

**Faculty of Education and Psychology: Gender-Based Agency
in the PISA Global Competencies Framework**

Nancy Vegas

Master's Thesis in Education, Development, and International Cooperation

Fall Term 2021

Faculty of Education and Psychology

University of Jyväskylä

1. ABSTRACT

Vegas, Nancy. 2021. Gender-Based Agency in the PISA Global Competencies Framework. Master's Thesis in Education, Development, and International Cooperation. University of Jyväskylä. Faculty of Education and Psychology.

Large-scale international assessments such as the Program for International Student Assessments (PISA) have tremendous impact on not just assessing student performance, but also guiding education policy and research. In 2018, PISA adopted a Global Competencies Framework (GCF), shifting its focus from the evaluation of mostly implicit economic citizenship skills to skills connected to a more explicit global citizenship curriculum. In this thesis, I aim to hold PISA accountable to its stated commitments to providing good data about gender equality. I specifically explored the way that PISA produces data about gender-based agency, or the extent to which a student is motivated to act politically on behalf of their gender group. The first research question examined the regimes of knowledge or discourses through which PISA GCF regulated gender subjectivities in their assessment of citizenship competencies. The second research question examined the ways that it measured the gender-based political agency and participation of students. These questions were explored through a Foucauldian discourse analysis and a quantitative content analysis.

The findings indicate that the 2018 GCF erased several elements of gender-based identity, agency, and action. Bibliographies and vocabulary can be used in large-scale international assessments to push an intercultural regime of knowledge and restrict gender subjectivities to heteronormative, cisnormative, and universal Western identity. In the case of the 2018 GCF, it resulted in the erasure of indigenous, Black, female and LGBTQ subjectivities, and in colorblind racism. This formulation of global competencies also resulted in the restricted measurement of gender-based political agency, the limitation of student understanding of political participation to interpersonal relations, and the encouragement of students to be neutral rather than to act for social justice.

Keywords: citizenship education, gender-based political agency, discourse analysis, content analysis, governmentality, large-scale international assessments, PISA.

2. CONTENT

Table of Contents

1. ABSTRACT	2
2. CONTENT	3
3. INTRODUCTION	4
4. LITERATURE REVIEW	6
4.1. GENDERED CITIZENSHIP EDUCATION.....	7
4.2. GENDER-BASED COMPETENCIES IN PISA	10
3. THEORY	13
3.1. RESOLVING THE ONTOLOGICAL ISSUE OF GENDER-BASED POLITICAL AGENCY.....	14
3.2. RESOLVING THE EPISTEMOLOGICAL ISSUE OF GENDER-BASED POLITICAL AGENCY.....	18
3.3. FRAMEWORK FOR GENDER-BASED POLITICAL IDENTITY, AGENCY, AND ACTION	19
4. RESEARCH PROBLEMS.....	22
4.1. PROBLEM STATEMENT	22
4.2. RESEARCH QUESTIONS	24
5. RESEARCH METHOD.....	25
5.1. METHODOLOGY	25
5.2. ETHICAL CONCERNS	29
5.3. DATA COLLECTION	30
5.3.1. <i>Material Selection</i>	30
5.3.2. <i>Data Collection</i>	31
5.4. DATA ANALYSIS	34
5.4.1. <i>Category Count</i>	34
5.4.2. <i>Analytics of Governmentality</i>	35
6. RESULTS	37
6.1. CATEGORY COUNT	37
6.2. FOUCAULDIAN DISCOURSE ANALYSIS	46
6.2.1. <i>Regime of Practices and Knowledge</i>	47
6.2.2. <i>Field of Visibility</i>	51
6.2.3. <i>Technical Aspect of Government</i>	53
6.2.4. <i>Governmentality as a Rational Activity</i>	58
6.2.5. <i>Formation of Identity</i>	66
7. DISCUSSION.....	71
7.1. RELIABILITY	78
7.2. LIMITATIONS	81
7.3. APPLICABILITY OF RESEARCH RESULTS AND CHALLENGES FOR FURTHER RESEARCH	83
7.4. CONCLUSION	84
8. REFERENCES.....	85
9. APPENDICES.....	103
9.1. APPENDIX A	103
9.2. APPENDIX B	104

3. INTRODUCTION

International large-scale quantitative assessments, such as the Program for International Student Assessment (PISA), have created a clear before-and-after moment in educational research and policy. Over the past two decades, these assessments, and the international comparative educational research they have inspired, have mostly focused on the measurement of educational competencies related to a competitive labor force. In recent years, perhaps due to the publication of the Global Sustainable Development Goals (SDGs), the PISA Framework adopted a Global Competencies Framework (GCF), shifting its focus from the evaluation of mostly implicit economic citizenship skills to skills connected to a more explicit civic curriculum.

In spite of these developments, and the research possibilities that this new data entails, a significant gender data gap remains. This gap is visible in the field of citizenship education (Sant, 2019), and in research and technology overall (Perez, 2019). The present thesis critically examines the production of data about citizenship education in international assessments by closely analyzing the way these exams make assessments of gender-based competencies. This research focus is based on the stated commitments of PISA to provide good quality data for the SDGs (OECD, 2016; OECD, 2019a). This data is significant because the PISA Framework is highly influential in international education policy and research (McGaw, 2008; Neumann, Fischer, and Kauertz, 2010; Carvalho, 2012; Pons, 2012; Bonal & Tarabini, 2013; Rutkowski, 2015; Cox & Meckes, 2016; Pons, 2017; Sellar & Lingard, 2013). Through a process called “PISA shock,” the new data could push countries to improve the quality of their citizenship education in connection to gender. In this thesis, I aim to improve data collection related to SDG indicators on gender equality, by evaluating the quality of gender-based PISA data.

In order to better assess the possible gender gap in data, I examined theoretical formulations of political agency, or decision-making, that influenced the rules of knowledge production about student citizens in both education research and international assessments. I pushed for an adjusted approach to the study of agency through the application of gender theory to widely accepted theoretical formulations of agency con-

constructs that inform both educational research and international assessments. First, I conducted a literature review of gender-based citizenship education research to identify the most pivotal gender-based citizenship curricula available. Then, I conducted a theoretical reformulation of agency constructs to account for queer, transgender, and gender non-conforming identities that are ontologically ambiguous. The development of the concept “gender-based political agency” in this thesis expanded and re-framed the subject-centered sociocultural definition of agency by placing a renewed focus on gender-based identity as a source of political action (Eteläpelto, Vähäsantanen, Hänn, and Paloniemi, 2013). Lastly, I outlined a theoretical framework for gender-based political participation through an explicit redefinition of gender-based political agency.

The method used to evaluate gender-based political agency in PISA originated in the fields of sociology of education and education policy. This study expanded on other analyses of the GCF that have identified bias (Ledger, Thier, Bailey, and Pitts, 2019), by focusing on gender-based competencies. I conducted a Foucauldian discourse analysis of the GCF assessment texts that included the tools of analytics of governmentality, and a micro-genealogy, or a partial genealogical (historical) study of how institutionally driven regimes of knowledge (i.e., disciplines and research fields) regulate subjectivities (Dean, 2009). Additionally, I conducted a quantitative content analysis using the category count method. The purpose was to reveal how these assessments formed regimes of knowledge that regulated the possibilities for gender-based political identity, agency, and action in citizenship education.

The research questions were as follows:

1. Through what kind of discourses or regimes of knowledge did 2018 PISA CGF regulate gender subjectivities in their assessment of citizenship competencies?
2. In what ways, if any, did the 2018 PISA GCF measure the gender-based political agency and participation of students?

4. LITERATURE REVIEW

In this literature review, I explored the discourses of gender and citizenship education relevant to an analysis of these concepts in large-scale international student assessments. Since I aimed to capture the dominant views in the literature, rather than using specific keywords on a search engine such as Google Scholar, I chose to instead search for publications in top ranked journals in the areas of education, gender studies, and citizenship studies. A list of the top 65 journals from each field was organized based on their rankings, according to: Scimago, Insite Journal, and Google Scholar Metrics. Then, a series of keywords were tested in each journal's search bar, including "citizenship education" in gender studies journals, and "gender" in citizenship education and education research journals. All the abstracts that appeared in the search results were closely examined, and from this list, the most relevant articles were extracted and inserted into a separate spreadsheet organized according to a list of themes.

