

Comments on Ethical Reporting and Interpretations of Findings in Hsu, Rosenthal, and Bailey's (2014) "The Psychometric Structure of Items Assessing Autogynephilia"

Jaimie F. Veale

Stigma and Resilience Among Vulnerable Youth Centre, School of Nursing, Faculty of Applied Science, University of British Columbia, Vancouver, Canada.

In their recent article in this journal, Hsu, Rosenthal, and Bailey (2014) reported the psychometric properties of a new measure of autogynephilia they had developed. Hsu et al. reported they had "assembled 22 items to assess the five types of autogynephilia that have been reported in the literature (Blanchard, 1991)." They conducted a factor analysis on these items to establish subscales. They then further assessed the validity of the scale and its subscales by assessing their relationships to other theoretically related variables. In this Letter, I call attention to the language used by Hsu et al. and how this fails to meet guidelines for scientific reporting using accurate, unbiased language that does not reinforce demeaning perspectives about groups being studied. Secondly, I interpret some of the findings in the article and reflect on the implications of these findings for Blanchard's (1991) autogynephilia theory and for transgender and gender-diverse people in the current political environment.

Misgendering Language

Hsu et al.'s (2014) language use does not meet current standards for unbiased reporting. They began their article by defining autogynephilia as "a man's paraphilic tendency to be sexually aroused by the thought or image of himself as a woman." Here, and throughout their article, they used masculine nouns and pronouns when referring to people who have these sexual attractions. The *Publication Manual of the American Psychological Association* (American Psychological Association, 2010), the style guide used by this journal, suggested that authors "respect people's preferences; call people what they prefer to be called" (p. 72). The manual also specifically referred to transgender people, suggesting "using words (proper nouns, pronouns, etc.) appropriate to the person's gender identity or gender expression, regardless of birth sex. For example, use the pronouns *he*, *him*, or *his* in reference to a female-to-male transgender person" (p. 74). The manual's supplemental material gave further guidance: "the nouns *woman* and *man* refer to gender identity or gender expression (e.g., a male-to-female transsexual can be referred to as a *biologic male* but should be called a *transsexual woman*, not a *transsexual man*)" (section 3.12).

Hsu et al.'s referring to their sample as men may be appropriate if this is their gender identity, but it's unclear from the article if this is the case. If this is so, then this would be the

first autogynephilia research that has not been conducted on transsexual women (a point which may have relevance to interpreting the study's findings, but Hsu et al. do not discuss this). By using the term *men* to refer to participants in previous studies and all people who they define as being autogynephilic, Hsu et al. consistently *misgendered* these people. That is, they used a gender pronoun that did not reflect gender identity of the people who they were referring to ("Misgender," n.d.).

Interpretations of the Findings

Assumptions About Autogynephilic Interests Being Limited to People Assigned Male at Birth

The idea that autogynephilic sexual attraction is something limited to people assigned male at birth is at best questionable given studies that have assessed this among cisgender women have found a notable level of affirmative responding (Moser, 2009; Veale, Clarke, & Lomax, 2008). The findings of Nuttbrock, Bockting, Rosenblum, Mason, and Hwahn (2011b) that cross-dressing sexual arousal is associated with demographic factors and more private than public feminine dressing indicate there may be a psychological process in which these dissonant feminine interests become sexualized (see also Veale, Lomax, & Clarke, 2010). Alternative explanations for these phenomena have been proposed by a supporter of Blanchard's theory. Lawrence (2009, 2010, 2011) suggested that the affirmative responding to autogynephilia measures in cisgender women could be due to something other than autogynephilia and that the relationship between demographic factors and cross-dressing sexual arousal is caused by both variables being correlated with sexual attraction to women.

Hsu et al. (2014) seem to work on the assumption that autogynephilia is limited to birth-assigned males and the specific autogynephilic sexual interests are of unusual or even exceptional things. Hsu et al. found factor analytic evidence for the five manifestations of autogynephilia that Blanchard (1991) proposed based on the items they created to assess these five manifestations. The manifestations are interpersonal, anatomic, transvestic (dressing as a woman), physiologic, and behavioral autogynephilia. Hsu et al. noted that some of the manifestations of autogynephilia are "seemingly trivial or mundane feminine activities."

