

Experimenting with Credibility in Refugee Adjudication: Gaydar^{*}

Sean Rehaag[†] & Hilary Evans Cameron[‡]

Canada offers refugee protection to sexual minorities facing persecution abroad. While success rates for sexual minority refugee claims have generally been higher than the overall average at Canada's Immigration and Refugee Board, hundreds of such claims are nonetheless turned down each year. The most common reason for denying these claims is that assertions about the claimants' sexual orientations are determined not to be credible. Scholars have raised concerns about how such credibility determinations are made. This article contributes to the critical literature in this area by exploring sexual minority refugee claim credibility assessments through an experimental study involving simulated refugee determinations. The experiment focuses on whether a claimant's appearance affects the simulated adjudicator's credibility determinations and written reasons provided to justify those determinations.

^{*} This research was supported through funding provided by the Social Sciences and Humanities Research Council of Canada. The authors are grateful for the able research assistance provided by Julia Kennedy, Elise Mercier and Yadesha Satheaswaran and for feedback provided at early stages of the research by William Cox. The authors would also like to thank the students who participated in the research and the instructor who integrated the research into her teaching.

[†] Sean Rehaag is an Associate Professor at Osgoode Hall Law School and the Director of York University's Centre for Refugee Studies.

[‡] Hilary Evans Cameron is an Assistant Professor at Ryerson University's Faculty of Law. When this article was written, she was a Social Sciences and Humanities Research Council of Canada postdoctoral fellow at Osgoode Hall Law School and the Bora Laskin National Fellow in Human Rights Research.

Le Canada offre l'asile aux minorités sexuelles confrontées à la persécution à l'étranger. Bien que le taux de succès des demandes de statut de réfugié des personnes issues des minorités sexuelles soit en général plus élevé que celui de l'ensemble des personnes qui font une demande auprès de la Commission de l'immigration et du statut de réfugié du Canada, des centaines de demandes de personnes issues des minorités sexuelles sont néanmoins rejetées à chaque année. Le motif le plus commun de rejet de ces demandes est le manque de crédibilité de l'affirmation du ou de la demandeur(e) quant à son orientation sexuelle. Nombreux chercheur(e)s sont préoccupés par la façon dont sont tirées ces conclusions relatives à la crédibilité. Cet article contribue à la littérature critique dans ce domaine, en explorant la façon dont la crédibilité est évaluée dans le cadre de demandes de statut de réfugié de personnes issues des minorités sexuelles. La méthode choisie est une étude expérimentale impliquant une simulation du processus de détermination du statut de réfugié. Cette étude se concentre sur la question de savoir si l'apparence du ou de la demandeur(e) a un impact sur les conclusions relative à la crédibilité et les motifs écrits fournis pour justifier ces conclusions.

I. Introduction

A refugee adjudicator walks into a hearing room at Canada's Immigration and Refugee Board (IRB) in downtown Toronto. The night before she reviewed materials submitted by a refugee claimant who says he faces persecution due to his sexual orientation as a gay man. Based on the materials submitted, the adjudicator is persuaded that sexual minorities¹ face persecution in the claimant's home country. There is, to her mind, only one question: is the claimant, in fact, gay? The adjudicator sits down and gets her papers in order. With everything in place, she is ready to start the hearing. She looks up at the refugee claimant and thinks to herself: well, he looks straight – this is going to be a hard case.

This hypothetical scenario raises a troubling question: does a refugee claimant's appearance affect whether refugee adjudicators believe their claim that they face persecution due to their sexual orientation? Or, to put the point more bluntly, do refugee adjudicators rely on some form of "gaydar"² when assessing sexual minority refugee claims – whether they say so in their written reasons or not?

This article grapples with that question through a research experiment conducted with approximately three hundred first-year students in a Canadian law school who acted as simulated refugee adjudicators. The experiment's main finding is that stereotypes about physical appearance significantly influenced whether simulated adjudicators believed a refugee claimant's asserted sexual orientation. However, the research also found that simulated adjudicators rarely referred to these stereotypes in the written reasons they provided to justify their decisions. In other words, there was a disjunction between the factors that affected the outcomes of simulated

¹ This article uses the term "sexual minorities" to refer to people who self-identify as something other than straight, including gay, lesbian, bisexual or queer. The authors' rationale for the choice of this term follows the explanation put forward by Rehaag in "Patrolling the Borders of Sexual Orientation: Bisexual Refugee Claims in Canada" (2008) 53 McGill LJ 59 at 62–63 [references omitted]:

Throughout this article I use the terms *queer* and *sexual minorities* interchangeably to cover a wide range of sexual and gender identities that challenge heteronormativity. I choose to avoid the more conventional label "LGBT" (Lesbian, Gay, Bisexual, and Transgender-Transsexual) due to a concern that such language is unnecessarily restrictive. The terms "queer" and "sexual minorities" serve my present purposes because their boundaries are blurred and explicitly invite contestation. As such, they can accommodate unconventional sexual and gender identities beyond those of gays, lesbians, bisexuals, and trans persons. Admittedly, any choice of terminology in this area is controversial. I acknowledge that many members of what I am calling queer and sexual-minority communities object to my chosen terminology on the grounds that it may underplay the long-standing political efforts of gays and lesbians to establish visible and politically recognized subject positions. Though I share this concern, I believe the advantages of nonexclusionary terminology outweigh its costs.

² For the purposes of this paper, we adopt the understanding of "gaydar" as the myth that one can reliably infer someone's sexual orientation by applying stereotypes about physical appearance, mannerism, fashion choices and the like. William Cox et al, "Inferences About Sexual Orientation: The Roles of Stereotypes, Faces and the Gaydar Myth" (2016) 53:2 J Sex Research 157.

refugee determinations (i.e. decision making) and the reasons provided by decision makers (i.e. decision writing). This finding, along with other ancillary findings from the experiment, have important implications for refugee status determination policy, oversight and training – and for future research in this area.

The article begins by addressing the Canadian socio-legal context in which the refugee claims of sexual minorities are adjudicated. Next, the article describes the research design of the study, noting some challenges and limitations. After discussing the results, the article ends by considering policy implications of the research findings and suggests further avenues for research.

II. Conceptual Framework

Canada was among the first countries in the world to recognize that people who face persecution due to their sexual orientation are eligible for refugee protection.³ Since this was established by the Supreme Court in 1993,⁴ thousands of sexual minorities have been recognized as refugees in Canada.⁵ Success rates in sexual minority refugee claims are similar to or slightly higher than average in Canada's refugee determination system.⁶ From 2013 to 2015, for example, refugee claims involving sexual orientation succeeded 70.5% of the time, compared to 62.5% for all claim types.⁷ While this indicates that many sexual minority refugee claimants are able to access protection, it also means that hundreds of these claims are turned down each year.

Prior research has found that the most common reason provided by Canadian refugee adjudicators for refusing sexual minority refugee claims in recent years is that the claimants were not credible in establishing their sexual orientation.⁸ In other words, adjudicators denying refugee protection to this group of claimants tend to do so on the grounds that claimants are lying about their sexual orientation.

However, extensive research has revealed problems with how refugee

³ The earliest decisions granting refugee protection to sexual minority refugees in Canada date back to 1991. See Nicole LaViolette, "The Immutable Refugees: Sexual Orientation in *Canada (A.G.) v Ward*" (1997) 55:1 UT Fac L Rev 1 at 3, 15. See also Sharalyn Jordan & Chris Morrissey, "On What Grounds? LGBT Asylum Claims in Canada" (2013) 42 Forced Migration Rev 13 at 13.

⁴ *Canada (Attorney General) v Ward*, [1993] 2 SCR 689 at 739, SCJ No 74.

⁵ For example, from 2004 to 2008, Canada granted refugee protection in 1,527 cases that were categorized by the IRB as involving persecution on account of sexual orientation. Sean Rehaag, "Do Women Judges Really Make a Difference? An Empirical Analysis of Gender and Outcomes in Canadian Refugee Determinations" (2011) 23 CJWL 627 at 644.

⁶ *Ibid* (indicating that, from 2004 to 2008, 51.7% of sexual orientation claims succeeded, compared to 50.2% for all claims).

⁷ Sean Rehaag, "Sexual Orientation in Canada's Revised Refugee Determination System: An Empirical Snapshot" (2017) 29 CJWL 259 at 275 [Rehaag, "Revised System"].

⁸ *Ibid* at 284.

adjudicators assess credibility in this area.⁹ One concern relates to adjudicators assessing the credibility of evidence regarding an asserted sexual orientation against stereotypes about sexual minorities. These include stereotypes about physical appearance, such as the inaccurate belief that all gay men appear effeminate and that all lesbians appear masculine.¹⁰ Sexual minority refugee claimants whose appearances do not match these stereotypes risk receiving mistaken negative decisions. This is a serious problem because, by definition, false negative decisions may result in refugees being deported to countries where they face persecution, torture or even death, in contravention of international refugee law.

Partly in response to these sorts of concerns, Canada's IRB has issued guidelines to adjudicators deciding cases involving sexual orientation, gender identity and gender expression.¹¹ The guidelines ask adjudicators to avoid, among other things, drawing on stereotypes relating to "feminized or masculinized appearance or mannerisms."¹² The United Nations High Commissioner for Refugees has also issued a guideline addressing the topic:

[D]ecision makers need to maintain an objective approach so that they do not reach conclusions based on stereotypical, inaccurate or inappropriate perceptions of LGBTI individuals. The presence or absence of certain stereotypical behaviours or appearances should not be relied upon to conclude that an applicant possesses or does not possess a given sexual orientation or gender identity. There are no universal characteristics or qualities that typify LGBTI individuals any more than heterosexual individuals.¹³

While explicit reminders about avoiding these sorts of stereotypes are encouraging, they are hardly new. Refugee adjudicators have received such

⁹ See e.g. Laurie Berg & Jenni Millbank, "Constructing the Personal Narratives of Lesbian, Gay and Bisexual Asylum Claimants" (2009) 22:2 J Refugee Stud 195; Jenni Millbank, "From Discretion to Disbelief: Recent Trends in Refugee Determinations on the Basis of Sexual Orientation in Australia and the United Kingdom" (2009) 13 Intl JHR [Millbank, "Discretion"]; Catherine Dauvergne & Jenni Millbank, "Burdened by Proof: How the Australian Refugee Tribunal has Failed Lesbian and Gay Asylum Seekers" (2003) 31:2 Fam L Rev 299; David Murray, "To Feel the Truth: Discourse and Emotion in Canadian Sexual Orientation Refugee Hearings" (2014) 3:1 J Language & Sexuality 6; Jenni Millbank, "Imagining Otherness: Refugee Claims on the Basis of Sexuality in Canada and Australia" (2002) 26:1 Melbourne UL Rev 144; Katherine Fobear, "Queer Settlers: Questioning Settler Colonialism in LGBT Asylum Processes in Canada" (2014) 30 Refuge 47 at 52-54; Thomas Spijkerboer, ed, *Fleeing Homophobia: Sexual Orientation, Gender Identity and Asylum* (New York: Routledge, 2013) 1 at 14-16.