Articles from this list were selected for the literature review based on their relevance to my research questions. Only studies that dealt with formal upper secondary education of young people, and had more than a passing mention of gender, were included. Only the articles based on the most pivotal publications and influential authors (i.e., Madeleine Arnot) were chosen, based on their relevance to the concepts of gender and citizenship in international comparative research. Even though there were several articles that delved into other paradigms of citizenship, such as sexual citizenship (i.e., Boryczka, 2009; McNeill, 2013; Carrara, Nascimento, Duque, & Tramontano, 2016; Illes, 2012), critical pedagogy (Wilkins, 2012; Seider, Tamerat, Clark, et al., 2017; Meyer, 2020; Mayo, 2013; Chan-Tiberghien, 2004); as well as a myriad studies of youth identity and subjectivity based on decolonial theory (Sabzalian, 2019; Howard, Dickert, Owusu & Riley, 2018), critical race theory (Garratt & Piper, 2010; Choules, 2006; Okello & Turnquest, 2020), intersectional research (Bondy, 2016; Francis, 2021; Yang, 2016; Moeller, 2021; Alemán, 2018; Love, 2017; Phoenix, 2009) and queer theory (Wells, 2017; Passani & Debicki, 2016; Cornu, 2016; Stucky, Dantas, Pochay et al., 2020; Pitoňák & Spilková, 2016); these were excluded based on their insufficient relevance to the analysis of the PISA assessments, or because they were not part of the pivotal gendered citizenship framework being explored in the following section.

In the sections that follow, the selected literature was reviewed revolving two themes: gendered citizenship concepts in gender education research, and previous analyses of data about gender equality in international large-scale assessments. These were relevant to my study because they helped establish the need for an explicitly defined and concrete theorization of gender-based citizenship that could serve as a basis for analysis of curricula such as the PISA 2018 Global Citizenship Framework.

4.1. Gendered Citizenship Education

My desire to establish a feminist theoretical position from which to evaluate the quality of gender assessments in PISA was challenged by the myriad analytical foci feminist education researchers used. Many studies have explored the inclusion of gender issues into civic education, but their focus on a particular gender issue in lieu of others has led to insufficient consistency. Russell, Lerch, and Wotipka (2018) conducted an international quantitative analysis of school textbooks, finding that the proportion of textbooks that mentioned gender-based violence (GBV) had grown in the previous sixty years from ten to thirty percent. In contrast, Bhog & Ghose (2014), as part of their curriculum study, participated in the creation of a feminist formal citizenship curriculum in India, yet decided to exclude GBV so as to not antagonize educators (pp. 58). This example highlights a contradiction between different education researchers. Even as some researchers attempted to systematically assess the state of feminist civic education by focusing on GBV, others resisted broaching GBV in their own curriculum. Since different researchers sometimes forego an alignment of focus in their study of international variance in feminist citizenship education, I searched for a theorization of gender-based citizenship that went beyond specific issues.

Madeleine Arnot created one such theorization in a pivotal publication that has influenced other researchers. A founder of the *Gender & Education* journal, and leading feminist citizenship education researcher (Abowitz & Harnish, 2006, pp. 667; David, 2015, pp. 931), Arnot wrote about postcolonial theories (Fennell & Arnot, 2008) and political agency (Arnot, 2009B), but it was her 2009A publication that summarized her “gender global gaze” (p. 130) approach. She expanded on Nussbaum's “global collective conscience from a gender perspective” (2009A, p. 122), under a “sisterhood” aware of race and class differences, and gender issues (i.e., gender inequity, sexual citizenship, global poverty, and GBV). I found her de-centering of Western subjectivity in favor of

local political contexts insufficient, and her abstract theorization of political action as global awareness about gender issues neglected gender-based motivation and concrete methods of civic participation. Arnot also erased trans identity in a discussion of sexual citizenship, mentioning only “sexual minority groups” (2009A, p. 126), focusing on reproductive rights, and describing gender discrimination as “homophobia” against “lesbians” (p. 127). Overall, she did not clearly define gender identity or formal civic participation, define gender citizenship education as competencies, outline their inclusion into formal education curricula, or determine how to measure student performance.

Issues with the application of Arnot's global gendered citizenship approach were visible in a qualitative study of citizenship in Arnot, Chege, & Wawire (2012). The researchers concluded from interviews that girls were concerned with participation in the private sphere, although some girls voiced concern for gender equality, and how GBV limited their public participation (2012, pp. 96, 99). This was in agreement with Fennel & Arnot's (2008, p. 520) attempt to portray the private sphere in political terms, but in contradiction to Arnot's (2009B, p. 242) call to not place women solely outside the political sphere. Although the study avoided the problematic Western “global” perspective, focusing on the concrete local context of Kenyan politics, it neglected concrete formal political participation. In their focus on the differences between gender groups, the researchers reinforced a gender binary, and also neglected to differentiate between young women's different degrees of participation in politics. Moreover, since Arnot did not fully define political agency as a competency, nor outline how to measure and include it in the classroom, many studies based on this paradigm did not adequately address these points. I evaluated these in the next paragraphs, disregarding those that were not relevant to gender (i.e., Kiwan, 2008, pp. 52; Chan-Tiberghien, 2004, pp. 195).

Many of the studies that followed Arnot's conceptualization shared this theoretical gap. Foulds (2014) studied gender responsiveness in the Kenyan curriculum, finding that certain images of female voters reinforced stereotypes, and that the suggestions in Fennel & Arnot (2008) needed a “clear framework explaining what gender responsiveness looks like” (Foulds, 2014, p. 668). Foulds mainly focused on an exploration of stereotypes, and although she called for better portrayals of women's political involvement, she did not address women's gender-based motivation. In a similar study, Carlson & Kanci (2017) analyzed gender regimes in textbooks in Turkey and Sweden. Their

critical evaluation of gender-based participation was particularly poignant in their distinction between the negative qualification of portrayals of women's participation that only served the state (i.e., motherhood, war, and labor), versus participation that served women's issues. However, the analysis was framed through a heteronormative subjectivity, was limited to comparison, and although they fully applied Arnot's call to recognize the role of women's activism (2009B), it did not explicitly describe a successful curriculum. Although both studies called for better portrayals of women's political involvement, they lacked quantifiable benchmarks to determine the degree to which curricula succeed at encouraging gender-based participation.

Other studies went beyond written curriculum to explore the discourse in the classroom. Tormey and Gleeson (2012) conducted a quantitative study in Irish schools, using the PISA sampling method. Their interpretation of Arnot's perspective was problematic. Their study did not originate in gender or global citizenship, these terms were retroactively applied during analysis, so students were not asked about gender equality, the term "Third World country" was used repeatedly, and women's holidays were excluded (pp. 634). Their Western-centric, heteronormative subjectivity was also evident in their interpretation of Nussbaum's global education as "globally sensitive patriotism" (pp. 629). They interpreted the global citizenship of Arnot & Dillabough (2000) as solidarity with the global poor, measuring how schools for boys and girls cared about the "Third World" differently, ignoring Arnot's repeated rejections of colonial discourse (Arnot 2009A, 2009B; Arnot & Fennel 2008). In spite of the attempt to set quantifiable benchmarks, the absence of issues related to gender equality, concrete gender-based political action, and critical reflection of gender stereotypes mirrors similar issues on other publications that followed Arnot's framework.

Gordon (2006) carried out a qualitative study of young women in Helsinki, focuses on embodiment in everyday school practices and student voice. She used the heteronormative definition of agency from the London Feminist Salon Collective (LFSC) (2004), or the "ability to make and carry out decisions," and applied Arnot only so far as to recognize school as a place where citizenship is formed. Gordon focused on the limitations put on girls' agency through the control of their bodies and behavior. Even though one of the participants mentioned the desire to vote unprompted, this was not included in the analysis, and Gordon did not frame identity in collective terms, nor explore students' gender-based motivation for political action. Gordon's treatment of

agency was abstract, without a concrete definition of agency as a competency, let alone its measurement or curriculum.