It would be worthwhile and interesting to look at whether specific interests related to enjoyment of appearing and behaving

femininely (whether these be sexual or nonsexual interests) actually differ between feminine people of different gender identities, birth-assigned sexes, and between those who report sexual arousal associated with it. There may actually be nothing exceptional or mundane about these types of interest amongst feminine people, as Hsu et al. seem to assume.

Correlations Between Items Measuring the Same Thing Interpreted as Positive Evidence

Hsu et al.'s (2014) measure of Interpersonal Autogynephilia was the only scale to predict non-heterosexual identity and number of male sexual partners. Hsu et al. claimed that this was consistent with a component of Blanchard's (1989b) theory, that sexual attraction to males in people with autogynephilia is a result of *pseudobisexuality*, an autogynephilic desire to be attractive to men, rather than an attraction to men's bodies. However, a much more parsimonious explanation is that part of this scale is actually measuring sexual interest in men. A large amount of the variation in two of the scale's four items, which ask how sexually arousing participants find *having a man take me out for a romantic evening* and *picuring myself as a woman having sex with a man*, could be attributed simply to sexual interest in men. Rather than replicating Blanchard's finding that bisexuals were more likely to be attracted to being admired as a women, Hsu et al. appear to have shifted the goalposts and lowered the bar significantly in arguing that this finding supports Blanchard's theory.

Similarly, Hsu et al. (2014) noted that their measure of transvestic autogynephilia was uniquely related to fetishism as it was the only autogynephilia subscale to significantly predict fetishism in a regression model. This finding is again likely to be due to overlap in item content. The Transvestic Autogynephilia subscale measures sexual arousal to wearing makeup, perfume, dresses, heels, women's underwear, and getting a women's hairstyle and the fetishism measure used was a single item measuring sexual arousal to "some nonhuman objects like shoes, rubber, latex, clothing, strap-ons, etc." There is clearly enough similarity in the items that it would be surprising if there was no relationship between these two variables.

Failure to Separate Diverse Sexual Interests

A "control group" of heterosexual men who had never cross-dressed was also used for comparison on a number of measures, including a scale they developed and labeled the *Paraphilic Interests Scale*. This scale was described as measuring a variety of sexual interests as a single construct: exhibitionism, fetishism, voyeurism, frotteurism, masochism, sadism, and transvestic fetishism. While Hsu et al. did not give any information about the validity of this scale or test this purported unidimensionality, they noted that it had an internal consistency of .68 among the sample they labeled autogynephilic. This level of internal consistency is generally considered questionably low (Nunnally & Bernstein, 1994). While the study found a significant difference between the heterosexual men and participants called autogynephilic on the Paraphilic Interests Scale, this difference had a small effect size. At this level of effect size, only around 59% of the sample called autogynephilic would score higher than the average score of the group labeled the control group. The

reason behind this group difference on the Paraphilic Interests Scale is likely to be because one of the items on the scale measured transvestic fetishism. Indeed, Hsu et al. reported a group difference for this item with a very large effect size. To a lesser extent, this group difference could also be due to the fetishism item asking about shoes and clothing. As I noted above, one would expect those who report sexual attraction to wearing women's clothing to also endorse this item.

It seems reasonable to conclude that the two groups did not differ on the other paraphilias measured as part of the same scale that involved sexual interest in non-consensual activity: exhibitionism, voyeurism, and frotteurism. If this is so, this finding was unresponsive of Blanchard's theory which suggests these varied paraphilic interests should be more common among those reporting autogynephilia (Freund & Blanchard, 1993).

This also has implications for those pushing for rights for transgender people. While previously, research using clinical and forensic samples noted a co-occurrence of transvestism with other paraphilias (e.g., Freund & Watson, 1990), it may be that studies using community-based samples do not find such a co-occurrence. This would mirror the history of homosexuality, previously thought to be associated with paraphilias (American Psychiatric Association, 1968) and found to be co-occurring with paraphilias in research using non-community-based samples (e.g., Chalkley & Powell, 1983; Whitener & Nikelly, 1964). A finding that those who report what Hsu et al. define as autogynephilic sexual arousal are no more likely to report sexual attractions to non-consensual paraphilias would be particularly important to note given ongoing arguments by opponents of transgender nondiscrimination laws. These laws can allow transgender people access to bathroom and change room spaces that align with their gender identity but opponents of them argue that they would allow supposedly predatory transgender women access to women's spaces (e.g. "Transwomen & Sexualized Violence," n.d.). Given arguments such as these were likely to be born from research using clinical and forensic samples and the ethical obligations that researchers have when conducting research on marginalized and vulnerable groups to ensure that their findings are not misrepresented or misused in a way that can cause harm to the group being researched (Connolly, 2003), it looks as though Hsu et al. (2014) missed the opportunity to anticipate and forestall these negative consequences of their research.