¹⁰ Jenni Millbank, "The Ring of Truth: A Case Study of Credibility Assessment in Particular Social Group Determinations" (2009) 21 Intl J Refugee L 1 at 7.

¹¹ "Chairperson's Guideline 9: Proceedings Before the IRB Involving Sexual Orientation and Gender Identity and Expression" (1 May 2017), online: *Immigration and Refugee Board of Canada* <irb-cisr.gc.ca/en/legal-policy/policies/Pages/GuideDir09.aspx#a3> [perma.cc/HTY4-C529] [IRB, "Guideline"].

¹² *Ibid* at 6.1.

¹³ "Guidelines on International Protection No 9: Claims to Refugee Status based on Sexual Orientation and/or Gender Identity within the context of Article 1A(2) of the 1951 Convention and/or its 1967 Protocol relating to the Status of Refugees" (23 October 2012) at no 60(ii), online (pdf): *United Nations High Commissioner for Refugees* <www.unhcr.org/publications/legal/50ae466f9/guidelines-international-protection-9-claims-refugee-status-based-sexual.html> [perma.cc/PH3J-DG8J] [UNHCR, "Guideline"] (footnotes omitted).

warnings on many occasions. Federal Court case law going back over a decade has established that it is a reviewable error for IRB adjudicators to make inferences about the credibility of sexual minority refugee claimants based on stereotypes about physical appearance.¹⁴ The Federal Court has gone so far as to say that relying on such stereotypes “reveals a level of ignorance and prejudice which is not only unusual in general, but is particularly astonishing on the part of a decision maker who is in a position to adjudicate sensitive claims.”¹⁵ For some time, training materials for IRB adjudicators have called on decision makers to avoid these stereotypes.¹⁶

There is, therefore, reason to be worried that stereotypes about physical appearance may be resistant to warnings in domestic and international guidelines. As Sarilee Kahn and Edward Alessi put it: “[T]here is a long way to go before directives are implemented on the ground. This will require a fundamental shift in the hearts and minds of adjudicators so they may move beyond preconceptions of what LGBT looks like.”¹⁷

Worse still, warnings about avoiding stereotypes regarding the appearance of sexual minorities may have unintended consequences. For example, adjudicators might continue to be influenced by these stereotypes in making their decisions, but realize that they must stop referring to them in their written reasons. Such decisions would be upheld on appeal. On the one hand, such a change may arguably be a positive development. Written reasons for decisions would not draw on offensive stereotypes and adjudicators would have to use other legally permissible reasons to justify their decisions. On the other hand, such a change may also decrease transparency and oversight in decision making, denying claimants written evidence necessary to overturn negative decisions on appeal.

This raises a tricky question for those of us interested in combatting the use of stereotypes in sexual minority refugee determinations. Assuming that (most) refugee adjudicators are now aware that they should not draw on stereotypes about the physical appearance of sexual minorities when they write their decisions, how can we know whether such stereotypes nonetheless continue to affect decision making? This study addresses this question.

¹⁴ *Herrera v Canada (Minister of Citizenship and Immigration)*, 2005 FC 1233 at para 12 [*Herrera*]; *Lekaj v Canada (Minister of Citizenship and Immigration)*, 2006 FC 909 at para 17.

¹⁵ *Herrera*, *supra* note 14 at para 17.

¹⁶ See e.g. Nicole LaViolette, “Sexual Orientation and the Refugee Determination Process: Questioning a Claimant About Their Membership in the Particular Social Group” (May 2004), online: SSRN <ssrn.com/abstract=2294763> [perma.cc/J42D-WDMZ]; Nicole LaViolette, “Sexual Orientation, Gender Identity and the Refugee Determination Process in Canada” (1 March 2015), online: SSRN <ssrn.com/abstract=2276049> [perma.cc/A284-5LWQ]. For a discussion of the importance of training, see Nicole LaViolette, “Overcoming problems with sexual minority refugee claims: Is LGBT cultural competency training the solution?” in Thomas Spijkerboer, ed, *Fleeing Homophobia: Sexual Orientation, Gender Identity and Asylum* (New York: Routledge, 2013) 189.

¹⁷ Sarilee Kahn & Edward Alessi, “Performing for their lives: LGBT individuals seeking asylum” (9 June 2018), online: OUPblog <blog.oup.com/2018/06/lgbt-refugee-seeking-asylum/> [perma.cc/7BPT-9LC5].

III. Research Design

This study uses a first-year Canadian law school class as subjects to examine the impact of stereotypes about the physical appearance of sexual minority refugee claimants on both decision making and decision writing. The experiment was integrated into a first-year course on legal procedure. Enrolled students received participation grades in the course for either participating in the experiment or completing an alternative assignment.¹⁸ Out of 294 students in the first-year cohort, five chose not to participate at all (because they either dropped the course or chose to forfeit participation points); one participated late (that data being excluded from the study); four elected to complete the alternative assignment and the remaining 284 completed the experiment.

Students who chose to participate in the experiment played the role of a simulated refugee adjudicator assessing the credibility of a refugee claimant. The simulation took place through online survey software. Simulated adjudicators began by reviewing brief guidelines with basic principles drawn from Canadian administrative and refugee law about how to make credibility assessments. Adjudicators were then told that a refugee claimant had established that gay men from his country faced severe persecution and were tasked with determining whether the claimant's testimony about his sexual orientation was credible. In other words, had the claimant established that he was, in fact, gay?

Simulated adjudicators were provided with excerpts of the refugee claimant's case file. The excerpts included evidence that supported the claimant's assertions about his sexual orientation as well as evidence that contradicted or otherwise weakened the claimant's assertions.

The excerpts were carefully calibrated so that simulated adjudicators applying the guidelines could reasonably believe or disbelieve the claimant. After reviewing the excerpts, adjudicators were asked to indicate whether, on a balance of probabilities, the claimant's assertions about his sexual orientation were credible.

The simulated adjudicators were incentivized to prepare written reasons that complied with the guidelines by awarding students who prepared such reasons with \$10 gift certificates redeemable at the campus coffee shop.

Simulated adjudicators were divided randomly into two groups. The two groups were given different guidelines, but both groups were given the same excerpts of the refugee file.

¹⁸ The alternative assignment (reading a law journal article and writing a brief reflection) was provided to ensure that participation in the research was voluntary, in keeping with norms established for social psychology research. American Psychological Association, *Ethical Principles of Psychologists and Code of Conduct*, Washington, 2017, s 8.04(b), online (pdf): American Psychological Association <www.apa.org/ethics/code/ethics-code-2017.pdf> [perma.cc/JF4V-S6Q9].

The first group, comprised of one half of the simulated adjudicators, were directed not to draw upon stereotypes about the physical appearance of sexual minorities in justifying the adjudicators' credibility assessments in cases involving sexual orientation.¹⁹ In contrast, the guidelines given to the second group did not mention sexual orientation.

The two groups were further divided into three subgroups. The first subgroup received, along with the claimant's written statement, a photograph of a man who reflected stereotypical markers of homosexuality in western cultures. The second subgroup received, along with the written statement, a photograph of a man who reflected stereotypical markers of heterosexuality in western cultures. The third subgroup received the statement alone with no photograph.

In choosing our photographs,²⁰ we drew on literature that investigates stereotypes about the physical appearance of gay and straight men that are most frequently held by people sharing the demographics of our simulated adjudicators – mostly university-educated students in their twenties living in a Canadian urban centre.

The literature suggests that people drawn from this demographic are more likely to conclude that a man is straight if they judge his facial features to be "more masculine"²¹ and are more likely to conclude that a man is gay if they judge his facial features to be "more feminine"²² – especially if he looks happy.²³ Individuals are also likely to draw inferences about sexual orientation based on gendered stereotypes about a man's hairstyle,²⁴ facial

¹⁹ The relevant provision of the guidelines said: "In cases in which a claimant's ethnic identity or sexual orientation are in issue, decision makers must be careful not to draw conclusions based on stereotypes about a claimant's mannerisms or physical appearance."

²⁰ We followed standard practice in the field and selected photos from online dating sites. For examples of other studies that use similar methods, see e.g. Nicholas O Rule et al, "Found in Translation: Cross-Cultural Consensus in the Accurate Categorization of Male Sexual Orientation" (2011) 37:11 *Personality & Soc Psychol Bulletin* 1499 at 1504; Joshua A Tabak & Vivian Zayas, "The Roles of Featural and Configural Face Processing in Snap Judgments of Sexual Orientation" (2012) 7:5 *Plos One* 1 at 2–3; Minna Lyons et al, "Detection of Sexual Orientation ("Gaydar") by Homosexual and Heterosexual Women" (2014) 43:2 *Archives of Sex Behaviour* 345 at 347; Susan M Hughes & Robert Bremme, "The Effects of Facial Symmetry and Sexually-dimorphic Facial Proportions on Assessments of Sexual Orientation" (2011) 5:4 *J Soc, Evolutionary & Cultural Psychol* 214 at 218. In retrospect, we would no longer use this methodology and would, instead, hire models. For a discussion of our reflections on why we would no longer use this methodology, see Sean Rehaag & Hilary Evans Cameron, "Integrating Research Experiments into Law School Courses" (on file with author).

²¹ Jonathan B Freeman et al, "Sexual Orientation Perception Involves Gendered Facial Cues" (2010) 36 *Personality & Soc Psychol Bulletin* 1318 at 1323. See also Malvina N Skorska et al, "Facial Structure Predicts Sexual Orientation in Both Men and Women" (2015) 44 *Archives of Sex Behavior* 1377 at 1378; Gerulf Rieger et al, "Dissecting "Gaydar": Accuracy and the Role of Masculinity-Femininity" (2010) 39 *Archives of Sex Behavior* (2010) 124.

²² Freeman et al, *supra* note 21 at 1323. See also Rieger et al, *supra* note 21; Skorska et al, *supra* note 21 at 1378.

²³ Nicholas O Rule, "Perceptions of Sexual Orientation From Minimal Cues" (2017) 46 *Archives of Sex Behavior* 129 at 134.

²⁴ Nicholas O Rule et al, "Accuracy and Awareness in the Perception and Categorization of Male Sexual Orientation" (2008) 95 *J Personality & Soc Psychol* 1019.

grooming²⁵ and fashion choices.²⁶

We selected images of subjects whose appearances matched these stereotypes and we took care to minimize other differences. For example, we chose subjects who appeared to be of the same race and age. We used photographs that were of similar size and quality in which the subject was in a similar pose and in which a similar proportion of his body was visible.

We must admit that we felt some trepidation in selecting images to reflect the very stereotypes about perceived sexual orientation that our research was hoping to combat. As we were working on the study, hotly contested debates were raging about research that uses machine learning algorithms involving facial patterns to predict a person's sexual orientation.²⁷ It is important to emphasize, however, that in conducting this research, we were not looking to contribute in any way to academic debates about the validity of "gaydar".

We sidestepped these debates because we began with the current state of Canadian refugee law which prohibits adjudicators from relying on stereotypes about physical appearance to assess credibility regarding sexual orientation. In other words, our study is agnostic on the empirical question of whether "gaydar" may, in constrained circumstances, offer some level of predictive power because refugee adjudicators may not lawfully rely on such predictions.²⁸

In this study, then, we are interested in whether our simulated adjudicators relied on these unlawful predictions and, if so, whether this was reflected in the written reasons they prepared.

