**Abstract**

In this commentary, the author reviews methodological and conceptual shortcomings of recent articles by K. D. Drummond, S. J. Bradley, M. Peterson-Badali, and K. J. Zucker (2008) as well as G. Rieger, J. A. W. Linsenmeier, L. Gygax, and J. M. Bailey (2008), which sought to predict adult sexual identity from childhood gender identity. The author argues that such research needs to incorporate a greater awareness of how stigmatization affects identity processes. Multidimensional models of gender identity that describe variation in children's responses to pressure to conform to gender norms are particularly useful in this regard (S. K. Egan & D. G. Perry, 2001). Experiments on the interpretation of developmental data are reviewed to evidence how cultural assumptions about sexuality can impact theories of sexual identity development in unintended ways. The author concludes that understanding the development of children presumed most likely to grow up with sexual minority identities requires a consideration of the cultural contexts in which identities develop and in which psychologists theorize.

**Keywords:** gender identity; gender identity disorder in childhood; lesbian, gay, bisexual, and transgender (LGBT) psychology; stigma

A recent special section of *Developmental Psychology* (Vol. 44, pp. 1–138) illustrates that developmental interest in sexual identity has “never been greater than it is today” (Patterson, 2008, p. 1). Over the last 35 years, the shift from a disease paradigm to a stigma paradigm in psychology has increased psychologists' understanding of the external risk factors and the internal strengths that shape sexual minority identity development in social contexts that are characterized by prejudice (Meyer, 2003). That paradigm shift involved the recognition of early nonpathologizing research (e.g., Hooker, 1957), the emergence of the modern lesbian and gay rights movement, the 1973 vote by the members of the American Psychiatric Association to de-pathologize homosexuality, and the subsequent resolution by the American Psychological Association for “all mental health professionals to take the lead in removing the stigma of mental illness that has long been associated with homosexual orientations” (Conger, 1975, p. 633; also see Bayer, 1981). As in any scientific revolution, old questions were abandoned for newer ones (Kuhn, 1970). Interest in the clinical assessment, psychological adjustment, and causes of homosexuality waned, and attention focused on many aspects of lesbian, gay, and bisexual lives that the disease paradigm had obscured, including stigmatization processes and adjustment to heterosexist societies (see Morin, 1977; Watters, 1986).

For the purposes of this article, I describe research that attributes particular features of lesbian, gay, bisexual, and transgender (LGBT) development to the influence of external stigmatization processes rather than inherent pathological factors as operating within a stigma paradigm. The shift to this paradigm has been incomplete; *ego-dystonic homosexuality* remained in the *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.; *DSM–III*; American Psychiatric Association, 1980), and *sexual disorder not otherwise specified* in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV*; American Psychiatric Association, 1994) still affords the diagnosis of unwanted homosexuality (Zucker & Spitzer, 2005). Moreover, cultural values always influence science, even in the stigma paradigm (e.g., Hegarty & Massey, 2006; Kitzinger, 2004) in which some groups have been foregrounded more than others. For example, research in the stigma paradigm has focused on men (Lee & Crawford, 2007), making recent models that center on
women's sexuality particularly welcome (e.g., Diamond, 2008; Thompson & Morgan, 2008). The stigma paradigm has also been applied primarily to adults and only rarely to children. The special section includes a study of the effects of stigmatization on adolescents who question their heterosexuality (Bos, Sandfort, de-Bruyn, & Hakvoort, 2008), but children who break gender norms were studied with the goal of predicting their adult sexual and gender identities and with little regard to the stigmatization that those identities might attract (Drummond, Bradley, Peterson-Badali, & Zucker, 2008; Rieger, Linsenmeier, Gygax, & Bailey, 2008).

These last two articles form part of a research tradition that has been little informed by the stigma paradigm. I call this paradigm the gender identity disorder in childhood (GIDC) paradigm. In an earlier meta-analysis, the senior authors of Drummond et al.'s (2008) and Rieger et al.'s (2008) articles examined how adults of diverse sexual identities recalled the gender conformity of their childhood experiences (Bailey & Zucker, 1995). In the earliest and largest retrospective study [page 896] that defined this paradigm, straight [Note1] adults reported more conformity than their sexual minority counterparts, and this effect was strongest among men and Whites (Bell, Weinberg, & Hammersmith, 1981). Bailey and Zucker (1995) similarly concluded that adult sexual identity predicted recall of childhood gender best among men, but they did not examine whether this trend varied across ethnic groups. Both Rieger et al. and Drummond et al. presented work within this tradition, and the coexistence of the GIDC and stigma paradigms exemplifies how multiple paradigms can operate concurrently in psychology (Driver-Linn, 2003).