Gunnarson (2019) conducted a more gender-responsive qualitative study of norm-critical pedagogy based on classroom observations. She cited Arnot in her discussion of how citizenship should be gendered, intersectional, and not universalized, alluding to Butler and Foucault, and what she called feminist post-humanist theory. She described a successful curriculum as one “breaking against binary gendered and cis normative logic” (pp. 45). Although Gunnarson described certain principles for a gender-based curriculum; such as, that it be transformative, address how inequality shifts in contradictory ways, and suggest possible futures; she did not describe formal political action, but political embodiment in the classroom. Ultimately, as a qualitative study, it did not provide quantifiable benchmarks for the assessment of success.

Based on the application of Arnot’s framework to the studies in this literature review, it seems that a more specific theorization of gender-based agency is needed for the execution of more consistent feminist civic education research. My goal was to articulate a more explicit theorization of gender-based political agency, in order to define it concretely as a competency, enable its inclusion into formal civic education, and facilitate research. In the theory section, I conducted a disambiguation of feminist theorizations of agency. Before that, in the next section I revisited an important debate about the ability of international assessments to produce good quality data about gender.

4.2. Gender-based Competencies in PISA

The suggestion that large-scale quantitative student assessments could be a good source of data for research based on critical theory is highly controversial; in fact, the two are usually portrayed as diametrically opposed. Most international comparative research based on critical pedagogy, dependency theory, or postmodern theory, have applied the theories of Freire, Gramsci, or Foucault to analyses of neoliberal hegemonies of the institutions that conduct these assessments (Stromquist, 2005, pp. 92-93). Stromquist (2005) described how poor funding has limited international feminist research to low-cost textbook analysis (pp. 98), and was skeptical of the “transnational sense of purpose” of global citizenship (pp. 102), but did not provide alternatives to international assessments. There seems to be an implicit understanding that critical feminist research and large-scale quantitative studies cannot coexist, allowing the claims

made by these large-scale assessments to remain unchallenged. Most of the articles that use International Civics CS (ICCS) from the International Educational Assessments (IEA) (Lee, 2003; Blaskóá, Dinis da Costaa, Vera-Toscanob, 2019; Godfrey & Grayman, 2014) or PISA (Engel, Rutkowski, and Thompson, 2019) gender data employ quantitative methods, and neglect a serious inquiry into the gendered dimension of these frameworks.

In this section, I reviewed the few studies that have critically evaluated Organization of Economic Development (OECD) and the IEA assessments. In their evaluation of an IEA study, Hyslop-Margison, Hamalian, and Anderson (2006) found that the framing of research questions can limit knowledge production, like establishing unnecessary empirical connections or reiterating established concepts like gender equality (pp. 401). However, they did not explicitly examine the treatment of gender in the IEA assessments. Hooghe & Stolle (2004) applied participation theory to their quantitative study of student responses, with a specific focus on gender. Their frame of inquiry around a gender gap construct limited the data produced about young women's motivations. They recognized that girls participated more in social movements without identifying the movements, or their motivations; instead, the authors gendered the behavior, claiming that since girls did more volunteering and fundraising, these actions should be considered as important as radical action (pp. 16). Their essentialist claims that ethnic groups participate "less intensively" (pp. 14), or that girls practice "good" obedient politics, while boys practice "bad" radical politics (pp. 16), were inconsistent with recent feminist mobilizations around GBV and Black Lives Matter. Neither study adequately evaluated gender-based agency.

The studies that critically evaluated OECD assessments were more thorough, and yet neither specifically conducted a detailed critical analysis of gender in the PISA GCF. Unterhalter (2017) conducted an exhaustive critical review of list-based gender equality measures, such as the Millennium Development Goals, and even Nussbaum's gender capabilities approach. She criticized reductive neoliberal measures that ignored the broader and more complex concept of gender equality "linked with wellbeing, agency, aspects of embodiment and lack of violence, knowledge and criticality, public good, social relationships and context" (pp. 7-8). These also excluded aspects that might be "unmeasurable" such as "gender relations, sexualities, and aspects of power" (pp. 8), and more specifically, "categories of gender identity (trans- or intersex-) that cross or confound these divisions" (pp. 9). She called for measures that more seriously consider

agency, autonomy, and voice of women; and dismissed the OECD's gender indicators that met these demands because they were less impactful than PISA on international education policy (pp. 9-10). In spite of this consideration of PISA, and her highly relevant benchmarks for measuring gender equality, Unterhalter ultimately did not focus on critically evaluating the 2018 PISA GCF measures, since it did not yet exist.

Ledger, Thier, Bailey, and Pitts (2019) more closely examined the OECD GCF that informed the 2018 PISA surveys, and yet they did not examine gender. They conducted a discourse analysis, concluding that, in spite of its claims for diversity, the framework constructed a privileged elite subjectivity, and cited material from a few of experts from the UK and the US, creating of a neocolonial hegemony. They specifically called for an analysis of the 2018 PISA global competency framework (pp. 34). It is clear from the articles Unterhalter (2017) and Ledger, Thier, Bailey, and Pitts (2019), that a study that not only examines PISA discourses of global competency, but also questions its measures of gender equality with serious regard to the “unmeasurable” concept of agency, is necessary. In the next section, I review the theorizations of gender-based agency, and fill-in the gap that has been created by the misalignment of agency theories from different fields of study.

3. THEORY

In order to conduct an analysis of gender concepts in large-scale international student assessments, it is important to apply a concrete framework for gender citizenship education concepts that can be measured as competencies. In this section I sought to build upon Arnot's and Ultehalter's proposed frameworks, and include a trans-inclusive understanding of gender, and an epistemological approach that allowed an appropriate empirical analysis based on this ontological stance.

There is an ongoing debate in citizenship education studies about the ontology of agency and the human subject that affects each concept of my research question. Ontology is the framework that the researcher applies to her understanding of reality and the "nature" of the human subject (Denzin & Lincoln, 2001, p. 52), while epistemology is her perception of the possibility of producing "truth" about reality (p. 56). Liberal researchers recognize a genderless, or universalist ontology of the citizen as both universal (Sant, 2019) and individualized (Arnot, 2009A), and take a positivist epistemological stance where "truth" can be determined through quantitative research (i.e., Oberle, 2012). Feminists regard the universal subject as a stand-in for white European men of the patriarchal imaginary (Erel, 2010), and call for particularity, or the emphasis on the differentiated experiences and collective identities of the oppressed (Erel, 2010; Sant, 2019). Postmodernist feminist researchers further theorize that identity is relative, and that the subject's identity exists in the social imaginary, dismissing the possibility of truth that is not socially constructed (Butler, 1999). Rather than taking a strong position, I chose to interpret gender theory as a space of continuous struggle, recognizing these important ongoing debates.

As an intersectional feminist researcher, I struggled with the contradiction between the postmodernist and emancipatory aspects of my academic stance, so I took a complex approach. This is a common experience for many feminist researchers, who are known to employ the full gamut of "constructivist, critical theory, feminist, new materialist, queer, and critical race theory" and are at the border of post-positivism and poststructuralism (Denzin & Lincoln, 2001, pp. 17). I wanted to improve data collection about young women's citizenship education, yet I took a deconstructive perspective which directly contradicted this stance. I was guided the practical need of feminist policymakers to leverage numbers to expert bureaucratic power (Springer, 2020); thus, I

supported the production of better quantitative data because it might create education policies that improve feminist citizenship education. At the same time, I employed a method of discourse analysis based on postmodern theories. I struggled with Foucault's rejection of state surveillance, or data as a tool of behavior control, and attempts to define the truth on behalf of "utopian" political goals that serve one group, and disadvantage others (Dean, 2009, pp. 83-87). I sought to resolve the ethical danger in producing "truth" about young women's political agency.

This danger involved the need for a clear definition of the gender identity of the citizenship subject being studied. The ontology of gender identity is at the core of feminist and queer debates. A key ethical concern in qualitative research is to avoid the further marginalization of the research subjects (Creswell, 2015, pp. 555). As a cis-gendered researcher, I had to be mindful of my privilege, and of the violence that cis-gendered academics have historically perpetrated against transgender people in gender research (Namaste, 2009). In advocating for young women, I risked contributing to an oppressive narrative of cis-gendered, heteronormative womanhood. Transgender, queer, and gender non-confirming students who are already excluded from data production at the OECD might be further marginalized by a poorly constructed definition of gender-based political agency.