²⁵ *Ibid*; Freeman et al, *supra* note 21 at 1321.

²⁶ William Cox et al, "Inferences About Sexual Orientation: The Roles of Stereotypes, Faces, and The Gaydar Myth" (2016) 53 J Sex Research 157 at 158.

²⁷ Yilun Wang & Michal Kosinski, "Deep Neural Networks are More Accurate than Humans at Detecting Sexual Orientation from Facial Images" (2018) 114 J Personality & Soc Psychol 246. But see *contra* Andrew Gelman, Greggor Mattson & Daniel Simpson, "Gaydar and the Fallacy of Decontextualized Measurement" (2018) 5 Sociological Sci 271; Arianne Miller, "Searching for Gaydar: Blind Spots in the Study of Sexual Orientation Research" (2018) 9:3 Psychol & Sexuality 1. See also Drew Anderson, "GLAAD and HRC call on Stanford University & responsible media to debunk dangerous & flawed report claiming to identify LGBTQ people through facial recognition technology" (8 September 2017), online: GLAAD <www.glaad.org/blog/glaad-and-hrc-call-stanford-university-responsible-media-debunk-dangerous-flawed-report> [perma.cc/4G5D-EWE2].

²⁸ Though we would hasten to add that, even if, as an empirical matter, "gaydar" has some level of predictive power, the context of refugee adjudication would make relying on such predictions extremely dangerous. The stakes involved in refugee adjudication are very high: false negatives mean that a person may be deported to face persecution, torture or even death in contravention of international law. Unless "gaydar" were to be shown to have a very high level of sensitivity (which seems unlikely), there would be a serious risk of false negatives. Moreover, refugee adjudication generally involves cross-cultural communication with claimants who may have had to hide their sexual orientation for years as well as those who may be suffering from mental health challenges resulting from the trauma that they have experienced. All of these factors would surely complicate the exercise of "gaydar." Beyond that, "gaydar" would likely be opaque in that it would often be difficult for adjudicators to offer reasons justifying their determinations that could be subject to meaningful oversight on appeal or review.

IV. Limitations

We would like to acknowledge that our study has several limitations. First, the simulated refugee adjudication experiment differs in several important respects from actual refugee adjudication. Our simulated adjudicators were first-year law students, who only received very general guidelines about how to make credibility determinations. For many simulated adjudicators, this was the first time that they encountered refugee adjudication. By contrast, actual Canadian refugee adjudicators are drawn from different demographics, go through prolonged training about refugee law, participate in an adjudicative culture at the IRB and build up extensive adjudicative experience. Moreover, the context of actual refugee adjudication is very different with actual adjudicators being well aware of the consequences of their real-world decisions. For our simulated adjudicators, the stakes were quite low (a \$10 gift certificate). Certainly, too, there is a significant difference between reading a brief excerpt of a case and conducting an actual refugee hearing with extensive supporting materials, hours of live testimony, the ability to ask questions and knowing that one's decision may be subject to appeal or judicial review, etc. All of this means one must be cautious in generalizing conclusions based on our findings.²⁹

Second, our ability to generalize is constrained by our limited sample size and the narrow range of photographs provided to our simulated adjudicators. For example, we chose only photographs of men. This means that our study does not deal with how this phenomenon might play out with photographs of women or people with other gender identities. Similarly, because we selected photographs of men who were white, we do not know how race may have intersected with our results. Also, we elected to use an example of a refugee claimant who identified as gay. Therefore, we do not know whether adjudicators would have responded to other sexual minorities in the same way. In addition, while we tried to limit confounding factors in selecting our images, it remains possible that some factors that are unaccounted for affected the responses of our adjudicators. Generalizations based on our findings, therefore, must be undertaken with caution.

Finally, we encountered a methodological hiccup in running the study. Each simulated adjudicator examined, in random order, two case files: the sexual orientation file that is the basis of this article and another case file that involved an experiment about credibility and demeanour in video-recorded testimony. The guidelines provided to adjudicators included either

²⁹ For example, recent research in the United States has found that experimental studies about judicial bias produce different results when research subjects are sitting judges, lawyers and law school students, see Dan Kahan et al, "'Ideology' or 'Situation Sense'? An Experimental Investigation of Motivated Reasoning and Professional Judgment" (2016) 164 U Pa L Rev 349.

instructions about avoiding stereotypes relating to the physical appearance of sexual minorities (relevant to our first study, but not to our second) or about not relying on demeanour in testifying, such as pauses, fidgeting and eye contact avoidance (relevant to our second study, but not our first). We assumed that the two case files would be entirely independent of one another. That is, we did not expect that seeing one experimental condition in one case file would have any impact on outcomes or reasons for the other case file.

It turned out, however, that this assumption was mistaken. On reflection, that is perhaps not surprising. Research into real-world refugee adjudication has revealed that adjudicators may adhere to the gambler's fallacy, which is the false belief that small subsets of a population will reflect distributions expected in the general population.³⁰ For example, if an equally weighted flipped coin turns up heads twice in a row, the gambler's fallacy may lead one to believe that tails is more likely on the third flip, even though tails is no more likely on the third flip than on the first or second. In the American context, research has shown that judges who grant a claim are more likely to deny immediately subsequent claims, as would be expected if judges adhere to the gambler's fallacy. In other words, some judges appear to think that after one or more well-founded claims they are "due" for an unfounded claim.³¹

Whether because of the gambler's fallacy or some other factor, our simulated refugee adjudicators were more likely to disbelieve the claimant in one case file if they believed the claimant in the other case file. Of the 132 simulated adjudicators who found the claimant to be not credible in our demeanour case study, 87.1% found the claimant in our sexual orientation case study to be credible. By contrast, of the 152 simulated adjudicators who found the claimant to be credible in the demeanour case study, 79.6% found the claimant in our sexual orientation case study to be credible.³²

Due to the study design, this potentially confounding factor is particularly problematic with regard to understanding the impact of the guidelines on the rationales offered for credibility assessments of simulated adjudicators. Specifically, we cannot necessarily infer that the warnings regarding stereotypes about sexual minorities in the guidelines were responsible for the differences that we observed between the decisions provided by adjudicators who received the guidelines about stereotypes and those who did not.³³

³⁰ Daniel Chen, Tobias Moskowitz & Kelly Shue, "Decision Making Under the Gambler's Fallacy: Evidence from Asylum Judges, Loan Officers, and Baseball Umpires" (2016) 131:3 QJ Econ 1181.

³¹ *Ibid.*

³² These differences are marginally statistically significant, at the $p < 0.1$ level. $C^2 (1, N = 284) = 2.842, p = .091$.

³³ Simulated adjudicators found the sexual minority claimant to be credible less frequently when they saw guidelines that instructed them to avoid applying stereotypes about sexual minorities (78.0%) than when they saw guidelines that did not mention sexual orientation (88.1%). This is the opposite of what one might expect, which suggests that further research would be warranted to see whether this outcome tells us something meaningful about the impact of guidelines or whether it is merely a quirk caused by the lack

This is due to a lack of independence. Each of the adjudicators who saw the warning about sexual minority stereotypes did not see the warning relevant to the demeanour study, and vice versa.

In the end, this methodological hiccup prompted us to redo the demeanour case study in which we were primarily interested in the impact of guidelines relating to demeanour on credibility assessments. For the sexual orientation study, however, the hiccup had a more limited effect. As demonstrated in the results below, it is clear that the simulated adjudicators already knew that they should not draw on stereotypes about the appearance of sexual minorities in their written reasons.³⁴ As a result, the key findings from this experiment relate not to the impact of the guidelines on written reasons, but to the effect of other experimental conditions. This includes findings related to the impact of the photographs of the sexual minority refugee claimants on outcomes, on written reasons provided to justify those outcomes,³⁵ as well as findings that flow from our qualitative review of the written reasons offered by the simulated adjudicators.³⁶

Taken together, we acknowledge several limitations of our study that make generalization from our findings difficult. We also acknowledge that, in hindsight, it would have been preferable to simplify our research design.

Despite these limitations, the study produced interesting implications which are detailed below. Moreover, the experimental hiccup produced noteworthy results (i.e. experimental evidence of interactions across multiple simulated refugee determinations possibly due to the gambler's fallacy) that may assist future researchers with their own study.

V. Research Results

A. Stereotypes Affecting Decision Making Not Reflected in Decision Writing

The table below summarizes the main findings of the study. The simulated refugee adjudicators were more likely to believe the asserted sexual orientation of the claimant as a gay man when the photograph contained stereotypical markers of homosexuality (88.0%) than if the photograph depicted stereotypical markers of heterosexuality (75.5%). Adjudicators in the control group who were not provided any photograph fell in between these two other groups (85.3%).³⁷

of independence in our research design.

³⁴ Only 1 out of 284 simulated refugee adjudicators referred to the claimant's physical appearance.

³⁵ Binary logistic regression was undertaken to ensure that our observations with regard to these experimental conditions were not explained by which version of the guidelines simulated adjudicators saw.

³⁶ See implications B through D, below.

³⁷ These differences are statistically significant at the $p < 0.05$ level: $C^2(2, N = 284) = 6.065, p = .048$. Focusing

Credibility Determinations by Claimant Photograph			
Claimant Photograph	Number of Adjudicators	Adjudicators Finding Claimant Credible (%)	Adjudicators Explicitly Relying on Physical Appearance (%)
"Gay" Presentation - Stereotypical Markers	95	88.0	1.1
"Straight" Presentation - Stereotypical Markers	94	75.5	0.0
Control (No Photograph)	95	85.3	0.0
Total	284	83.1	0.4

While stereotypes about the appearance of sexual minorities affected credibility assessments (i.e. decision making), this was not reflected in the written reasons provided to justify those credibility assessments (i.e. decision writing). In fact, only one simulated refugee adjudicator explicitly referred to the physical appearance of the claimant in the written reasons justifying their credibility assessment. That adjudicator wrote: "I have to say that providing the picture of [the claimant] made me first assess his sexuality based on his facial appearance, and I judged him to be homosexual. Then I read his statement which made me think he is again homosexual."³⁸

This adjudicator then went on to offer a close assessment of the various factors in the claimant's file that justified a positive credibility determination regarding the claimant's sexual orientation.³⁹

None of the other 284 simulated refugee adjudicators made any reference to the claimant's physical appearance.

One finding of this study, then, is that while the decision making of the simulated refugee adjudicators appeared to be affected by stereotypes about the physical appearance of sexual minorities, such stereotypes were seldom discernible in the written reasons. It is important to note, however, that due to the design limitations of the study noted above – including the interaction

specifically on the comparison that is of most interest, adjudicators who saw the "straight" presentation photos and those who saw the "gay" presentation photos, the differences are statistically significant with a slightly higher level of confidence: $C^2(1, N = 189) = 5.321, p = .021$. Binary logistic regression on the subset of adjudicators who saw either the "gay" or "straight" presentation photo confirms that the difference between adjudicators seeing either photo remains statistically significant ($p = 0.023$) when one controls for all other experimental conditions seen by this subset of adjudicators, including the version of the guidelines seen, the version of the demeanour case file seen, the credibility determination in the demeanour case file and the order in which the two case files were seen.

