to the (identifiable) personal data gathered, to what they have said during a study or an interview, as well as to any photograph

that was taken. Hence it is important that all participants gave their informed consent in writing prior to inclusion in the study. Identifying details (names, dates of birth, identity numbers and other information) of the participants that were studied should not be published in written descriptions, photographs, and genetic profiles unless the information is essential for scientific purposes and the participant (or parent or guardian if the participant is incapable) gave written informed consent for publication. Complete anonymity is difficult to achieve in some cases, and informed consent should be obtained if there is any doubt. For example, masking the eye region in photographs of participants is inadequate protection of anonymity. If identifying characteristics are altered to protect anonymity, such as in genetic profiles, authors should provide assurance that alterations do not distort scientific meaning. The following statement should be included:

Informed consent: "Informed consent was obtained from all individual participants included in the study."

If identifying information about participants is available in the article, the following statement should be included:

"Additional informed consent was obtained from all individual participants for whom identifying information is included in this article."

Appendix G - Brief feedback summary report

The aim of Section A of this MRP was to investigate and evaluate the evidence for an association between gender nonconformity (GN) and depression. Of the papers reviewed, five related to childhood GN (CGN) and eight to adolescent or adult GN. Several papers found an association between GN and depression with some finding biological sex differences, where stronger associations were found in males. As eleven papers included a cross sectional design, it was not possible to infer causal relationships between study variables in these studies. Other methodological issues, such as limited sampling, and issues with measures of key variables, have been discussed.

Many studies also investigated the mechanisms which seek to explain the association between GN and depression, such as bullying victimisation, abuse and quality of parental relationships and several significant associations with GN and depression were identified. This review highlighted the need for more research, including the requirement for more longitudinal studies, UK based studies, studies in non-westernised societies and those containing robust, reliable and valid measures. There was only one study which investigated CGN and childhood depression and this highlights the need for further studies in this area. Community, public health and school-based interventions will help improve awareness and tackle GN stigmatisation. Additionally, mental health services should be aware of the heightened distress that GN individuals often experience and ensure that this is taken into account when developing individual interventions. This includes parental support.

In Section B of the MRP the research question was related to whether childhood gender nonconformity (CGN) was associated with childhood depression and whether this association was mediated by self-esteem. While other studies have found associations between CGN and depression in adolescents and adults, only one considered childhood depression (Petterson et al., 2016). The current study accesses longitudinal data to investigate whether

self-esteem mediates the association between early CGN and childhood depression. The dataset was obtained from the Avon Longitudinal Study of Parents and Children. As there was a large proportion of missing data, multiple imputation of missing variables was performed. However, comparisons between the original data and imputed data revealed negligible differences to the analysis. Therefore the analysis was performed on the original data as levels of missing data were too high to ensure validity of the process.

A mediation analysis (Hayes, 2012) showed that self-esteem mediated the association between CGN and childhood depressive symptoms in females. However the effect size was very small. In males the mediation model demonstrated an association between CGN and depressive symptoms which approached significance but this was not mediated by self-esteem. Again, the effect size was very small. Study limitations were discussed, such as high levels of missing data. The study highlighted a need for further UK based, preferably longitudinal studies exploring the mechanisms which explain the association between CGN and childhood depression. This should help to develop effective interventions for this client group. Additionally, community psychology, public health and school interventions have been discussed.