Further Findings That Were Contrary to Blanchard's Theory

Anatomic autogynephilia was unrelated to gender dysphoria

Hsu et al. (2014) noted that some of their findings were contrary to Blanchard's autogynephilia theory. The most significant of these was that their measure of anatomic autogynephilia was not significantly related to gender dysphoria. This is in direct contradiction with a central tenet of Blanchard's theory, that anatomic autogynephilia is the main cause of gender dysphoria in this population (Blanchard, 1993). To explain this finding, Hsu et al. suggested that it may be necessary to assess anatomic autogynephilia relative to other types of autogynephilia in order to generate more meaningful

conclusions.” Indeed, this is what Blanchard (1993) originally did.

Hsu et al. (2014) had the data to do this and, upon request, Hsu kindly conducted this analysis for me (personal communication, February 25, 2015). This analysis showed that there were also no significant differences in the Pure Gender Dysphoria Scale score between those who endorsed picturing oneself as a nude woman ($M = 5.41$), a partially clothed woman ($M = 5.04$), or a fully clothed woman ($M = 5.43$) as most strongly associated with sexual arousal, $F(2, 140) < 1$. This represents a failure to replicate Blanchard's (1993) findings.

Another potential explanation for the inconsistency between Hsu et al.'s (2014) finding and Blanchard's (1993) finding that Hsu suggested in the personal communication (February 25, 2015) is that Blanchard's sample was more gender dysphoric. However, Hsu et al.'s sample still reported notable gender dysphoria ($M = 5.36$ on a scale of 0 to 9) and although Blanchard's sample had a higher average score (6.81),¹ 28% of Hsu et al.'s sample still scored higher than the mean of Blanchard's sample.²

Lack of asexual participants

Also in contradiction to Blanchard's theory was the low proportion of participants in Hsu et al.'s (2014) study who identified as asexual. Hsu et al. did not reflect this point, but the number of participants who identified as asexual (1 out of 149) was even fewer than the number who identified as homosexual (four)—even though by definition the latter group are not autogynephilic. In contrast, Blanchard (1989a) suggested and found evidence that asexuals (but not those who Blanchard called homosexual) are one of the three sexual orientation categories that are autogynephilic (the other two being bisexual and those who Blanchard calls heterosexual). Blanchard (1989a) proposed that asexuals (or, as he called them, analloerotics) “represent those cases in which the autogynephilic disorder nullifies or overshadows any erotic attraction to women” (p. 324).

Hsu et al.'s (2014) lack of asexual participants was in accordance with other findings of transsexual women using nonclinical samples that the prevalence of asexuality is lower than that found in clinical samples (Nuttbrock et al., 2011a) and this might even be no different to the prevalence of asexuality in the general population. An advantage of studies conducted on community samples such as Hsu et al.'s is less risk of response bias, as participants do not perceive their responses could impact their access to gender-affirming medical care. Studies like Blanchard's may be at risk of participants being more likely to respond in a way they think the researcher/clinician/gatekeeper wants to hear (e.g., responding as asexual).

¹ Calculated by combining the means of the subgroups reported by Blanchard (1993): nude $M = 7.8$, $n = 94$; partially clothed $M = 6.8$, $n = 67$; and fully clothed $M = 5.6$, $n = 77$.

² The overall mean score for Blanchard's sample corresponds to $z = 0.57$ in Hsu et al.'s sample. The percentage figure was calculated from obtaining the area under the normal curve. While this assumes a normal distribution for the Pure Gender Dysphoria scale in Hsu et al.'s sample, Hsu et al. did not note that this scale was skewed, but they did note that other scales were skewed, so this seems like a reasonable assumption.