³⁸ Adjudicator 179 ("Gay" presentation).

³⁹ *Ibid.*

between our two case files⁴⁰ and the lack of independence regarding different guidelines seen by simulated adjudicators⁴¹ – some caution should be exercised in interpreting this finding to avoid inferences about how much influence stereotypes about “gaydar” have on assessments of credibility. We do not know whether the differences observed may have been amplified or reduced by the confounding factors.

B. Stereotypes About Sexual Behavior Reflected in Decision Writing

While our simulated adjudicators seldom drew explicitly on stereotypes about the physical appearance of sexual minorities in their written reasons, frequently, other problematic stereotypes about sexual behavior were evident. These stereotypes arose primarily in relation to evidence in the case file indicating that when he was a teenager, the claimant had engaged in a sexual relationship with a woman, producing a child. Several simulated adjudicators viewed this as strong evidence that the claimant was lying about being gay. For example:

The most important point for my decision is his former girl friend’s statement in her visa application that he is the father of her son and [the claimant] acknowledged that it was possible that he was the boy’s father. I have some knowledge of gay people and they cannot have sex with girls.⁴²

The fact that he could have been sexually aroused enough to impregnate a woman heavily suggests that he is not indeed a homosexual.⁴³

His admittance of potentially being the father adds further doubt.⁴⁴

The only evidence available relating to his sexuality points towards heterosexuality (the fact that he has a child).⁴⁵

[There is] strong evidence of his heterosexuality, which is the fact that an admitted ex-girlfriend has claimed him to be the father of her child.⁴⁶

Ultimately, what made me think that he made up the story and is not being truthful about being homosexual is at the very end when [the claimant] acknowledged that it was possible that he was the boy’s father. So, my question is why would someone

⁴⁰ *Supra* note 32 and accompanying text.

⁴¹ *Supra* note 33 and accompanying text.

⁴² Adjudicator 24 (“Gay” presentation).

⁴³ Adjudicator 35 (“Gay” presentation).

⁴⁴ Adjudicator 58 (“Straight” presentation).

⁴⁵ Adjudicator 67 (“Straight” presentation).

⁴⁶ Adjudicator 68 (“Straight” presentation).

take part of sexual activity if they are not interested in the opposite gender? As that would be the only way that he could be the possible father.⁴⁷

The fact that he has produced a child with his girlfriend [justifies a negative credibility finding].⁴⁸

[T]he fact that he admitted that he may be the father raises a reasonable doubt as to his homosexuality claims.⁴⁹

The incriminating evidence is [the claimant's] admission that it is possible he was the father of his girlfriend's baby meaning he had sexual relations with females.⁵⁰

These views are all prefaced on misconceptions about human sexuality, including the complex relationship between sexual orientation and sexual behavior.⁵¹ Compared with these students' findings, the IRB's guidelines adopt a more sophisticated understanding of that relationship. The guidelines warn adjudicators not to base their credibility assessments on over-simplified stereotypes about how sexual minorities may be expected to behave, noting in particular that adverse credibility findings should not be based on the mistaken presumption that sexual minorities "would not have had heterosexual sexual experiences" or that sexual minorities would not have had children.⁵² Case law confirms that drawing on such presumptions constitutes a reviewable error.⁵³

While some simulated adjudicators drew explicitly on these mistaken presumptions about sexual behavior, many others expressed in their written reasons strong objections to negative credibility inferences based on evidence in the file about the claimant's different-sex sexual relationship as a teenager. For example:

I find it contemptible that one would consider that [the claimant's] claim for refugee status would be cast into doubt simply on the basis that he had potentially fathered a child with his teenage girlfriend. The presumption that sexuality is a static

⁴⁷ Adjudicator 73 ("Straight" presentation).

⁴⁸ Adjudicator 77 ("Straight" presentation).

⁴⁹ Adjudicator 185 ("Gay" presentation).

⁵⁰ Adjudicator 238 (No picture).

⁵¹ See e.g. Brian Mustanski et al, "The Association Between Sexual Orientation Identity and Behavior Across Race/Ethnicity, Sex, and Age in a Probability Sample of High School Students" (2014) 104:2 Am J Pub Health (finding that there is a disconnect between sexual orientation and sexual behavior in US adolescents, with many youth who identify as gay or lesbian engaging in different-sex sexual behavior and with patterns in this disconnect differing across racialized groups); Juliet Richters et al, "Sexual identity, sexual attraction and sexual experience: the Second Australian Study of Health and Relationships" (2014) 11 Sex Health 451 at 454-56 (finding that, in Australia, around 50% of men who identified as homosexual had had sexual experiences with female partners and that around 75% of women who identified as lesbian had had sexual experiences with male partners).

⁵² IRB, "Guideline", *supra* note 11 at 6.1.

⁵³ See e.g. X (*Re*), 2016 CanLII 39702 (CA IRB) at para 63.

dichotomy appears to be one which is ingrained in our society yet demonstrably fallacious. There is no reason that this evidence need be considered against the claimant; teenagers experiment.⁵⁴

Sexuality is not as black and white as some people make it out to be and simply because [the claimant] might have had a sexual relationship with a woman [...] does not mean that he does not actually identify as homosexual.⁵⁵

[T]he fact that he slept with a woman in the past (or even actively) has no bearing on the fact that he is homosexual. Many people choose to be sexually active with both sexes. This would not mitigate the fear he holds as the result of his attraction to the opposite sex.⁵⁶

The idea that [the claimant] can't simultaneously be a homosexual and have fathered a child with a woman is simply incorrect and premised on the problematic idea that sexuality is binary.⁵⁷

Having sexual intercourse with a partner of a different sex/gender has no bearing on your sexual orientation.⁵⁸

I do not find the fact that he could be the father of a child of a previous girlfriend compelling evidence that he is not homosexual. In fact, there are many cases where men come out after years of marriage with children. In terms of anatomy and the way the human anatomy functions, I find it entirely plausible that he may be the father of the child and also homosexual.⁵⁹

It is exceedingly common for homosexual men and women to have opposite sex relationships from time to time, either in confusion or under duress. It strikes me as inevitable that children may also occasionally arise from these relationships. I don't find this evidence to be material at all.⁶⁰

The second finding of the study, then, is that some simulated adjudicators drew on stereotypes about human sexuality in their written reasons and specifically on the inaccurate belief that homosexuals never engage in different-sex sexual behavior. Other simulated refugee adjudicators strongly disagreed with relying on such stereotypes.

⁵⁴ Adjudicator 25 ("Gay" presentation).

⁵⁵ Adjudicator 54 ("Straight" presentation).

⁵⁶ Adjudicator 60 ("Straight" presentation).

⁵⁷ Adjudicator 98 (No picture).

⁵⁸ Adjudicator 128 (No picture).

⁵⁹ Adjudicator 171 ("Gay" presentation).

⁶⁰ Adjudicator 210 ("Straight" presentation).

C. Similar Evidence Treated Differently by Different Adjudicators

Another striking finding from the study is that identical evidence was treated very differently by different adjudicators.

For example, the case file included an inconsistency in the timing of when the claimant came out to his mother as gay. The claimant's narrative and testimony indicated that he came out to his mother when he was 17; whereas an affidavit provided by his mother in support of his claim indicated that he came out to her when he was 15. The claimant explained the discrepancy by saying that his mother was simply mistaken about the timing.

For some simulated adjudicators, this discrepancy was among the strongest evidence in the file demonstrating that the claimant was not credible. For example:

I find it unlikely that his mother would not remember the age at which she found out her son was gay, given that this was seemingly a very important event in their lives.⁶¹

[T]he inconsistency in his age when he [came out to] his mother is implausible [...]. Such an event is lifechanging, and being 15 or 17 is a large difference in age that would be easily distinguishable.⁶²

[I]t is very unlikely that a mother would forget how old her son was when he told her about his sexuality.⁶³

The most persuasive evidence to me is that his mother got the age he told her she was gay wrong. Given the attitude to homosexuality in his home country, it is doubtful that she would forget such a significant event.⁶⁴

I find the fact that his mother mixed up the age that she found him to be homosexual very strange. I am a mother myself, and it is hard to believe an event like this about your child can be mistaken especially when his safety is on the line.⁶⁵

By contrast, other simulated adjudicators did not think that this discrepancy detracted from the claimant's credibility at all. For instance:

[I]f his mother were lying in the affidavit then their facts would have been the same because he would have spoken [to] her about it.⁶⁶

The errors are minor regarding his age when his mother discovered he was gay.⁶⁷

⁶¹ Adjudicator 52 ("Straight" presentation).

⁶² Adjudicator 58 ("Straight" presentation).

⁶³ Adjudicator 78 ("Straight" presentation).

⁶⁴ Adjudicator 190 ("Straight" presentation).

⁶⁵ Adjudicator 251 (No picture).

⁶⁶ Adjudicator 12 ("Gay" presentation).

⁶⁷ Adjudicator 33 ("Gay" presentation).

Not unusual for parents to be mistaken about an age at which something happened, especially if it's something distressing for the parent (ex. my mother does not recall having conversations about sexuality with me in high school and insists that this only happened after I came out - in my final year of university).⁶⁸

[M]y mother often forgets what exact year I was born, mistaking it for a year or two earlier, let alone the exact date that she found out a personal fact about me – so this did not deter me from finding that he is homosexual.⁶⁹

The biggest thing in [the claimant]'s favour is the affidavit from his mother. Given the time lapse, it is possible an error was made in the age (and regardless would not change its veracity).⁷⁰

A second example of simulated refugee adjudicators interpreting similar evidence differently was a falsified newspaper account submitted by the claimant that was intended to provide additional support for his claim. The fraudulent newspaper article described the claimant's father as a prominent police officer with violently homophobic views. The claimant acknowledged the forgery during his testimony.

Some simulated adjudicators relied heavily on the fraudulent evidence to find that the claimant was not credible:

I would interpret his newspaper fraud to be more than just a mere exaggeration, but a grossly dishonest act that raises serious doubt about the core of his story.⁷¹

[T]he falsified newspaper shows a pattern of dishonesty.⁷²

My opinion is also shaped by the [...] history of fraud that [the claimant] has displayed [...] in the forgery of the newspaper article.⁷³

Other adjudicators found that the forgery did not detract from the claimant's credibility, but rather constituted further evidence in support of the claimant's fear of persecution:

[The claimant] producing the false newspaper article [attests] to his fear [and] doing everything he can to avoid having to go back to his country.⁷⁴

The fact that he forged documents to me was something that made me believe he was homosexual because I find that you only go through such lengths when you have a

⁶⁸ Adjudicator 47 ("Gay" presentation).

⁶⁹ Adjudicator 200 ("Straight" presentation).

⁷⁰ Adjudicator 232 ("Straight" presentation).

⁷¹ Adjudicator 116 (No picture).

⁷² Adjudicator 174 ("Gay" presentation).

⁷³ Adjudicator 256 (No picture).

⁷⁴ Adjudicator 145 ("Gay" presentation).

reason to be scared (resort to desperate measures).⁷⁵

[H]e did engage in fraud by forging the [...] newspaper article. However, I think this forgery substantiates his fear.⁷⁶

A third example of simulated adjudicators treating the same evidence differently involves photographs submitted by the claimant showing him attending the Pride parade in Toronto.