Why do these paradigms remain separate? First, affirmative psychologists have only rarely examined how sexual identity based stigma affects children who break gender norms (e.g., Hill & Willoughby, 2005; Martin, 1995). Second, both Bell et al.'s (1981) study and the canonical prospective study linking childhood gender role to adult sexual identity (Green, 1987) were initiated prior to the de-pathologizing of homosexuality, but both were published some years later. Third, lesbian and gay activists who protested the diagnosis of adult homosexuality were absent from the later debates that shaped the diagnosis of GIDC, which defines some children as mentally ill as a result of their gendered behaviour (Bryant, 2006). Fourth, advocates of the GIDC paradigm have often dismissed their many critics who argue that gender is, at least in part, socially constructed and that gender variance is no grounds for a psychiatric diagnosis—often without regard for the merit in their critiques (see, e.g., Bradley & Zucker, 1998; Zucker, 1997; Zucker & Spitzer, 2005).

Sandford (2005) further noted that researchers in the stigma paradigm are often averse to research that emerges from retrospective and prospective studies linking children who break gender norms with sexual minority adults (see, e.g., Hill, Rozanski, Carfagnini, & Willoughby, 2005; Minter, 1999; Paul, 1993). By closely attending to the recent articles by Drummond et al. (2008) and Rieger et al. (2008), I hope to explain such aversion and to suggest how the stigma paradigm could assist the understanding of the development of those children who—according to the scientists who study them most determinedly—are most likely to be grow up lesbian, gay, bisexual, or transgender. A switch away from conceptualizing gender as essentially a masculinity–femininity (MF) personality dimension, toward multidimensional models of gender identity that allow for individual variation in the interpretation of gender-related behaviors, will be key to this paradigm shift (Egan & Perry, 2001). Like all arguments about paradigm shifts, mine assumes that different interpretations can be made of the same data depending on the prevailing theoretical framework. Accordingly, I also return to Hegarty’s (1999) experiments to show how unacknowledged cultural knowledge affects the explanation of data about the development of sexual identity.

Rieger et al. (2008): A Critique
Childhood gender nonconformity is recalled by only some adults with sexual minority identities (e.g., Diamond, 1998; Savin-Williams, 1998), and many straight women recall
tomboy pasts (Hyde, Rosenberg, & Behrman, 1977). Retrospective studies in this domain are vulnerable to the obvious criticism that people systematically misremember the past (see, e.g., Gottschalk, 2003; Paul, 1993; Ross, 1980). As a clever response to this problem, Rieger et al. (2008) asked raters to watch video clips of lesbian, gay, and straight adults as children and as adults. Participants completed self-report measures of MF both for their current adult selves and for their past childhood selves. Straight participants judged themselves, and were judged by others, to be more gender conforming than their lesbian and gay counterparts, both as adults and as children. Rieger et al. presented these results as supporting earlier retrospective studies (e.g., Bailey & Zucker, 1995) and interpreted positive correlations between MF measures as evidence that there is an “essence” to this dimension that varies systematically across sexual identity groups (p. 54).

This last conclusion is overstated. For the childhood period, gay men’s self-reported MF was significantly correlated with observers’ ratings ($p < .05$), but the scores of all other groups were not (all $ps > .05$). For adulthood, only straight men’s self-reported and observer-rated MF scores were significantly correlated ($p < .05$). Those of all other groups were not (all $ps > .05$). Self-reported adult MF and observer-rated childhood MF were correlated for men only. To support the argument that these moderate relationships between self-report and observer ratings evidence an MF personality dimension, Rieger et al. (2008) cited Pedhazur and Tetenbaum (1979). However, those authors also found only a weak relationship between self-reported masculinity, femininity, and other gender-related traits.