Conclusion

Despite concerns being raised about the controversial nature of Blanchard's autogynephilia theory (Dreger, 2008), Hsu et al. (2014) have missed an opportunity to improve the standing of the scientific community. Researchers studying transgender and gender-diverse people have an obligation to use accurate and unbiased language to refer to the group they are studying and to report their findings in a way that cannot be misused in a way that can perpetuate negative consequences to the group.

References

- American Psychiatric Association. (1968). *Diagnostic and statistical manual of mental disorders* (2nd ed.). Washington, DC: Author.
- American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.
- Blanchard, R. (1989a). The classification and labeling of nonhomosexual gender dysphorias. *Archives of Sexual Behavior*, 18, 315-334.
- Blanchard, R. (1989b). The concept of autogynephilia and the typology of male gender dysphoria. *Journal of Nervous and Mental Disease*, 177, 616-623.
- Blanchard, R. (1991). Clinical observations and systematic studies of autogynephilia. *Journal of Sex and Marital Therapy*, 17, 235-251.
- Blanchard, R. (1993). Varieties of autogynephilia and their relationship to gender dysphoria. *Archives of Sexual Behavior*, 22, 241-251.
- Chalkley, A. J., & Powell, G. E. (1983). The clinical description of forty-eight cases of sexual fetishism. *British Journal of Psychiatry*, 142, 292-295.
- Connolly, P. (2003). *Ethical principles for researching vulnerable groups*. Belfast, Northern Ireland: Office of the First Minister and Deputy First Minister. Retrieved from http://www.ed.ac.uk/polopoly_fs/1.71235!/fileManager/EthicsPr-Vulnerable.pdf
- Dreger, A. D. (2008). The controversy surrounding The Man Who Would Be Queen: A case history of the politics of science, identity, and sex in the internet age. *Archives of Sexual Behavior*, 37, 366-421.
- Freund, K., & Blanchard, R. (1993). Erotic target location errors in male gender dysphorics, paedophiles, and fetishists. *British Journal of Psychiatry*, 162, 558-563.
- Freund, K., & Watson, R. (1990). Mapping the boundaries of courtship disorder. *Journal of Sex Research*, 27, 589-606.
- Hsu, K. J., Rosenthal, A. M., & Bailey, J. M. (2014). The psychometric structure of items assessing autogynephilia. *Archives of Sexual Behavior*. doi:10.1007/s10508-014-0397-9
- Lawrence, A. A. (2009). Something resembling autogynephilia in women: Comment on Moser (2009). *Journal of Homosexuality*, 57, 1-4.
- Lawrence, A. A. (2010). A validation of Blanchard's typology: Comment on Nuttbrock et al. (2010). *Archives of Sexual Behavior*, 39, 1011-1015.
- Lawrence, A. A. (2011). Further validation of Blanchard's typology: A reply to Nuttbrock, Bockting, Rosenblum, Mason, and Hwang (2010). *Archives of Sexual Behavior*, 40, 1089-1091.
- Misgender. (n.d.). In *Oxford dictionary online*. Retrieved from http://www.oxforddictionaries.com/definition/american_english/misgender
- Moser, C. (2009). Autogynephilia in women. *Journal of Homosexuality*, 56, 539-547.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- Nuttbrock, L., Bockting, W., Mason, M., Hwang, S., Rosenblum, A., Macri, M., & Becker, J. (2011a). A further assessment of Blanchard's typology of homosexual versus non-homosexual or autogynephilic gender dysphoria. *Archives of Sexual Behavior*, 40, 247-257.
- Nuttbrock, L., Bockting, W., Rosenblum, A., Mason, M., & Hwang, S. (2011b). Sexual arousal associated with private as compared to

-
- public feminine dressing among male-to-female transgender persons: A further response to Lawrence (2011). *Archives of Sexual Behavior*, 40, 1093-1096.
- Transwomen & Sexualized Violence. (n.d.). Retrieved from <http://genderidentitywatch.com/transwomen-sexualized-violence/>
- Veale, J. F., Clarke, D. E., & Lomax, T. C. (2008). Sexuality of male-to-female transsexuals. *Archives of Sexual Behavior*, 36, 586-597.
- Veale, J. F., Lomax, T. C., & Clarke, D. E. (2010). The identity-defence model of gender-variant development. *International Journal of Transgenderism*, 12, 125-138.
- Whitener, R. W., & Nikelly, A. (1964). Sexual deviation in college students. *American Journal of Orthopsychiatry*, 34, 486-492.