Some adjudicators disregarded these photographs because there is no necessary connection between attending Pride events and one's sexual orientation:

[A] picture at pride is hardly strong evidence of one's homosexuality (I have a few photos at Pride and am not homosexual).⁷⁷

[The] Pride photo is not very strong evidence. People go to Pride without being gay, especially when they're new to the city and the phenomenon of Pride.⁷⁸

Watching the Pride parade is not persuasive evidence. Everyone goes to Pride!⁷⁹

Other simulated adjudicators, however, found that the Pride photographs strengthened the claimant's credibility, particularly as the case file indicated that the Pride photographs were taken before the claimant was in contact with the lawyer who assisted him with his claim:

[T]he fact that he attended the Pride parade prior to meeting his lawyer adds more credibility to his story.⁸⁰

I believe his attendance at Pride prior to speaking to the lawyer is important as well as it demonstrates his pursuit of LGBT activities.⁸¹

The most important factor in credibility I believe is the pictures at the pride festivities pre-dating his [meeting with his lawyer].⁸²

A final example of differential treatment of the same evidence by different simulated adjudicators relates to the claimant's explanation for his delay in making his refugee claim. According to evidence in the case file, the claimant travelled using a fraudulently obtained visitor visa and worked unlawfully in

⁷⁵ Adjudicator 205 ("Straight" presentation).

⁷⁶ Adjudicator 147 ("Gay" presentation).

⁷⁷ Adjudicator 68 ("Straight" presentation).

⁷⁸ Adjudicator 93 ("Straight" presentation).

⁷⁹ Adjudicator 190 ("Straight" presentation).

⁸⁰ Adjudicator 195 ("Straight" presentation).

⁸¹ Adjudicator 211 ("Straight" presentation).

⁸² Adjudicator 239 (No picture).

Canada. He made a refugee claim only after he was detained on immigration grounds and after speaking to a lawyer. In his testimony, the claimant explained his delay in seeking refugee protection by saying that he believed refugees were people fleeing armed conflict and did not know that it was possible to make a refugee claim based on sexual orientation.

Some simulated adjudicators found that the delay in claiming refugee protection negatively impacted the claimant's credibility and rejected his explanation that he did not know that a refugee claim could be made based on sexual orientation:

The fact that he did not declare his intent to seek refuge in Canada upon his arrival, and only did so after he was caught working for some months without authorization makes me believe that he does not have a genuine claim for protection on the basis of homosexuality, but was rather trying to pursue any avenue available to stay here for economic opportunities.⁸³

As an educated person I would think that [the claimant] would look into if homosexual people can make refugee claims or not.⁸⁴

Also, he didn't apply for refugee status right away (which he could have easily googled if he was unsure what it meant).⁸⁵

For other simulated adjudicators, however, the claimant's explanations for his delay in seeking refugee protection were persuasive:

I believe his story that he did not know that one could apply for refugee status for being homosexual. I did not know this was possible and I certainly don't think it is common knowledge.⁸⁶

I believe the claims that he did not know he could have applied for refugee status until after speaking to a lawyer. As a newcomer to the country, it is unlikely that he would have known the full extent of legal options available to him until instructed by a lawyer.⁸⁷

[I]t would make sense that he did not know he could make a refugee claim on the basis of sexual orientation. As something that was completely unacceptable in his home country, it makes sense that he would not have considered that it might be protected in another country.⁸⁸

It should be noted that these are not the only examples of differential treatment of the same evidence. A large proportion of the facts set out in the

⁸³ Adjudicator 79 ("Straight" presentation).

⁸⁴ Adjudicator 83 ("Straight" presentation).

⁸⁵ Adjudicator 135 (No picture).

⁸⁶ Adjudicator 62 ("Straight" presentation).

⁸⁷ Adjudicator 64 ("Straight" presentation).

⁸⁸ Adjudicator 250 (No picture).

case file led to disagreement between some simulated adjudicators about how a particular fact supported, detracted from or had no impact on the claimant's credibility. It is also worth noting that, in reasoning about how to interpret pieces of evidence, many adjudicators explicitly referred to their own life experiences, raising the possibility that different life experiences may be one of the factors accounting for divergent interpretations of similar evidence.

D. Some Simulated Adjudicators Uncomfortable Assessing Sexual Orientation

A final finding from the study is that some simulated refugee adjudicators expressed discomfort with the idea of assessing the veracity of someone's assertions of sexual orientation. For example:

This is a very difficult thing to prove, even if totally true.⁸⁹

Its ridiculous to question someone's sexual identity, especially as someone who grew up in a different cultural context than them.⁹⁰

To begin, I believe the manner in which [the claimant]'s sexual orientation is investigated is extremely invasive.⁹¹

I would say that flipping a coin would give me a better determination of this man's sexual orientation than my reasoned reflection would.⁹²

This claim also raises an ethical issue: to what extent should government authorities scrutinize the sexual orientation of a person? If [the claimant] identifies himself as a homosexual and there is evidence that he behaved in a way consistent with that identity, it does not seem ethical or even possible to truly test the matter on some objective criteria.⁹³

Another adjudicator's written reasons consisted of only two words to justify the finding that the claimant had established his sexual orientation on a balance of probabilities:

[S]elf identification.⁹⁴

Finally, another adjudicator who was uncomfortable assessing whether the claimant was in fact gay suggested an alternative approach:

⁸⁹ Adjudicator 23 ("Gay" presentation).

⁹⁰ Adjudicator 155 ("Gay" presentation).

⁹¹ Adjudicator 180 ("Gay" presentation).

⁹² Adjudicator 204 ("Straight" presentation).

⁹³ Adjudicator 262 (No picture).

⁹⁴ Adjudicator 223 ("Straight" presentation).

I believe that the question can be further clarified – whether or not he actually is OR PERCEIVED TO BE homosexual in his home country? Since no one in his home country can prove or disprove his sexuality any more than we in Canada can, all [the claimant] requires for his case is to demonstrate whether people in his home country perceive him to be homosexual.⁹⁵

Comments were also provided by the same adjudicator in the open-ended question at the conclusion of the experiment:

“[I]s he homosexual” [is] [...] not the right question [...]. [W]hether or not [the claimant] is homosexual is not really relevant, and lends to the larger (irrelevant) debate about what is homosexuality anyway?⁹⁶

It would seem, then, that at least in the view of some of our simulated refugee adjudicators the exercise of assessing the credibility of an asserted sexual orientation is either impossible, offensive or a distraction.

VI. Implications and Discussion

A. More Research Needed on Decision Making versus Decision Writing

The study has several implications. The first relates to the disjunction between decision making and decision writing observed in the experiment. At least some of the time, the written reasons did not fully reflect the factors affecting credibility assessments. For example, photographs of the claimant impacted credibility assessments of several adjudicators, but the written reasons of only one adjudicator mentioned the photograph.

We acknowledge that this study leaves many questions relating to the disjunction between decision making and decision writing unanswered. First, limitations in the research design suggest that caution is required when drawing inferences from the disjunction observed in this experiment for actual refugee adjudication. We do not know whether the observed disjunction is a feature of refugee adjudication, generally, or of the specific circumstances of the experiment.

Second, this study does not examine whether our simulated adjudicators were conscious of the various influences on their decision making. We do not know whether simulated adjudicators knowingly drew inferences about the claimant’s sexual orientation from the photographs, but intentionally left those inferences out of their written reasons or, instead, were simply unaware that their decisions were influenced by the photographs.

Despite these limitations, the disjunction between decision making and

⁹⁵ Adjudicator 179 (“Gay” presentation) (capitalization in original).

⁹⁶ *Ibid.*

decision writing observed in the experiment nonetheless raises provocative questions.

One interesting question is methodological: why are examinations of written reasons for decisions the dominant research methodology among refugee law scholars, often without any articulated justification? We do not suggest that such research is unhelpful. However, using this method is a choice that should be explained and the limitations that accompany this choice should be acknowledged. This is especially important if, as this study suggests, the connections between reasons offered for a refugee decision and the factors that led to a particular decision may be complicated.

The observed disjunction between decision making and decision writing also raises questions about what written reasons for refugee determinations should aspire to do. Should written reasons be prepared with the aim of offering a legally valid justification of a decision that can withstand scrutiny on review – even if those reasons do not reflect the actual decision making process followed by the adjudicator? Alternatively, should adjudicators instead attempt to make best efforts to describe their actual thought process in coming to a particular conclusion? Many might contend that, in principle, these two exercises ought to be the same (i.e. the actual thought process employed by adjudicators should be fully described *and* the described thought process should reflect a legally valid justification for the outcome). However, the disjunction observed in this study indicates that, in some cases, these exercises may not align. What should be done in such circumstances?⁹⁷

The extent to which this last question is pressing may hinge on how often one thinks there are multiple outcomes for which legally valid justifications could be offered. If one is of the view that there is only one legally valid outcome in a given case, then possible disjunctions between decision making and decision writing may not be troubling. That is so because, regardless of factors that may influence decision making, the decision writing exercise will discipline the outcome (i.e. if there is only one legally valid outcome, any legally valid decision-writing will necessarily produce that outcome). If, however, one is of the view that there are frequently multiple legally valid outcomes in a given case, then decision writing does not necessarily discipline decision making. In that case, disjunctions between decision making and decision writing raise serious fairness and rule of law problems. Our sense is that refugee credibility assessments often involve the latter scenario of multiple potentially legally valid outcomes due partly to the high level of

⁹⁷ For an argument that refugee adjudicators should strive not to write decisions that will be upheld on review but rather to facilitate review and to attempt to articulate as clearly as possible the thought process that led to the decision, see Sean Rehaag, “‘I Simply do not Believe...’: A Case Study of Credibility Determinations in Canadian Refugee Adjudication” (2017) 38 Windsor Rev Legal Soc Issues 38 at 67–69 [Rehaag, “Do not Believe”].

deference shown to the factual findings of first-instance decision makers upon review and partly to the conflicting and permissive grounds upon which negative credibility findings can be made.⁹⁸

In our view, existing scholarship in the refugee law field has not yet adequately grappled with these difficult questions. Much more research is needed to enhance understanding of the connections and disjunctions between refugee decision making and decision writing – and to provide guidance to refugee adjudicators on these issues. Further experimental research looking specifically at credibility assessments in sexual minority refugee adjudication is warranted – including research examining whether the findings of this study would be replicated in more tightly controlled circumstances (i.e. without the gambler’s fallacy problem) and research looking into the way that guidelines affect sexual-minority refugee decision making and decision writing.

B. Ongoing Need to Address Stereotypes in Refugee Adjudication of Sexual Minorities

The experiment confirms the ongoing need for guidance – in the form of training, guidelines and the like – for refugee adjudicators on avoiding stereotypes about sexual minorities.

The study shows that many of our simulated refugee adjudicators held misconceptions about sexual minorities that influenced their credibility assessments. Some of these misconceptions were explicit in written reasons for decisions (e.g. simplistic and incorrect understandings about the relationship between sexual orientation and sexual behavior) and some affected decision making without being reflected in the written reasons (e.g. stereotypes about what sexual minorities look like).