Rieger et al. (2008) evidenced an awareness that their correlational results fell short of support for a one-dimensional MF model when they attributed the nonsignificant correlations between the MF scores to low statistical power. They also averaged correlations between all possible MF scores, and when these averaged correlations did not differ significantly by group, they were interpreted as further evidence that the one-dimensional MF model applied equally well across all groups. However, averaged correlations between MF scores were somewhat higher among straight and gay men ($r_s = .46$ and .32, respectively) than among lesbian and straight women ($r_s = .24$ and .25, respectively). Tellingly, the authors did not conjecture whether a more powerful study might have yielded significant gender differences in the strength of these correlations. Such a finding would, of course, suggest that the MF model works best among men (see Rieger et al., 2008, p. 52).

Although the authors’ interpretations of their correlations between MF scores are unusual, such moderate correlations are precisely what multidimensional models of gender identity would predict (e.g., Egan & Perry, 2001; Spence, 1985). MF models and androgyny measures (e.g., Bem, 1996) share the assumption that the masculinity and femininity of traits and behaviors are relatively fixed. In contrast, multidimensional models assume that people use gendered behavior to form subjective impressions of MF in variable ways, leading to moderate, variable relationships between MF ratings and specific behaviors. Longitudinal research in which multidimensional models were used suggests how the kind of retrospective studies reviewed by Bailey and Zucker (1995) might be improved by stepping outside of the MF model. For example, Carver, Egan, and Perry (2004) found that children who question their heterosexuality engage in fewer same-sex activities than their peers, but they do not engage in a greater number of opposite-sex activities. Such children also come to see themselves as more atypically gendered as a consequence of questioning their heterosexuality. Researchers who seek to understand when people correctly or falsely remember the gendered behaviors of their childhood would do well to study such identity processes.

Ironically, Rieger et al.’s (2008) new methodology casts doubt on the earlier retrospective studies reviewed by Bailey and Zucker (1995): only gay male adults, it seems, can be trusted to recall their childhoods with an accuracy approaching that of videotape. Rieger et al.’s results do not fit an MF model, but they do fit a multidimensional model that assumes
that people perceive others’ genders quite differently from the ways that they perceive their own (see also Spence, 1985). More generally, the items used by Rieger et al. routinely constructed masculinity and femininity as opposites. For example, the observers rated the videos using a scale anchored by the terms *masculinity* and *femininity*, and self-report items presented masculine and feminine behaviors as mutually exclusive (e.g., “As a child I preferred playing with girls than with boys”). As a result, Rieger et al. assumed, but could not test, that masculinity and femininity are persistently perceived as logical opposites by ordinary people.

**Drummond et al. (2008): A Critique**

Concerns about memory distortion were less pressing for Drummond et al. (2008); these authors used a longitudinal design to study the link between childhood gender identity and adult sexual and gender identity prospectively. Unlike earlier, larger prospective studies of boys (e.g., Green, 1987), Drummond et al. asked whether girls labeled with GIDC would take on sexual and gender minority identities as adults. Few did, but more did than population estimates would predict. These authors accessed the files of the world’s largest clinic for children diagnosed with GIDC. Only 71 girls altogether were referred to this clinic for GIDC between 1975 and 2004, and only 36 of these were above the cutoff of 17 years required for study participation. Another participant was contacted opportunistically. Of these 37 eligible people, only 25 were contactable, and only 15 had met the threshold for a GIDC diagnosis in childhood. The difficulties surpassed in constructing this sample show how firmly the GIDC paradigm is focused on boys, and how rarely it is applied to girls; GIDC is largely a diagnosis of atypical femininity and only rarely a diagnosis of atypical masculinity.

Drummond et al. (2008) and Rieger et al. (2008) share the assumption that the clinical literature on GIDC can serve as a model for the development of minority sexual identity more generally. Drummond et al. noted that their data show “at least some convergence with data from retrospective studies” (pp. 42–43), whereas Rieger et al. motivated their study with the question of “how much the findings” of prospective studies “generalize to the development of most homosexual people” (p. 47). LGBT affirmative theorists’ reluctance to draw on this work (see Sandford, 2005) may be due to its reliance on medicalizing language when talking about sexual and gender minorities. Drummond et al.’s study fails to take the lead “in removing the stigma of mental illness that has long been associated with homosexual orientations” (Conger, 1975, p. 633). Instead, Drummond et al. conflate statistically common gender patterns with normativity in their first sentence (p. 34); describe “girls with potential problems in their sexual identity development” prior to describing those girls’ likelihood of growing up lesbian, bisexual, or transgender (p. 42); describe the breaking of gender norms as a *risk factor* for adult transgender status (p. 42); and note that a high “dosage” of gender transgression is likely to cause bisexual and lesbian identities (p. 43). Rieger et al. did not use such medicalized language, but by citing such GIDC work as the framework they wished to extend to understand sexual identity development more generally, their work might similarly augment stigma against LGBT people rather than attenuate it. Of course, work within the GIDC paradigm is also in tension with the increasing number of emerging analyses of the stigma experienced by transgender adults (e.g., Clements-Nolle, Marx, & Katz, 2006; Grossman & D’Augelli, 2007; Hill & Willoughby, 2005; Lombardi, Wilchins, Priesing, & Malouf, 2001; Tee & Hegarty, 2006).