In the following section, I mapped the theoretical debates in gender theory that have shaped this theoretical chasm between postmodernist relativity and emancipatory feminist materialism. In the ontology section, I identified a theoretical approach that came close to resolving this dilemma while at the same time minimizing the possible ethical repercussions of an emancipatory pursuit. Rather than taking the postmodernist dichotomy at face value, I re-framed it as a debate about the ontology of gender identity. Outlining an explicit ontological stance on gender identity was necessary before critically evaluating whether the 2018 PISA GCF could quantitatively measure data about ontologically ambiguous subjects.

3.1. Resolving the Ontological Issue of Gender-Based Political Agency

In a pivotal article in a leading education journal, Eteläpelto, Vähäsantanen, Hänn, and Paloniemi, (2013) claimed to have found a cure for research stagnation caused by

postmodern rejection of truth, naming it the Subject Centered Sociocultural (SCSC) approach to agency. Their multidimensional approach provided a concrete and practical framework for agency research, resolving the ontological conflict between postmodern theories of socially constructed identity and critical realist theories that maintained the existence of a pre-discursive self. Their solution was centered around feminist debates about realism and postmodernism, in combination with Margaret Archer's critical realism, and her three levels of reality: natural, involving embodied knowledge; practical, involving practical knowledge; and social, involving discursive knowledge. Life-course theory served to provide a temporal dimension to this model. The authors described agency as a mediating mechanism through which individuals could manage their relationship between social suggestions and individual desires defined by a pre-discursive self.

The dismissal of "strong" poststructuralist theories was at the center of agency constructs in the SCSC approach. Eteläpelto, Vähäsantanen, Hänn, and Paloniemi, (2013) positioned Butler and Foucault as "radical," citing feminist criticism of "strong" poststructuralism as deterministic, "divorced from social reality," positioned against "emancipatory" feminism, prone to "political nihilism and fatalism," unethical, and leading to "political and ethical paralysis" (p. 52). They cited McNay (2004), who had a subject-centered focus on embodiment, and who dismissed Butler's concepts as structural abstractions; and also, Clegg (2006), who saw poststructuralism as incompatible with agency, since its unintelligible narratives dissolved humanity into apolitical, disembodied textualism. This misleading portrayal of poststructuralism as antagonistic to agency theories and emancipatory feminist politics disregarded the centrality of queer theory, and its emancipatory aims, in the work of both Butler and Foucault. Most importantly, it left open questions about the placement of queer or transgender identity in agency constructs not based on postmodern theory. In this theory section, I sought to address this gap through a re-framing of a realist-postmodern debate as a conflict between cisnormativity and trans identity.

The London Feminist Salon Collective (LFSC) originated the feminist compromise on agency, forming "a 'viewpoint' for the journal *Gender & Education*" (pp. 26), which rejected gender ambiguity based on a misreading of Butler's work as nihilistic and threatening. Butler's prediction that she would be seen as threatening feminism (Butler, 1999, vii), and that her work would produce panic (ix) was reflected in the LFSC's fear of the "fragmentation/dissolving of the women's movement," and "retreat

into epistemic communities” (LFSC, 2004, pp. 31). Accusing Butler of nihilism meant erasing her involvement in queer, postcolonial, and race studies (Butler, 1999, ix), and her goals to discredit the delegitimization of queerness, challenge homophobia, and address a “crisis in ontology” related to trans and lesbian identities (pp. xi). The misrepresentation was glaring enough for some LFSC members to feel an obligation to “defend” Butler from being “misinterpreted,” yet in response, one presenter mentioned that these discourses should be resisted even at the risk of normativity (LFSC, 2004, pp. 27). The fear of queerness was evident from the suggestion that diverse gender identities threatened feminism, as well as from the dismissal of queer emancipatory goals. The LFSC treatment of Butler, and her goals, was thus transphobic and homophobic.

The unacknowledged cisnormative bias in the LFSC was significant given its influence on the SCSC approach to agency. The gender, race, sexuality, and number of LFSC members were omitted in the text. It was mentioned that “the participants in the salon were all women” (LFSC, 2004, pp. 28) in a cisnormative discussion about the dueling “men” and “women academics.” One presenter wondered, “can *we* ever speak ‘for’ *other* women (e.g., across class, ethnicity, sexuality)?” (p. 27; emphasis added). I interpreted the “we” as a Western, white, and cis-gendered female point of view. Given this privileged position, and the evidence for transphobia, I found it disconcerting that Clegg (2006) described her theorization of agency as a direct continuation of the LFSC (2004, pp. 309), and narrated a history of feminism that excludes queer movements (pp. 312). Since Clegg’s critique of postmodernism was ingrained in a history of feminist cisnormativity, I considered that my application of the SCSC needed to be adjusted for cisnormative bias.

Cisnormative bias seems to be marked by insufficient consideration of trans experiences. For example, Archer (2000) did not consider trans identity in her work, and held a leadership position in a religious institution that persecuted Butler (Francis, 2017; Gessen, Shteir & Mishra, 2020, February 09). While trans academics ask the research community to consider the effect of mortality rates on their work (Pearce, 2020) the LFSC (2004) members discussed agency, buoyed by the feeling of “safety” and “even the beneficial effects of alcohol!” in a discussion that was “enjoyable, stimulating, amusing, and even liberating” (pp. 25-26). Overall, cis academics question concepts obvious to people with delegitimized identities (Butler, 1999, pp. viii), because they do not engage with trans experiences with violence, criminalization, mortality, legal recognition, or how histories of slavery shaped medical definitions of gender (Snorton, 2017).

Cis-feminist theorizations of trans people often involve “hostility,” otherizing, objectification, denial of agency (Heyes, 2013, pp. 211), failed attempts to globalize (Halberstam, 2016), centering of anglophone white subjectivity (Bhanji, 2013), and intersex erasure (Whittle & Turner, 2016). I therefore find it problematic to fail to consider the embodied experiences of trans people when theorizing on gender-based identity.

In an attempt to consider these realities, I reviewed trans studies literature, and found many ontological approaches to gender. Sandy Stone founded trans studies, combining postmodernism, intersectional brown feminism, queer theory (Stryker, 2008), and Foucauldian genealogy; she described gender as the discursive act of *passing*, being perceived as cis, and the trans body as a battlefield for medicine, feminist rage, and trans experience (Stone, 2006, p. 230). C. Riley Snorton (2009), a Black “pre- and non-operative, no-hormone transsexual” (pp. 82), broadened *passing* to include a psychic dimension, a material, pre-discursive trans body (physical) and identity (psyche) (pp. 87) independent of social recognition (pp. 87, 89). Transgender activists, facing increased violence, grew suspicious of academics, particularly Butler’s abstractions, and called for relevant and embodied research (Halberstam, 2016; Namaste, 2009). Some trans theorists viewed gender as material, some as embodied, some as discursive, and others as all these; thus, trans studies seemed to be a thriving field, not a unified theory.

Certain feminists and trans theorists describe positions that allowed me to incorporate trans studies to the SCSC approach. Grace Lavery (2019) conducted a Foucauldian genealogical review of trans identity, finding that all theorizations of trans people, either materialist or discursive, were delegitimized, concluding that trans identification was a political act of survival, not an academic theoretical inquiry. Sebastian Jansen (2016) rejected the agency dichotomies of Archer’s critical realism and Butler’s post-structuralism, describing the latter as mind essentialism; ultimately, intersex and trans people seek legitimacy, freedom from categorization, and the right to self-determination (pp. 127, 134, 136). Zimman (2019) suggested that political self-determination is limited without institutional recognition of trans identity. I felt that these practical understandings of trans identity as a political matter of self-determination could be applied to the SCSC approach. Tuijja Pulkkinen, a feminist philosopher, described identity as a social movement. To her, agency meant viewing gender identity as a political question; agency is a specific yet mobile location; a self-reflective, environmentally-responsive nationhood that has agency enough to look at itself with particularity, and question the

nature of its construction (Pulkkinen, 2000). So gender-based agency could be considered a political matter determined by individuals and social movements who demand state recognition.

While the ontological debate of gender identity may not be easily resolved given the diversity of perspectives in transgender studies, I can attempt to adapt these viewpoints into the SCSC approach to agency. The SCSC approach established a model for the research of agency, but because it is based on a cisnormative feminist compromise on gender identity, I decided to adapt it to integrate transgender theories. In this thesis I understand agency partly through the lens of SCSC approach, but integrating the right to self-determination of people with diverse gender identities, and demanding that they be recognized by institutions. Gender-based political agency is the way that we look at the particular role our gender identity plays in our lives as a political agent that we can influence, and that determines the nature of our public existence. The level of our awareness about this political agent that mediates our experience with politics might influence our ability, desire, and actions we take to participate politically. The development of gender-based political agency is the process of forming one's particular gender-based identity and being willing to act on behalf of gender equality. The next section explores how it might be evaluated as a competency.