In our opinion, the fact that many of our simulated adjudicators – a demographic that is likely inclined to adopt relatively progressive and sophisticated understandings of human sexuality⁹⁹ – continue to hold these misconceptions, suggests that the IRB must continue with its efforts to challenge these inaccurate misconceptions among its adjudicators.

These efforts should include education, training and clear guidelines that

⁹⁸ Hilary Evans Cameron, *Refugee Law’s Fact-Finding Crisis: Truth, Risk, and the Wrong Mistake* (Cambridge: Cambridge University Press, 2018) at 161–71.

⁹⁹ Robert Andersen & Tina Fetner, “Cohort Differences in Tolerance of Homosexuality” (2008) 72:2 Pub Opinion Q 311 (finding substantial generational differences in tolerance of homosexuality in Canada and the US, with more recent generations being more supportive); E. Glenn Schellenberg, Jessie Hirt & Alan Sears, “Attitudes Toward Homosexuals Among Students at a Canadian University” (1999) 40:1/2 Sex Roles 139 (finding that attitudes towards gay men improved with increased post-secondary education); Naomi Carniol, “Cheers for queers”, *Canadian Lawyer 4Students* 6:1 (March 2011) 12, online: <www.canadianlawyermag.com/news/general/cheers-for-queers/268156> [perma.cc/LT62-78KQ] (citing University of Toronto law professor Brenda Cossman’s observation that she has seen dramatic progressive shifts in views related to LGBTQ issues among law students over her more than 20 years of teaching).

written reasons should not, as a matter of law, include stereotypes and other prejudicial inferences relating to sexual minority refugee claimants.

In addition, training for adjudicators should discuss unconscious bias and highlight that adjudicators may be affected by stereotypes unconsciously.¹⁰⁰ That training should stress active measures to combat unconscious reliance on these stereotypes in decision making. To this end, adjudicators should review literature on implicit bias¹⁰¹ and participate in active learning exercises involving implicit association tests designed to expose such biases.¹⁰² Along similar lines, the IRB should encourage adjudicators to participate in experimental research projects like this study and to regularly engage in processes of critical self-reflection.¹⁰³ Moreover, in order to suggest ways to improve decision making, legal scholars who have identified the many problems evident in sexual minority refugee decision writing¹⁰⁴ should consider adopting research methods that can help identify problematic features of sexual-minority refugee decision making about which adjudicators may not be aware. This research could then be shared with adjudicators.

C. More Research Needed on Credibility Assessments

A third implication of the study is that more research is needed on how refugee adjudicators make credibility assessments in general. This implication flows from the study's finding that different simulated adjudicators interpreted identical evidence differently. Some adjudicators found that a particular piece of evidence detracted from a claimant's credibility, while others found that the same evidence enhanced the claimant's credibility. For some other adjudicators, the same evidence had no bearing on the claimant's credibility. In addition, adjudicators frequently referred to their own life experience in considering how to interpret particular evidence.

¹⁰⁰ The IRB's guidelines currently tell adjudicators not to draw on stereotypes and list some of the prohibited stereotypes but do not explain why these stereotypes are problematic. IRB, "Guidelines", *supra* note 11 at 6.1. The UNHCR's guidelines, by contrast, explain why some of the stereotypes discussed are problematic. See e.g. UNHCR, "Guidelines", *supra* note 13.

¹⁰¹ See e.g. Mahzarin Banaji & Anthony G. Greenwald, *Blindspot: Hidden Biases of Good People* (New York: Delacorte Press, 2013); Anthony Greenwald, Debbie McGhee & Jordan Schwartz, "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test" (1998) 74:6 J Personality & Soc Psychol 1464. For an overview of research involving implicit bias regarding sexual minorities, Gregory Herek & Kevin McLemoresee, "Sexual Prejudice" (2013) 64 Ann Rev Psychol 309 at 313–15; Pasquale Anselmi et al, "Implicit and Explicit Sexual Attitudes Across Genders and Sexual Orientations" (2015) 15:1 J Bisexuality 40. For a discussion of implicit bias in the immigration decision-making context, see e.g. Fatima Marouf, "Implicit Bias and Immigration Courts" (2011) 32 Immig & Nat'lity L Rev 775.

¹⁰² Implicit Association Tests, including tests involving sexual orientation, are available online: *Project Implicit* <implicit.harvard.edu/implicit> [perma.cc/C5V8-SVEG].

¹⁰³ Audrey Macklin, "Truth and Consequences: Credibility Determination in the Refugee Context" in *The Realities of Refugee Determination on the Eve of a New Millennium: The Role of the Judiciary* (Haarlem: International Association of Refugee Law Judges, 1999) 134 at 139–40.

¹⁰⁴ *Supra* note 9 and accompanying text.

Research on how refugee adjudicators interpret (and should interpret) evidence involves many methodological challenges, particularly for legal scholars. Legal scholars have used standard legal doctrinal research methods, focusing mainly on appellate level judicial reviews of administrative decisions to examine refugee credibility assessments.¹⁰⁵ However, at least in the Canadian context, such methods are limited because of the high degree of deference shown by courts to factual findings made at first instance¹⁰⁶ and because, due to procedural barriers, only a small proportion of likely unrepresentative cases involving refugee determinations make their way to appellate level courts.¹⁰⁷

Research focusing on examinations of large numbers of written reasons at both the IRB and first-instance Federal Court levels seems more promising.¹⁰⁸ Yet, first-instance IRB refugee decisions at the Refugee Protection Division are seldom published,¹⁰⁹ a declining proportion of Refugee Appeal Division decisions are published¹¹⁰ and the vast majority of Federal Court judicial reviews of IRB refugee decisions are disposed of without written reasons.¹¹¹ Moreover, if researchers were to obtain large numbers of unpublished decisions,¹¹² a review of these decisions might tell us something about rationales offered for credibility assessments in decision writing. However, we could not necessarily assume that this tells us much about credibility

¹⁰⁵ See e.g. James Hathway & Michelle Foster, *The Law of Refugee Status*, 2nd ed (Cambridge: Cambridge University Press, 2014) at 136–61.

¹⁰⁶ Cameron, *supra* note 98 and accompanying text.

¹⁰⁷ The large majority of applications for judicial review involving refugee determinations are brought by unsuccessful refugee claimants and, thus, Federal Court case law is mostly made up of refugee claims that were denied at first instance. Only a small proportion of applications to the Federal Court involving refugee claims are determined on the merits at the Federal Court, and there are further restrictions on when even that small proportion of cases can be appealed beyond the Federal Court. Sean Rehaag, “Judicial Review of Refugee Determinations: The Luck of the Draw?” (2012) 38:1 *Queen’s LJ* 1 at 9–10 [Rehaag, “Luck”].

¹⁰⁸ For examples in the sexual-minority context, see e.g. Millbank, “Discretion”, *supra* note 9; Dauvergne & Millbank, *supra* note 9.

¹⁰⁹ A search performed on 9 March 2019 on Quicklaw (Lexis Advance) in “Canada Immigration and Refugee Board, Refugee Protection Division Decisions” indicated only 284 decisions in 2015, 136 decisions in 2016, 123 decisions in 2017 and 52 published decisions in 2018. During the same period, there were thousands of RPD decisions each year. “Refugee claims statistics” (last modified 22 May 2020), online: *Immigration and Refugee Board of Canada* <irb-cisr.gc.ca/en/statistics/protection/Pages/index.aspx> [perma.cc/X4R8-MHJV].

¹¹⁰ A search performed on 9 March 2019 on Quicklaw (Lexis Advance) in “Canada Immigration and Refugee Board, Refugee Appeal Division Decisions” indicated 1,883 decisions in 2015, 1,602 decisions in 2016, 467 decisions in 2017 and 324 decisions in 2018. During the same period there was an increase, rather than a decrease, in the number of decisions made at the RAD, with over 4,000 finalizations in 2018. “Refugee appeals statistics” (last modified 22 May 2020), online: *Immigration and Refugee Board of Canada* <irb-cisr.gc.ca/en/statistics/appeals/Pages/index.aspx> [perma.cc/RDA9-GPQG].

¹¹¹ Rehaag, “Luck”, *supra* note 107 at 23.

¹¹² Redacted decisions can be obtained from the Immigration and Refugee Board through access to information requests. For examples of research using such requests to study refugee decision-making, see e.g. Sean Rehaag, Julianna Beaudoin & Jennifer Danch, “No Refuge: Hungarian Roma Refugee Claimants in Canada” (2016) 52 *Osgoode Hall LJ* 705 at 729–31; Rehaag, “Do not Believe”, *supra* note 97.

assessments in decision making given the disjunction observed in this study.¹¹³

Quantitative research on decision making drawing on data about outcomes in refugee adjudication can help shed light on a variety of factors that might influence decision making,¹¹⁴ but such research cannot easily lead to inferences about credibility assessments in particular because this research focuses on outcomes, irrespective of the specific issues raised by individual cases.

Qualitative research methods involving interviews and observational studies in the field would be especially helpful in shedding light on questions related to credibility assessments, in the context of both sexual minority claims¹¹⁵ and refugee claims, generally.¹¹⁶ However, in addition to challenges relating to research ethics and access,¹¹⁷ these methods also raise challenges related to generalizability from small samples.

Taken together, the main research methods employed by most refugee law scholars and by interdisciplinary scholars have much to contribute to our understanding of how credibility assessments work, but also carry important limitations. In our view, the findings of the current study suggest that these research methods can be helpfully complemented by further research using experimental methods to examine how adjudicators make credibility assessments and how those assessments can be improved.¹¹⁸ The current study highlights the need for further research about the impact on credibility assessments of diverse lived experiences among refugee adjudicators and about the ways that refugee adjudicators collectively can bring their varied

¹¹³ An interdisciplinary team of scholars has tried to circumvent these problems by working with counsel to identify problematic cases and then to review not just the written reasons but the whole file. Such research has produced valuable insights, but the methodology generates challenging questions about how to generalize from small and unrepresentative samples. Cécile Rousseau et al., "The Complexity of Determining Refugeehood: A Multidisciplinary Analysis of the Decision-making Process of the Canadian Immigration and Refugee Board" (2002) 15:1 J Refugee Stud 1.

¹¹⁴ For examples of quantitative studies of outcomes in refugee adjudication, see e.g. Jaya Ramji-Nogales, Andrew Schoenholtz & Philip Schrag, *Refugee Roulette: Disparities in Asylum Adjudication and Proposals for Reform* (New York: New York University Press, 2009); Sean Rehaag, "Troubling Patterns in Canadian Refugee Adjudication" (2008) 39:2 Ottawa L Rev 335; Innessa Colaiacovo, "Not Just the Facts: Adjudicator Bias and Decisions of the Immigration and Refugee Board of Canada (2006-2011)" (2013) 1:4 J Migration & Hum Security 122.

¹¹⁵ See e.g. David Murray, *Real queer?: sexual orientation and gender identity refugees in the Canadian refugee apparatus* (London: Rowman & Littlefield, 2016); Nathalie Ricard, "Testimonies of LGBTIQ refugees as cartographies of political, sexual and emotional borders" (2014) 3:1 J Language & Sexuality 28.