Again, multidimensional models of gender identity suggest where useful alternatives to the GIDC paradigm might lie. Egan and Perry’s (2001) model predicts that children who break gender norms are not *intrinsically* maladjusted, as the psychiatric diagnosis of GIDC presumes. Indeed, among these children, those that feel less pressure to conform to gender roles fare the best (Yunger, Carver, & Perry, 2004). The diagnosis of GIDC assumes that children experience distress, but negative reactions to children’s gender nonconformity magnify that distress as the multidimensional model would predict (Cohen-Kettenis, Owen, Kaijser, Bradley, & Zucker, 2003; Zucker & Bradley, 1995, pp. 108–110). As Lev (2005) has
noted, to the extent that children’s feelings of distress come from external sources, the GIDC diagnosis is inconsistent with DSM criteria that specify that psychiatric distress must result from inherent pathology rather than reactions to social deviance (see also Hird, 2003; Wilson, Griffin, & Wren, 2002). Indeed, attributing GIDC to a child, or delivering therapy aimed at reinforcing conformity to gender norms (e.g., Green, 1987; Zucker & Bradley, 1995), may augment the pressure children feel to conform to gender norms. As such, the GIDC model may undermine, rather than support, that child’s well-being, as critics of the diagnosis have long argued (e.g., Sedgwick, 1991).

Stigma researchers may also be averse to GIDC work because more recent psychiatric alternatives acknowledge the possibility that acceptance of children’s atypical gender behavior might promote their well-being. Menvielle and Tuerk (2002) have interpreted evidence of developmental links between childhood gender conformity and adult sexual identity as an ethical imperative to assist parents to develop less heterosexist and more gender inclusive scripts for family life. Their interventions assume that all children, particularly boys, are vulnerable to homophobia and that acceptance of gender variant behaviors supports well-being among gender nonconforming children (cf. Carver et al., 2004; Yunger et al., 2004). Consistent with research on the effectiveness of lesbian and gay parents (e.g., Patterson, 2006) and the experiences recounted by their children (Goldberg, 2007; Saffron, 1998), Menvielle and Tuerk have recognized that adherence to gender norms is not essential to familial well-being. Finally, the intervention does not posit an “essence” to MF but recognizes that gender nonconformity may persist or desist and may lead to gay/lesbian, bisexual, or straight adult identities (cf. Carver et al., 2004). In short, the intervention is interested in the child’s well-being, understands familial bonds to be key to that well-being, and disattends to anxious questions about the child’s eventual gender[page 898] identity or sexual identity on the basis of his or her current gender nonconformity.

Who Needs To Be Explained?
In arguing for a paradigm shift in both research and clinical practice around children who break gender norms, I have assumed Kuhn’s (1970) premise throughout, that data about children can be interpreted differently within different paradigms and that such paradigms can be shaped by larger cultural contexts. For this reason, it was particularly disappointing that Drummond et al. (2008, p. 35) cited my unpublished dissertation (Hegarty, 1999) as an argument that “the sex-typed behavior-sexual orientation association is nothing more than participants recalling behaviors that adhere to cultural stereotypes and expectations.” Contrary to this suggestion, Hegarty (1999) reported inconsistent results about effects of cultural expectations on straight students’ memories (see Hegarty, 2001, for a discussion). However, published experiments from that dissertation (Hegarty & Pratto, 2001; see also Hegarty & Pratto, 2004) showed the effects of cultural meaning systems on data interpretation with regard to retrospective recall studies.