3.2. Resolving the Epistemological Issue of Gender-Based Political Agency

My purpose in deconstructing the quantitative data produced by the neoliberal regimes of knowledge in international assessments was not to end data collection, but to improve it. This objective was based on a practical understanding that the production of better quantitative data that serves the need of feminist education policy leaders is necessary in order to influence national educational systems and policy debates (Unterhalter, 2017, pp. 9-10). Springer (2020) described the “paradox of quantified utility,” in which feminist either gather quantitative data or exhaust themselves trying to fight it. Feminists operationalize quantitative data in practical political strategies to “overcome the emotional fatigue and structural marginalization they endure” in their pursuit of genuine gender equality. In the “age of evidence-based decision making,” data is needed to get resources, and those who “adopt quantified knowledge production” are rewarded

(Springer, 2020, pp. 73). Thus, based on the current context of evidence-based decision making, and considering the practical needs of feminist policy makers, I considered it more appropriate to call for an adjustment of the data being produced about gender equality, than to completely deconstruct and dismiss it.

Before pursuing this objective, I had to address whether it was possible to quantitatively measure gender concepts related to ontologically ambiguous subjects. Unterhalter (2017) described this enigma as a “tension between what is easily measurable, but may not be significant, and what is of major importance, but cannot be measured (p. 2).” She concluded that in some instances, the benefits of an inaccurate, yet reasonably approximate data overcomes the political consequence of being erased from data measures. In such a situation, data can be “used to stand as [proxy] for complex relationships that are really unmeasurable” (p. 2). In my case, I took a calculated wager in this thesis that the cost for marginalized groups of inaccurate, but approximate, portrayals in international assessments is lower than the cost of a total lack of representation in data measures. My hope was that this compromise could serve the emancipatory aims of feminist research, while lowering the risk of oppressive consequences to marginalized LGBTQ groups (Dean, 2009, pp. 83-87).

3.3. Framework for Gender-Based Political Identity, Agency, and Action

To answer the first research question, how the 2018 GCF regulated gender subjectivities, I made judgments about the quality of their portrayal of students in comparison to a gender-based curriculum based on gender research. To answer the second research question, whether the 2018 GCF measured the gender-based political agency or action, I also outlined a set of benchmarks for high gender-based political agency and participation. Both questions were based on certain assumptions. I posited that the authors of PISA would limit gender identity, and that some students and some type of information about them, could be excluded from the assessments. Since I expected the frameworks to be lacking in feminist agency, I needed to concretely describe a successful gender-based curriculum (Foulds, 2014, pp. 668) to help me identify abstract, implicit, or “missing” information. The main theory I applied was the SCSC approach of

Eteläpelto, Vähäsantanen, Hänn, and Paloniemi, (2013). This approach has a multi-dimensional ontology with three levels of reality (material, practical, and discursive), a temporal dimension (employing life-course theory); where agency acts as a mediating mechanism between the social and pre-discursive selves. In this section, I outlined the theoretical framework.

My model for a gender-based citizenship curriculum originated in various streams of feminist citizenship education literature. The understanding of gender equality includes wellbeing, embodiment, social relationships and context (Unterhalter, 2017, pp. 7-8), and takes the agency, autonomy, and voice of women seriously (pp. 9-10). It centers local political contexts, not a Western viewpoint (Fennell & Arnot, 2008), and promotes the awareness of race, class, and gender issues such as gender inequity, sexual citizenship, global poverty, and GBV (Arnot, 2009A); however, it prioritizes formal political participation over abstract awareness. The curriculum is transformative, and deconstructs binaries and cisnormativity; it portrays citizenship identity not as a universal, but as gendered, intersectional, context-specific (Gunnarson, 2009), explicit, and particular, without erasing any identities (Erel, 2010; Sant, 2019). Citizenship identity is also collective, and based on social movements, or specific, shifting locations regulated by activist discourse, collective self-awareness, and responses to political environments (Pulkkinen, 2000, pp. 137). Thus, it is intersectional, inclusive, and particular.

The inclusion of trans identity and the priorities of the field of transgender studies is particularly important. A gender-based citizenship curriculum acknowledges all levels of identification that a trans student might have, such as pre-discursive physical and psychic gender identity, and social identity influenced and regulated by their embodied experience of *passing* (Snorton, pp. 87, 89). Since theoretical disagreements about the nature of the gendered self are conflicts over recognition of queer, transgender, and intersex identity (Jansen, 2016) the curriculum should provide institutional recognition of self-determined identities (Zimman, 2019). This concept should also be flexible, taking into account the transitional aspect of identity and agency normal for young people as they develop their understanding of the constraints of the world and their own rights and freedoms. This means that categories of gender, like “queer,” “trans,” “female,” “feminist,” or “gender non-conforming” should be acknowledged and welcomed. Most importantly, identity concepts should be determined through an inclu-

sive and open dialogue with academics, activists, and members of civil society who represent these groups. Inclusivity in this way extends not just to specific identity terms, but to the communities that establish them.

The specific barometer employed in this thesis to judge whether the political action is being assessed is “gender-based” originates in the concept of gender-based voting developed by Holli & Wass (2010). This concept was defined as “voters preferring and actually casting their votes for candidates of their own gender in elections” (pp. 599). Gender-based political agency and participation is thus defined as a current or expected action of formal political participation of the student that is taken in the interest of their own gender group, and on behalf of gender equality. Another aspect involves its definition as a competency to be developed through formal education. The curriculum should be concrete, explicit, action-specific, and based on a specific understanding of political action made explicit in relation to gender equality at an individual and collective level. While understandings of political action take informal action into account, these understandings were out of the scope of this thesis. A gender-based curriculum is thus built around specific formal political actions that students might take for gender equality in a way that is specifically attached to their own gender identity.

4. RESEARCH PROBLEMS

4.1. Problem Statement

The low-level of global participation of women in legislatures is a big problem. Only 25% of all legislators are female, and there is a significant variance between different countries (IPU, 2020), so international comparative research is needed to understand this phenomenon. Some countries have made commitments to meeting SDG indicators on gender, such as increasing the participation of women in leadership (Target 5.5), and international comparative research needs to hold them accountable for the extent to which their civic curricula serve these commitments. However, gender does not feature as a central focus in critical educational research on citizenship education (Sant, 2019), reflecting a general trend in data collection that excludes the priorities of women (Perez, 2019). Insufficient research has been conducted to hold governments accountable for the extent to which their civic curricula prepares citizens to fulfill their commitments to gender equality.

Large scale international assessments are some of the primary agents that hold governments accountable for skills development, and it is important for researchers to hold PISA accountable. The Organization for Economic Cooperation and Development (OECD), which creates and coordinates the implementation of the PISA exams, is an organization focused on the economic development of its member nations. The purpose of PISA is to work as a GPS for the education policy direction of countries, and help them determine their needed areas of development (OECD, 2016). Some worry about assumptions made in PISA frameworks about which competencies are important to measure, which influences how they are prioritized in educational systems. Recently, PISA adopted a Global Competencies Framework (GCF), which they connected to SDG target 4.7: improving student skills related to promoting the SDGs related to, among many issues, gender equality (OECD, 2019a). Assessments for civics competencies are more difficult than for math or science, because civics curricula are affected by different political systems, do not focus on the same skills, and are not equally explicit. The civic competencies related to gender equality did not seem to be as explicit in the GCF as the PISA math skills (OECD, 2019a). It is important for research to hold PISA accountable for its assessment of gender-based competencies.

Many studies have been critical of the PISA or the GCF, but so far, none have conducted an in-depth analysis of the gender-based subjectivities or competencies in the GCF. Much of the critical research evaluating international assessments have focused on questioning the overt neoliberal focus on economic competencies of the PISA framework, and the undemocratic method through which it is elaborated (Liesner, 2012; Meyer, 2014). Some have questioned why the framework is insufficiently aligned with the Sustainable Development Goals (SDGs) (Vaccari & Gardiner, 2019), and others have questioned the Eurocentric perspective of the so-called “intercultural” elements of the global competencies related to inclusion (Simpson & Dervin, 2019). Engel, Rutkowski, and Thompson (2019) mentioned gender as one of the identities within the fourth key target dimensions of the GCF, but gender did not play a relevant role in their analysis. Overall, the research about gender-based identities and competencies in PISA has been inadequate.