¹¹⁶ See e.g. Sule Tomkinson, "Who are you afraid of and why? Inside the black box of refugee tribunals" (2018) 61:2 Can Pub Admin 184; François Crépeau & Delphine Nakache, "Critical Spaces in the Canadian Refugee Determination System: 1989-2002" (2008) 20 Intl J Refugee L 50.

¹¹⁷ Sule Tomkinson, "Doing Fieldwork on State Organizations in Democratic Settings: Ethical Issues of Research in Refugee Decision Making" (2015) 16:1 FQS Art 6.

¹¹⁸ For examples of recent studies that adopt these methods, see e.g. Sarah Filone & David DeMatteo, "Testimonial inconsistencies, adverse credibility determinations, and asylum adjudication in the United States" (2017) 3:2 Translational Issues in Psychological Sci 202; Hannah Rogers, Simone Fox & Jane Herlihy, "The Importance of Looking Credible: the Impact of the Behavioural Sequelae of Post-Traumatic Stress Disorder on the Credibility of Asylum Seekers" (2015) 21:2 Psychol Crime & L 139.

lived experiences to bear in order to assist each other to broaden their perspectives.¹¹⁹

D. Need New Approaches to Refugee Adjudication of Sexual Minorities

A final implication of the study is that it may be time to reconsider some of the foundational norms in sexual minority refugee adjudication. This conclusion flows from our finding that several simulated refugee adjudicators were uncomfortable with trying to go beyond a claimant's asserted sexual orientation to establish the veracity of those assertions. One reason for this discomfort is that making factual determinations about a person's sexual orientation necessarily implicates refugee adjudicators in deeply contested normative and empirical terrain.

On the normative side, the very idea of questioning someone's stated sexual orientation or identity is highly offensive among sexual minority communities and their allies. Sexual minorities have extensive experience with their bodies, behaviour, identity, and sexuality being heavily scrutinized – all too often in contexts in which such scrutiny carries with it the threat of heterosexist violence. Questioning asserted sexual orientations and identities is frequently used as a bludgeon against sexual minorities: “there's no such thing as ____”,¹²⁰ “it's just a phase”,¹²¹ “he just needs therapy”,¹²² “she just needs a good fuck”.¹²³ Similar types of questioning in the transgender context is currently a favoured strategy deployed by those seeking power and validation through public displays of transphobia.¹²⁴ It is thus not surprising that, when asked to engage in the exercise of scrutinizing a refugee claimant's asserted

¹¹⁹ For arguments about how that can work in the refugee law context, see Macklin, *supra* note 103. For a general argument about how encounters with different lived experience can improve judgement, see Jennifer Nedelsky, “Embodied Diversity and the Challenges to Law” (1997) 42:1 McGill LJ 91.

¹²⁰ Marjorie Garber, *Bisexuality & the Eroticism of Everyday Life* (New York: Routledge, 2000) at 16 (critically noting that there is a “common wisdom” that “[t]here is no such thing as bisexuality”). See also Kenji Yoshino, “The Epistemic Contract of Bisexual Erasure” (2000) 52 Stan L Rev 353 (arguing that straight and homosexual communities are invested in discrediting bisexuality).

¹²¹ Michael Woodford et al, “The LGBTQ Microaggressions on Campus Scale: A Scale Development and Validation Study” (2015) 62:12 J Homosexuality 1660 at 1672 (finding that being told that one's sexuality is “just a phase” is a common microaggression confronted by sexual minorities on US college campuses).

¹²² Peter Gajdics, “I experienced ‘conversion therapy’ – and it's time to ban it across Canada” (6 June 2018), online: *Maclean's* <www.macleans.ca/opinion/i-experienced-conversion-therapy-and-its-time-to-ban-it-across-canada/> [perma.cc/S9HH-VVNL] (describing the author's experience with years of therapy designed to eradicate his homosexuality).

¹²³ Didi Khayatt, “What's to Fear: Calling Homophobia into Question” (2006) 41:2 McGill J Educ 133 at 133, 141 (describing the author's experience as a lesbian teacher confronted by a homophobic student asserting what she “really needed is ‘a good fuck’”).

¹²⁴ Brenda Cossman, “Gender identity, gender pronouns, and freedom of expression: Bill C-16 and the traction of specious legal claims” (2018) 68:1 UTLJ 37 (discussing a well-known professor who constituted himself as an authority and drew a great deal of public attention by asserting that he would not defer to requests by students to use pronouns matching self-identified gender identities).

sexual orientation, some simulated adjudicators in our study pushed back.

On the empirical side, longstanding debates about how to understand sexual orientation continue unabated among scholars who study human sexuality. For example, there is little consensus on what sexual orientation means,¹²⁵ on the relation between sexual orientation and sexual behaviour,¹²⁶ or on what tools (if any) can be used to measure a person's sexual orientation.¹²⁷

If scholars cannot agree on what sexual orientation means, let alone how to measure it in controlled research settings, and if many in sexual minority communities and their allies find attempts to assess whether a person "really" is a member of a sexual minority to be objectionable, how should refugee adjudicators approach these questions?

In our view, sexual minority refugee adjudication (and refugee adjudication more generally) needs to involve decision makers carefully calibrating how they engage in factual findings for the specific context in which they are making their decisions. One of us has recently attempted to set out a systematic argument about what fact-finding would look like if the context of refugee adjudication were taken seriously.¹²⁸

In brief, refugee adjudicators should recognize that they can never hope to definitively establish the facts in a case. Instead, adjudicators can only ever make a best guess about those facts. As such, adjudicators should not make determinations about each individual fact and then draw conclusions based on those facts. Rather, they should examine the totality of the evidence and ask whether that evidence, taken together, demonstrates a serious risk of persecution. In doing so, it should never be sufficient for adjudicators to say simply that claimants have failed to establish a necessary fact. Rather, adjudicators who disbelieve the claimant's account should have a clearly articulated counter-theory that provides a more persuasive account of the

¹²⁵ Karine Igartua et al, "Concordance and Discrepancy in Sexual Identity, Attraction, and Behavior Among Adolescents" (2009) 45 *J Adolescent Health* 602 at 603 (noting that sexual orientation "is a multidimensional construct... the prevalence of nonheterosexual orientation varies widely depending on whether the definition includes identity, attraction, or behavior"). For summaries of scholarship on this topic, see e.g. Lisa Diamond & Clifford Rosky, "Scrutinizing Immutability: Research on Sexual Orientation and U.S. Legal Advocacy for Sexual Minorities" (2016) 53:4/5 *J Sex Research* 363 at 365; J Michael Bailey et al, "Sexual Orientation, Controversy, and Science" (2016) 17:2 *Psychol Sci in the Pub Interest* 45 at 61.

¹²⁶ See e.g. Preeti Pathela et al, "Discordance between Sexual Behavior and Self-Reported Sexual Identity: A Population-Based Survey of New York City Men" (2006) 145:6 *Ann Internal Med* 417 at 419 (finding a substantial discordance between self-reported sexual identity and sexual behavior among men in New York City – including that 78% of men who had sex with men in the prior year identified as straight); Anjani Chandra et al, "Sexual Behavior, Sexual Attraction, and Sexual Identity in the United States: Data From the 2006-2008 National Survey of Family Growth" (2011) 36 *CDC Nat'l Health Statistics Reports* (finding that sexual attraction and sexual identity are closely but not perfectly correlated in a large national survey in the US).

¹²⁷ See e.g. Randall Sell, "Defining and Measuring Sexual Orientation for Research" in Ilan H Meyer & Mary E Northridge, eds, *The Health of Sexual Minorities* (Boston: Springer, 2007) 355 at 370 ("Researchers wanting to measure sexual orientation today have a number of measurement tools from which to choose... None of these is completely satisfactory"); *supra* note 27 and accompanying text.

¹²⁸ Cameron, *supra* note 98.

totality of the evidence than the one offered by the claimant.

In addition, throughout this process, adjudicators must keep in mind that they may be mistaken about their assessments of the evidence and about their counter-theory. Adjudicators should be attentive to the very high stakes involved if claimants are incorrectly disbelieved (i.e. requiring refugees to return to face risks of persecution, torture or even death). In these circumstances, refugee adjudicators should resolve doubt in favour of claimants.

Applying this approach, if a sexual minority refugee claim is denied on credibility grounds, typically it will be based on the counter-theory that the claimant is heterosexual and fraudulently asserting a different sexual orientation for the purposes of obtaining refugee protection. If a refugee adjudicator wishes to make this finding, they ought to articulate why the evidence is best interpreted to mean that the claimant is heterosexual, despite the benefit of the doubt accorded to the claimant, the challenges of reliably assessing sexual orientation and the extremely high stakes if the decision maker gets this wrong. In our view, such a finding would be appropriate only under this model in clear cases, which would help reduce the number of false negative sexual minority refugee determinations. We also believe that this model appropriately encourages adjudicators to scrutinize their own reasoning about why a claimant is likely heterosexual, rather than scrutinizing only assertions made by sexual minority refugee claimants about their sexual orientations.

This approach to weighing evidence – one in which adjudicators assess risks of persecution holistically while giving the benefit of the doubt to claimants – fits nicely within an approach to sexual minority refugee adjudication that one of us has termed “queering Canadian refugee law”.¹²⁹ Under this approach, adjudicators should focus less on the sexual orientation of claimants and more on the various risks of persecution faced by claimants. The starting point for queering refugee law is for adjudicators to shift their gaze from scrutinizing sexual minorities to scrutinizing heteronormativity.¹³⁰ One way this can be achieved is by considering how heterosexuality is produced and enforced in the claimant’s country of origin, that is, by examining the many different practices that constitute compulsory heterosexuality.¹³¹ Rather than viewing the persecution at play as violence targeting an unpopular minority – thus raising the question of whether the claimant is, in fact, a member of

¹²⁹ Sean Rehaag, “Patrolling the Borders of Sexual Orientation: Bisexual Refugee Claims in Canada” (2008) 53 McGill LJ 59 at 93 [Rehaag, “Bisexual”].

¹³⁰ One of the central tenets of queer theory is to focus on incoherencies and slippages in the production of heterosexuality as normative or natural. Annamarie Jagose, *Queer Theory: An Introduction* (New York: New York University Press, 1996) at 3. See also Judith Butler, *Gender Trouble: Feminism and the Subversion of Identity* (New York: Routledge, 1999).