In those experiments, undergraduates explained, in their own words, the results of retrospective studies linking adult sexual identity to childhood gendered behavior. Explanations overwhelmingly took lesbian and gay development as the effect to be explained, and heterosexual development as the implicit norm, even when straight adults were described as recalling childhood gender transgressions. Only when explicitly instructed to do so did participants write explanations that focused explicitly on straight people. These experiments evidence heteronormative thinking (Butler, 1990; Warner, 1993)—the tendency to conflate heterosexuality with the way people universally are and will universally remain. Similar studies show that men and White people are taken as the implicit norm when gender and race differences are explained (Hegarty & Buechel, 2006; Miller, Taylor, & Buck, 1991; Pratto, Hegarty, & Korchmarios, 2007).

Spontaneous explanations of retrospective studies were further shaped by stereotypes (Hegarty & Pratto, 2001, 2004). When participants were told that gay men or lesbians
recalled childhood nonconformity, participants’ explanations largely assumed that childhood behavior had been accurately recalled. However, even when straight adults were said to recall gender nonconformity, explanations remained focused on lesbians and gay men and now specified that gay men or lesbians had lied, misremembered their past, or otherwise reconstructed their past experiences. Hegarty (1999) did not present evidence that culture impacted the way that sexual minority people recall their childhoods as Drummond et al. (2008) suggested (but see Carver et al., 2004). Rather the published studies from that dissertation clearly show that cultural stereotypes and expectations lead people to invoke the question of memory bias selectively when they are presented with results that do not conform to their stereotypes.

These asymmetric patterns of explanation are important because many authors delineate good from bad developmental science on sexual identity on the basis of whether research “treats heterosexuality, bisexuality and homosexuality as equally requiring explanation” or “seeks a cause only for homosexuality and omits heterosexuality from the investigation” (Herek, Kimmel, Amaro, & Melton, 1991, p. 958; for similar statements, see Bem, 1996; Freud, 1905; Zucker & Bradley, 1995). However, heteronormativity can affect scientific explanations in spite of our best intentions (Hegarty, 2007). Bell et al. (1981) similarly argued that “any parsimonious and valid scientific explanation must account for the rule as well as the exception” and that “any satisfactory theory of sexual orientation must explain with equal ease both the majority pattern, heterosexuality, and the less-common homosexual pattern” (p. xi). However, the explanations of group differences in their canonical book focused on sexual minority groups more than sexual majority groups by a ratio of 3:1—much like those of experimental participants (see Hegarty & Pratto, 2001). Bell et al. are not unusual in this regard. Heteronormativity also frames developmental explanations of sexual identity in popular journalism (Hegarty, 2003) and psychology textbooks (Barker, 2007).

Conclusion
Explanations of developmental processes that focus selectively on lesbians, gay men, or bisexual people are not just influenced by the larger heteronormative culture. Rather, as developmental psychologists continue to theorize sexual-identity-related differences in ways that favor heterosexuality (Savin-Williams, 2008), those explanations also influence and reinforce that culture. Questions about the causes of homosexuality are not “closely associated with technologies that have attempted to correct or prevent the development of the ‘condition’” as they once were (Morin, 1977, p. 633). However, heterosexual development is all too easily taken to be the implicit benchmark for sexual development rather than one-among-many sexual identities that healthy people grow into. Like the children of LGBT parents, some developmental psychologists are increasingly recognizing that widening the range of gender roles that are taken as normal may have benefits for all (see, e.g., Eisenberg, Martin, & Fabes, 1996). This recognition implies that any greater gender role flexibility afforded by LGBT identities and communities might become a valued resource for healthy development rather than an anomaly to explain, a “dosage effect”, or a natural experiment to explore whether gender is essentially tied to sexual identity. We will not get the measure of the developmental relationships between sexuality, gender, and well-being by repetitively deploying an MF thermometer to predict adult sexual identity as an outcome variable. Instead, we need to see growing people as actively managing relationships between their behavioral preferences, emergent sexual identities, and the double bind between performing gender in accord with heterosexual norms and risking the consequences of transgression. Ironically, even developmental psychologists have a better understanding of how these processes work among adults than among children at this point in time.

Note 1
I use the term “straight” rather than “heterosexual” for the same reason that I use “gay” and “lesbian” rather than “homosexual”; it is a term derived from sexual minority communities and not a medicalizing term.

References


Watters, A. T. (1986). Heterosexual bias in psychological research on lesbianism and male homosexuality (1979–1983), utilizing the bibliographic and taxonomic system of