This gap is significant because of the high influence that PISA has on education policy and research. As international assessments have become the main way through which countries have been made accountable for their education quality, they have also jettisoned research and policy efforts to improve it. PISA is the largest international exam used to assess student performance, and it has become, more than a tool, an international mechanism of influence that works as the standard for international education quality, and a catalyst for reform. This phenomenon is sometimes labeled as PISA shock, considered by some to influence the process of knowledge creation through its own hidden curriculum. The overall consensus in the field of educational policy is that the PISA Framework is highly influential in international policy, educational discourses, and educational research (McGaw, 2008; Neumann, Fischer, and Kauertz, 2010; Carvalho, 2012; Pons, 2012; Bonal & Tarabini, 2013; Rutkowski, 2015; Cox & Meckes, 2016; Pons, 2017; Sellar & Lingard, 2013). The important question at this time is not “if” the PISA framework acts as a hidden curriculum, but “how” this curriculum is shaping important aspects of young people’s development, such as their citizenship identity and civic competencies.

It is unclear to what extent the 2018 PISA GCF utilized gender and feminist research to determine the most important competencies that students need to develop in order to address gender inequality. My goal in this thesis was to apply gender theory to evaluate the quality of the 2018 PISA GCF, because improving the assessment of gender-based political agency could ignite PISA shock, and influence education policy. I

applied the theories of gender and feminist research to the analysis of gender-based political identity, agency, and participation. In other words, my goal was to illustrate how data production generated discourses about gender-based political agency that might shape policies about civic education for gender equality.

Both research questions fill the research gaps that were made by the insufficiently focused gender-based evaluation of the PISA GCF. To answer the first research question, I applied gender-based theoretical framework to evaluate the portrayal of gender-based identities and subjectivities, and understand the extent to which they recognized collective feminist, racial, or LGBTQ identities. To answer the second question, I again applied gender-based theoretical framework, and the benchmarks for political participation outlined in the previous section, to evaluate the extent to which political agency and participation was based in gender and related to gender equality.

4.2. Research Questions

The research questions were as follows:

1. Through what kind of discourses or regimes of knowledge did 2018 PISA CGF regulate gender subjectivities in their assessment of citizenship competencies?
2. In what ways, if any, did the 2018 PISA GCF measure the gender-based political agency and participation of students?

5. RESEARCH METHOD

This thesis consisted of content analysis of the 2018 PISA GC Framework and Questionnaires, which I treated as policy documents, or qualitative data. I chose a mixed-method study design because I sought to critically evaluate the concepts promoted by the institutions behind these international assessments. My research objective was to identify the hegemonic discourses that framed gender-based citizenship, and how these discourses were used by institutions to regulate gender identities of students and define the realm of political possibility available to them. My approach emerged from the postmodern and critical feminist paradigms, which are critical about the type of data it is possible to collect about gendered subjects. A positivist quantitative analysis would have only reinforced discourse. I considered a mix of Foucauldian critical discourse analysis and quantitative content analysis to be the most appropriate methods to study the influence of relations of power on data production in international education policy.

The most influential approach in this thesis is within the field of Sociology of Education Policy, which is focused on the connection between education, governance, and relations of power. Many studies have focused on applying Foucault's concept of governmentality to the study of education policy (Popkewitz & Brennan, 1997), especially in the analysis of the influence of neoliberalism and international institutions on education (Peters et al., 2009, p. xxvii). I applied the tool of analytics of governmentality, a Foucauldian approach to discourse analysis as interpreted by Mitchel Dean (2009), which has been applied to a few studies in education (Fejes, 2009; Kessler, 2009; Ball, 2013). I also applied Stephen Ball's (2013, 1994) approach to Foucauldian genealogical analysis in education policy. I understood genealogy and the analytics of governmentality as analytical tools (or "methods") that are forms of Foucauldian discourse analysis grounded in the study of neoliberalism. In this section I first made some necessary methodological clarifications; then, I described the method of data collection; and finally, I described the method of data analysis.

5.1. Methodology

The identification and application of a well-established method is a basic requirement for qualitative research, but the primary purpose of Foucauldian methodology

is to disrupt research through resistance to pre-determined concepts. In Foucauldian research, one should resist definition to prevent methodologies from “congeal[ing] into a kind of dogma,” avoid applying the methods of others, and instead “borrow, fashion, and refashion” them in relation to one’s own context (Dean, 2009, pp. 25, 58). I navigated the delicate tension between resistance to definition and empiricism by finding agreement about the existence of certain analytical tools, methods, or problematizations, such as: archaeology, genealogy, critical discursive critique, and governmentality (Ball, 2013, pp. 16, 18, 26, 28; Hook, 2007, p. 3). I was thus pushed to concretely define tools according to my own interpretation, and adapt them in relation to the social environment, the research questions, and empiricism.

Of all the Foucauldian tools, the genealogical and linguistic approach to discourse analysis have both been well-established in education research. Foucault’s genealogies of the human subject and analysis of discursive practices have helped researchers problematize educational concepts and institutions (Peters, 2007, pp. 181). Given the relative complexity of genealogies, however, I considered applying critical discourse analysis (CDA). Much of the understanding of CDA emerged from the work of Norman Fairclough (2014) in the field of critical linguistics, and though some of its traditions were relevant to my study (Lazar, 2005; Wodak & Corson, 2012), these required sociolinguistic expertise. Some educational researchers have stripped these theories of linguistics, applying their political elements to the analysis of international education policy (i.e., Olssen, Codd, and O’Neill, 2004; Harden, 2009; Ledger, Thier, Bailey, and Pitts, 2019). I chose not to take a linguistic approach due to my lack of linguistic expertise, and some weaknesses in non-linguistic political CDA approaches.

I sought to avoid the common mistake of misinterpreting discourse analysis as language analysis, while ignoring power and history. Derek Hook (2007) made a distinction between the different foci and characteristics of authentic and inauthentic approaches. Authentic Foucauldian analysis is focused on the material consequences of the historical creation of systems of thought, and the contextualization of external social conditions under which statements are recognized as truth. Text-based analyses have an internal textual concern with disembodied semantics, or “structural linguistics, deconstructionism and semiotics,” ignoring the social dynamics of power and forming a “decontextualized set of hermeneutic interpretations” that can be easily dismissed. Hook saw genealogy as epistemologically stronger because of its socio-historical foundation,

which moves it beyond linguistics, into the formation of a discipline: as a scholarship (i.e., sociology, biology), and an institution (i.e., prison, school). (pp. 5, 132-133) I thus interpreted genealogy as a more authentic, rigorous, and historically focused version of Foucauldian discourse analysis.

One study of the OECD GC framework illustrated the weak power analysis of text-focused CDA. Ledger, Thier, Bailey, and Pitts (2019) seemed to apply the CDA method based on Fairclough in a reductive manner, as a tool of language analysis rather than discourse analysis.¹ They stated their purpose was to “make meaning of text,” deconstruct language, and “acknowledge the power of language within the construction of policy documents” (p. 7), rather than exploring the power dynamics that formed the discourse. This emphasis on meaning also contradicted the authors they cited. For example, Stephen Ball (2015), declared that researchers failed in CDA when they misinterpreted discourse as the meaning in text rather than the regulation of text, and when they assumed that critical detachment and language analysis was sufficient to access it (p. 311). This failure was evident in Ledger, Thier, Bailey, and Pitts (2019), since they were concerned with interpreting underlying meaning through keyword analysis. In one case, they conflated the importance of diversity as the intended meaning of the authors, and also the results of CDA (p. 26), reducing CDA to interpretation. To me, this further confirmed the superiority of genealogy, which focuses not on meaning but on methods of institutional control.