¹³¹ Adrienne Rich, “Compulsory Heterosexuality and Lesbian Existence” (1980) 5:4 Signs 631.

that unpopular minority – adjudicators would instead ask how the claimant has and may in the future be impacted by compulsory heterosexuality, recognizing that these impacts may have multiple connections to a variety of refugee convention grounds.¹³²

For example, sexual minority claimants may not be free to express their opinions about state policies and practices that enforce compulsory heterosexuality or they may be targeted for mistreatment because of expressed or imputed political views.¹³³ Additionally, sexual minority claimants may be targeted for violating religious norms relating to compulsory heterosexuality or they may be forced to avow or display adherence to such norms.¹³⁴ For instance, they may be mistreated for failing to meet socially accepted gender norms that accompany and partly constitute compulsory heterosexuality, be forced to display compliance with such norms,¹³⁵ or be subject to persecution for violating compulsory heterosexuality regardless of their sexual orientation.¹³⁶ Sexual minority claimants may be mistreated not for failing to present themselves as heterosexual, but for failing to enforce compulsory heterosexuality adequately in others, such as a sibling, parent or child.¹³⁷ There may also be scenarios where the claimant is a child whose sexuality has not yet developed, but who has the right not to be coerced into heterosexuality.¹³⁸

In all these circumstances, compelling cases can be made that claimants face forms of persecution related to compulsory heterosexuality that can attract refugee protection without adjudicators having to make a definitive factual finding about the claimant's sexual orientation. Instead, adjudicators can focus on holistically assessing the risks of heterosexist persecution faced

¹³² Refugee claimants must show that they face persecution on account of "race, religion, nationality, membership of a particular social group or political opinion." *Convention Relating to the Status of Refugees*, 28 July 1951, 189 UNTS 2545, Can TS 1969 No. 6 (entered into force 22 April 1954), art 1.A(2). See also *Immigration and Refugee Protection Act*, SC 2001, c 27, s 96.

¹³³ IRB, "Guideline", *supra* note 11 at 8.4.1; UNHCR, "Guideline", *supra* note 13 at 50.

¹³⁴ IRB, "Guideline", *supra* note 11 at 8.4.1; UNHCR, "Guideline", *supra* note 13 at 42–43.

¹³⁵ Rehaag, "Bisexual", *supra* note 129 at 95–97; Nicole LaViolette, "UNHCR Guidance Note on Refugee Claims Relating to Sexual Orientation and Gender Identity: A Critical Commentary" (2010) 22:2 Intl J Refugee L 173 at 181–83.

¹³⁶ IRB, "Guideline", *supra* note 11 at 8.3.1; UNHCR, "Guideline", *supra* note 13 at 41.

¹³⁷ IRB, "Guideline", *supra* note 11 at 8.3.2.

¹³⁸ There are various ways to connect such coercion to refugee convention grounds. One possibility is to frame this coercion as persecution on account of intersections between age and gender, both of which can constitute particular social groups. For discussions about gender and age as grounds for refugee protection, see generally "Chairperson Guidelines 4: Women Refugee Claimants Fearing Gender-Related Persecution" (13 November 1996), online: *Immigration and Refugee Board of Canada* <irb-cisr.gc.ca/en/legal-policy/policies/Pages/GuideDir04.aspx> [perma.cc/K5XB-MR3E]; "Guidelines on International Protection: Gender-Related Persecution within the context of Article 1A(2) of the 1951 Convention and/or its 1967 Protocol relating to the Status of Refugees" (7 May 2002), online: *United Nations High Commission for Refugees* <www.unhcr.org/publications/legal/3d58ddef4/guidelines-international-protection-1-gender-related-persecution-context.html> [perma.cc/XJA9-HZUQ]; Alice Edwards, "Age and Gender Dimensions in International Refugee Law" in Erika Feller, Volker Türk & Frances Nicholson, eds, *Refugee Protection in International Law: UNHCR's Global Consultations on International Protection* (Cambridge: Cambridge UP, 2003) 46.

by the claimant.

A queer approach to refugee adjudication has other advantages. One advantage is that this approach encourages intersectional analysis. Focusing less on the sexual orientation of the claimant and more on the experienced and feared persecution draws the adjudicator's attention to the unique circumstances of the claimant, including circumstances that may amplify or mitigate risks faced by the claimant. This approach can help adjudicators pay closer attention to ways in which gender, race, class, wealth, education, (dis)ability, social capital and other factors may lead to distinct encounters with compulsory heterosexuality.¹³⁹ A queer approach can also help adjudicators pay attention to differences between groups of sexual minorities in terms of the forms of persecution they may experience (rather than adopting one archetype of a persecuted sexual minority and then finding accounts of those whose experiences differ from that archetype to be non-credible).¹⁴⁰

Another advantage is that this approach encourages adjudicators to reflect on compulsory heterosexuality, not only in the claimant's country of origin, but also in Canada.¹⁴¹ Doing so can help counteract forms of heteronormativity that may otherwise be imposed during the refugee determination process itself. Such reflection can, for example, help adjudicators to recast some of the norms found in guidelines, not as instructions to be followed in decision writing in order to avoid having decisions overturned on review, but instead as recommendations to help adjudicators avoid acting unwittingly as agents of heteronormativity in the refugee determination process, including in their decision making.¹⁴²

¹³⁹ For discussions about the importance of paying attention to intersections, see e.g. Jenni Millbank, "Gender, Sex and Visibility in Refugee Claims on the Basis of Sexual Orientation" (2003) 18 *Geo Immigr LJ* 71; Rachel Lewis, "'Gay? Prove it': The politics of queer anti-deportation activities" (2014) 17:8 *Sexualities* 958; Jasmine Dawson & Paula Gerber, "Assessing the Refugee Claims of LGBTI People: Is the DSSH Model Useful for Determining Claims by Women for Asylum Based on Sexual Orientation?" (2017) 29:2 *Intl J Refugee L* 292; Efrat Arbel, Catherine Dauvergne & Jenni Millbank, eds, *Gender in Refugee Law: From the Margins to the Centre* (New York: Routledge, 2014). See also IRB "Guideline", *supra* note 11 at 8.5.2.

¹⁴⁰ Rehaag, "Bisexual", *supra* note 129; Millbank, "Discretion", *supra* note 9; Berg & Millbank, *supra* note 9 at 207–15; David Murray, "Learning to be LGBT: Sexual Orientation Refugees and Linguistic Inequality" (2018) 6:1 *Crit Multilingualism Stud* 57 at 68–69.

¹⁴¹ Such reflection can serve to push back against what some scholars have termed "homonationalism" in the way that sexual minority refugee claims are framed discursively as a one-way flight from homophobia in rights-violating countries abroad to safety in rights-respecting Canada, thereby shoring up Canada's place in various global hierarchies, forwarding colonialist nation-building projects and obscuring ongoing homophobia in Canada. See e.g. Edward Ou Jin Lee, "Responses to Structural Violence: The Everyday Ways in Which Queer and Trans Migrants With Precarious Status Respond to and Resist the Canadian Immigration Regime" (2019) 10:1 *Intl J Child Youth & Fam Stud* 70 at 73–74; David Murray, "Real Queer: 'Authentic' LGBT Refugee Claimants and Homonationalism in the Canadian Refugee System" (2014) 56:1 *Anthropologica* 21; Gary Kinsman, "Policing borders and sexual/gender identities: queer refugees in the years of Canadian neoliberalism and homonationalism" in Nancy Nicol et al, eds, *Envisioning Global LGBT Human Rights: (Neo)colonialism, Neoliberalism, Resistance and Hope* (London: Human Rights Consortium, Institute of Commonwealth Studies, 2018) 97.

¹⁴² See the text accompanying note 101.

Yet another advantage is that this approach avoids reinforcing the presumption that the world is sharply divided into heterosexuals and sexual minorities.¹⁴³ Such a sharp distinction invariably casts heterosexuality as a natural inclination of the majority, an inclination not naturally shared by sexual minorities (who are deserving of protection against mistreatment on the basis of natural inclinations for which they are not responsible).¹⁴⁴ But this presumption makes it difficult to understand the persecution that sexual minorities face. If the presumption were true, there would be no plausible purpose for the extensive, sustained and all-too-often violent social, economic and political efforts to produce, reproduce and enforce compulsory heterosexuality. By contrast, if one understands compulsory heterosexuality as something fragile that must be actively and violently (re)produced in countless social locations, then one can better understand why sexual minorities may face multiple (and intersecting) forms of persecution.

Of course, the approach that we have set out here, both at the level of assessing the evidence (approaching the evidence holistically and giving claimants the benefit of the doubt) and at the level of the reason for protection (focusing more on persecution related to compulsory heterosexuality and less on asserted sexual orientations), is not the only way that one might try to reform sexual minority refugee adjudication. There may be other approaches that would be worth considering. Our primary goal in this section is to show that there are options available for improving sexual minority refugee adjudication which address the discomfort expressed by some simulated adjudicators about trying to determine whether or not a refugee claimant “really” is a member of a sexual minority. That is a discomfort that we share, and that we think is shared by many real-world refugee adjudicators. In our view, there is an urgent need for further research examining and proposing options for rethinking the foundation of sexual minority refugee adjudication in this way.

VII. Conclusion

In this article, we have set out the results of an experiment examining simulated sexual minority refugee adjudication. In our experiment, a form of “gaydar” was relied upon by simulated adjudicators in their assessments of the credibility of a sexual minority refugee claimant. However, the written reasons provided by simulated adjudicators did not refer to the claimant’s

¹⁴³ Alfred Kinsey famously critiqued models of human sexuality that only have a place for exclusive heterosexuals and exclusive homosexuals, saying “[t]he world is not to be divided into sheep and goats.” Alfred Kinsey et al, *Sexual Behavior in the Human Male* (Philadelphia: W.B. Saunders, 1948) at 639.

¹⁴⁴ See generally Janet Halley, “Sexual Orientation and the Politics of Biology: A Critique of the Argument from Immutability” (1993–94) 46 *Stan L Rev* 503.

appearance. This gap highlights that decision making and decision writing are distinct and that one cannot assume that factors discussed in the latter are the same factors that influenced the former. The experiment revealed also that many other types of prejudicial stereotypes and misunderstandings about human sexuality were relied upon explicitly in the written reasons of the simulated adjudicators. Another key finding is that for much of the evidence in the case file, simulated adjudicators disagreed on whether particular pieces of evidence detracted from the claimant's credibility and, indeed, sometimes the same evidence relied upon by some adjudicators as detrimental to the claimant's credibility was viewed by other simulated adjudicators as enhancing the claimant's credibility. Finally, the experiment showed that at least some simulated adjudicators were uncomfortable with the very idea of scrutinizing a claimant's asserted sexual orientation.

In our view, these findings have four important implications.

First, refugee law scholarship must grapple more directly with possible disjunctions between decision making and decision writing through research methods that go beyond standard doctrinal legal research because simply examining written reasons for decisions may not tell us much about decision making.

Second, ongoing efforts must be made to ensure that adjudicators do not rely on prejudicial and inaccurate stereotypes about sexual minorities in decision making and decision writing. Efforts also should be made to help adjudicators avoid unconsciously being affected by these stereotypes in their decision making.

Third, more research into how adjudicators make credibility assessments in general is needed, including how adjudicators reason about evidence and how those assessments can be improved. Various research methods, including experimental methods, should be adopted more widely.

Fourth, and finally, our experiment illustrates the problems of scrutinizing a claimant's stated sexual orientation. Further research is needed about how sexual minority refugee adjudicators can adopt approaches that de-centre the practice of scrutinizing assertions about sexual minority sexuality and identity. This research should develop ways to move away from questions about whether a refugee claimant is "really" a member of a sexual minority towards questions about how violence operates in the claimant's country of origin to enforce heteronormativity and how that violence has affected and may, in the future, affect the claimant. In this regard, our experiment provides further support for "queering" Canadian refugee law.