Given these advantages, the first method I applied for discourses analysis was genealogy. I considered genealogy most appropriate for my research questions because it explores the power dynamics in the creation of discourses and regimes of knowledge, which is crucial to understanding how international assessment frameworks create and regulate gender subjectivities. Foucault did not offer a methodology of genealogy, but “a methodological rhythm” of “defamiliarizing procedures” (Hook, 2007, pp. 138, 172). Following this rhythm through close readings of Foucault’s work, I sought to defamiliarize the universal or heteronormative gender subjectivity that is taken from granted in the GC Framework. My application of genealogy was based on Stephen Ball’s (2013) understanding of genealogy as an ontology of the present that destabilizes historical continuities, uncovering both the visceral effect of discourse on the human body and the displacement of the subject. Though this tool, I troubled the hegemonic gender histories

¹ They also distanced their work from Foucault, and applied the theoretical framework of one of their own authors (Thier, 2015), who did not acknowledge his hidden bias as an employee of the International Baccalaureate (IB), an arguably neoliberal institution.

of these regimes of knowledge through a “[retelling of] the histories of disciplines, institutions and practices drawing on excluded and hidden texts” (pp. 33-34). Genealogy helped me identify the regimes of knowledge that established the hegemonies of certain gender identities, and explore alternative, subjugated, and excluded discourses.

I adapted this genealogy in a way that was appropriate for a master’s thesis, reducing the scope to a “micro” genealogy of the gendered subject of global competencies in the 2018 GCF. Through this micro-genealogy of subject positions, I explored the formation of discourses that regulated gender subjectivity and gender-based political agency and action. This was not limited to the text, but also to the scholarly debates around the formation of the text, and the material consequences of these struggles on the gendered bodies of both students and researchers. It was “micro” because it was not a complete historical narration of the formation of gender concepts in the West, but rather a limited historical genealogy of the gender subjectivities that were excluded from the PISA 2018 GC framework. In my genealogical analysis, I focused on how the variety of discourses came to be formed, the systems of constraint to which they were related, and their specific norms, or conditions of appearance, growth and variation. The idea was to identify the rules used to decide which types of discourses and subjectivities were allowed to take part of the PISA GC framework, and which were not.

The second method I applied was a category count method, which was integrated with the micro-genealogy. I used quantitative data to measure the relevance of discourses and connect them to a community of authors. I employed a process of content analysis called category counts, in which words are classified into categories and then counted, under the assumption that a higher relative count (percentage) reflects a “higher concern with the category” (Weber, 1990). I used descriptive and inferential statistics to illustrate general trends and patterns in the data, and to check the significance of my findings. While Ledger, Thier, Bailey, and Pitts (2019) also used quantitative data to triangulate their CDA of the OECD framework, their Social Network Analysis (SNA) used to identify significant communities and brokers within the cited authors (p. 2), did not sufficiently address relations of power. SNA only illustrated “what is happening in relationships” (p. 8), limiting their conclusion to, for instance, assessing whether government agency publications were more salient than peer-reviewed journals, without identifying different discourses, nor their level of influence. The category count method of content analysis deepened the information about the relations of power.

The third method I applied was the analytics of governmentality, which I used as a framework to guide my analysis. This tool was most relevant to my research questions due to their focus on governance, regimes of knowledge, and the regulation of discourse. This tool explores how human populations are governed within regimes of practice. Regimes of practice are “the organized practices through which [humans] are governed,” and the specific conditions under which these different regimes, and their subjects, emerge, continue to exist, and are transformed (Dean, 2009, pp. 59, 63). Foucault’s inquiry of governmentality observes power not as top-down domination, but as the plurality of relations of power, or discourses, at all levels of government that shape the conditions for human behavior through the economy, or most efficient and rational management of the population (Dean, 2009, pp. 45, 59; Foucault, 2007, p. 201). Like many in governmentality studies, I interpreted this tool as a line of questioning in discourse analysis that explores neoliberalism (Peters et al., 2009). I followed its structure to guide the micro-genealogy and organize my analysis and presentation of results.

I diverged from Dean’s take on governmentality by incorporating a feminist lens. As was mentioned in the theory section, I applied the perspective of the SCSC approach to identity, which allowed me to use of Foucault’s work as a middle ground between extreme relativism and emancipatory work. In fact, many authors have questioned the view that Foucault’s concept of discourse devolves into extreme relativism (Hook, 117; Olssen, Codd, and O’Neill, 2004, p. 33-36). Peters et al. (2009) described Foucauldian analysis as contextualism rather than textualism, because it is situated within the local sociopolitical context. Dean’s (2009) interpretation deviated from this view, since stating “utopian” political goals to be achieved through better governance by any group risks further marginalizing other groups (p. 83-87). I took the epistemological position that applying a feminist lens to governmentality analysis, while embracing emancipatory political ideals of equity, can be authentic to Foucauldian work if I use self-reflexivity to take ethical risk of this position into consideration.

5.2. Ethical Concerns

Since my data collection did not involve ethnography, interviews, or direct interactions with students, my biggest ethical concern was about positionality during data analysis. I carefully considered the influence of my positionality on my research, or how

my social identity as a cis-gendered, heterosexual, feminist Woman of Color shaped my research approach, and my motivations to conduct this study. This was most relevant in my definition of the theoretical concept of gender-based agency, and the methodological lens I employed to analyze this concept. My struggle with the contradiction post-modernist and feminist emancipation resulted in theoretical and methodological developments that allowed me to navigate this intersection in a more ethical way.

Theoretically, I rejected how Eteläpelto, Vähäsantanen, Hänn, and Paloniemi, (2013) portrayed Butler and Foucault as poststructuralism as deterministic, and re-framed them a realist-postmodern debate as a conflict between cisnormativity and trans identity. To avoid the ethical risk of marginalizing my research subjects (Creswell, 2015, pp. 555; Namaste, 2009) while defining a truth about “girls” to serve utopian feminist goals (Dean, 2009, pp. 83-87), I avoided a narrative of cis-gendered, heteronormative womanhood. Instead, I defined the subjects of gender-based agency as ontologically ambiguous. Methodologically, this allowed me to take a feminist stance in my analysis, rather than a fully deconstructive, or deterministic, stance. In other words, this approach created space for me to imagine a better future for data production. It allowed me to critically analyze current approaches to data production, without eliminating the possibility for future international large-scale assessments to gather better quality data about ontologically ambiguous subjects. Overall, it allowed for an ethical emancipatory stance that did not devolve into determinism.

5.3. Data Collection

5.3.1. Material Selection

In discourse analysis, there is a distinction between the collection of material, and the collection of data. To be converted into data, materials go through a process of consideration, organization, and selection that is influenced by the theoretical approach as well as the themes of the study (Taylor, 2001, p. 24). The data sample in discourse analysis consists of the relevant data that is extracted (i.e., the most relevant key words and text passages) after the material has been filtered. The materials in this case were the frameworks, questionnaires, and bibliography, as well as a small sample of the most relevant books, policy texts, and articles selected from the bibliography. The materials were selected based on their relevance to gender-based identity, agency, and political

participation. As stated previously, the PISA GC assessment materials were selected because PISA is one of the few international student surveys, and the one with the most power over education policy. The texts analyzed were the following:

- PISA 2018 Global Competence Framework: Text (OECD, 2019a)²
- Bibliography Sample
- Author Sample
- Cognitive Test Questions: “Illustrative examples of scenarios for the cognitive assessment of global understanding” (OECD, 2019a, Annex 6.A., pp. 210-211)
- Student Questionnaire Questions: “Questions related to global competence in the student questionnaire” (OECD, 2019b, Annex C, pp. 50-61)

The 2018 Global Competence Framework Text is the section of the PISA 2018 framework which deals with global competencies. I sometimes refer to this as the GCF Text, the GCF, the 2018 GCF, or the framework. There were two versions of this text found in the OECD (2019a, 2019b) website, with some differences, which I discuss in the results section. The most relevant of the works cited in the bibliography (of both versions of the text) were selected for quantitative category count analysis, and I henceforth refer to this narrower number of works as the Bibliography Sample. This narrowed list of materials was further filtered for qualitative discourse analysis (as part of the micro-genealogy), selected based on their influence on the knowledge regimes relevant in the GCF Text. In the results section, I do not refer to this as a separate sample- I simply mention the works I analyzed. I also analyzed a selection of texts from authors influential to either the GCF or to different sub-fields of gender studies, and call this sample the Influential Authors Sample. This sample was also included in the micro-genealogy and category count analysis in order to add further depth. Lastly, I analyzed the questionnaire given to students (excluding the questionnaire given to teachers and school principals), and the sample of cognitive test questions included in the framework.

5.3.2. Data Collection

The first phase of the data collection process involved the selection of works for the Bibliography Sample. Rather than conducting an exhaustive analysis of every single

² There were a few versions published. Unless otherwise noted, references to the GCF Text (i.e., OECD, 2019a) refer exclusively to the original PDF version published in the following link: https://www.oecd-ilibrary.org/education/pisa-2018-assessment-and-analytical-framework_b25efab8-en

text that was referenced in the framework, which I considered to be beyond the scope of a master's thesis, I focused on the most significant texts in which the presence of gender was salient. I selected the data after several phases of coding in order to increase their relevance to the research question. During the first phase, I transferred the complete list of references to a spreadsheet organized in alphabetical order, categorizing it into color-coded groups labeled based on discipline or field of study, theme, and type of publication. I categorized the sources by skimming their abstracts to extract the main points and relevant themes. For example, sources that covered the study of intercultural education were highlighted in orange, and sources that covered education research were highlighted in green.

In the second phase, I gathered data about the institutional weight of the whole population of sources cited in the bibliography. The purpose of this genealogical information was to illustrate the institutional relations that influenced the creation of the text. I collected information about its authors, and institutions providing funding. I recorded the level of education of the author, their university department of affiliation, and their field of study. To assess the influence of each source, I recorded the number of times an article was cited in Google Scholar, the journal in which it was published, the country that the journal was affiliated with, and the ranking of the journal according to Scimago. This information was available through the professional social media website, LinkedIn, or verified in their publications, or their university webpage. I then sorted and grouped these sources in different ways, and later identified the most salient categories they belonged to. These steps reduced the human error in the categorization of, for instance, determining which discipline was most salient to an interdisciplinary study.

During the third phase of data collection, I compared the spreadsheet codes to the 2018 PISA GCF Text. I cross-referenced the labels and color codes of the spreadsheet against the relevant passages and in-text citations in the GCF Text. This helped ensure that the samples of text chosen were faithful to the text, and relevant to the research questions. I examined the salience of categories that were relevant to the most important aspects of the GCF Text, the creation of a regime of knowledge, the gender identity of the students, and institutional power. The development of the final groupings of disciplines and types of sources lay the groundwork for the next phase.

During the fourth phase of data collection, I officially selected the most relevant articles for content analysis and created the Bibliography Sample. There were two

rounds of data collection for two methods of analysis. The first round consisted of quantitative content analysis based on keyword search, and the second round consisted of qualitative content analysis based on discourse analysis. Any source that was included in the Bibliography Sample was selected based on its influence upon the creation of a knowledge regime for gendered political agency within the framework. I excluded many of the texts that dealt with topics that were irrelevant to the research questions. For instance, texts dealing with environmental issues, school climate (related to bullying or violence, and not to democratic climate), or the economy were out of scope.

The first round of this phase of data collection prepared the sample for the content analysis method of quantitative category count. Only articles that were cited in key sections of the framework, where the influence upon the cognitive test and questionnaire design was high, were analyzed. I searched for and extracted relevant keywords electronic articles, documents, and books with live text. I used computer programs for the numerical count, using the “Find” function of Apple’s “Preview” program, “PDF Expert,” and other virtual reader applications. All the material was retrieved electroni-

Table 1

Terms used for category counts

Gender-based categories	LGBQ	Trans	Feminist	Female	Gender & Sexuality
	lesbian(s)	transgender		woman(en)	gender
	gay	transexual	feminist	female(s),	sex
	queer	intersex	feminism	girl(s)	sexuality
	bisexual	hermaphrodite		feminine	sexual ¹
	homosexual(s)				
Race-based categories	General	LatinX	Indigenous	Black	White
	race			Black	
	racist(ism)	Hispanic,	indigenous	African	white
	ethnic	Mexican,	Native	African-	caucasian
	ethnicity	Latino/a,	American ²	American	
	brown	Latinx		Afro-	
	Asian ³				
Cultural terms	Cultural				

¹ Terms were mutually exclusive: “homosexual” or “bisexual” were excluded from “sexual”

² Mentions of “native” language speakers were excluded from counts of “Native American”

³ “African,” “Asian,” and “Mexican” were considered race-based based on conventional understandings

⁴ The word “cultural” instead of “culture” was used because it was more related to intercultural education, and less likely to be confused with other uses of culture, such as learning or leadership cultures

cally. I excluded scanned versions of certain documents that did not have live, or “clickable” text, because of time constraints, and the increased possibility of human error. I repeated this process several times, for accuracy. The keywords included in each category are detailed in **Table 1**.

The second round of content analysis was part of the micro-genealogy or discourse analysis. A few articles were selected based on significance. I carefully examined the selected materials, doing a close reading of the context around the keywords to extract the most relevant passages. This was done to triangulate the data. My aim was to identify relative changes in the representation of each term depending on the author’s institutional affiliation, discipline, geographical location, or type of source. Ultimately, I chose texts that were representative of different types of sources (as opposed to disciplines or geographical locations), because they provided the most relevant institutional information. These inquiries informed the micro-genealogy and helped me establish a narrative about the creation of the knowledge regimes.

In the fifth and final round of data collection, I selected texts from the GCF Text, the GCF Questionnaire, and the sample questions for the cognitive test. This is the data I used to conduct a discourse analysis based on the analytics of governmentality. I did a close reading of the entire GCF Text, selecting the most important sections for closer analysis. I selected passages based on their exploration of themes: political participation, identity, gender equality, human rights, values, or motivations for political participation. Excerpts were also selected where gender identity should have been present, but was excluded, such as sections discussing discrimination. I selected the sample questions for the student cognitive test that were most relevant to gender equality or discrimination. I also selected sections of the student questionnaire that dealt with gender equality and political participation, excluding sections for teachers and school leaders.

5.4. Data Analysis

5.4.1. Category Count

A process of content analysis called category counts was employed, in which words are classified into categories and then counted under the assumption that a higher relative count (percentage) reflects a “higher concern with the category” (Weber, 1990).

The use of descriptive and inferential statistics was employed for the purpose of illustration and to indicate general trends and patterns in the data. After grouping together different categories based on geography, institutions, authors, themes, and types of publication, I calculated the relative frequency of gender-based terms, comparing the results to find patterns and trends. This analysis made visible the relative relevance of certain discourses.

Besides a category count, I also conducted a Chi Square Goodness-of-Fit test. This test helped me evaluate whether the sample data “fits” into a theoretical distribution. It compares the sample data to an idea about how the data values are distributed. (JMP, 2020). First, I explored whether the category count data of the GCF Text “fits” with the same distribution as in the Bibliography Sample. In other words, whether the GCF Text is more or less biased towards gender. Then, I evaluated whether the data “fits” within an ideal, “unbiased,” or equal distribution, and in this case, I evaluated all groupings of sample texts against this ideal distribution.

I thus used this method to test two sets of hypotheses:

H_{a1}: The relative frequency distribution of the category counts for the GCF Text will not match those of the Bibliography Sample³

H₀₁: The relative frequency distribution of the category counts for the GCF Text will match those of the Bibliography Sample

H_{a2}: The relative frequency distribution of the category counts will not be equally distributed regardless of the sample, source type, discipline, or author

H₀₂: The relative frequency distribution of the category counts will be equally distributed regardless of the sample, source type, discipline, or author

5.4.2. Analytics of Governmentality

Generally, studies in international education that implement this tool use three or five questions to examine a policy (Fejes, 2009; Kessler, 2009). I used five questions, or steps. According to Dean’s (2009) interpretation of Foucault’s work and of the broader

³ Please note that for the first hypothesis, only the GCF Text was compared to the Bibliography Sample. For the second hypothesis, all the samples (including the GCF Text and Bibliography Sample) were compared against an “ideal” or equal distribution. Also, the sources grouped by discipline or source type are all part of the Bibliography Sample.

trends in governmentality research, these studies are based on “how” questions that help identify the regime of knowledge. This resume is composed of four dimensions: the field of visibility, technologies of governance, rationale of governance, and the formation of identity.

The first step was to identify the regime of knowledge. In this step, I outlined the objectives of the framework, the diagnosis of existing ills, and the schemata for evaluation. The second step consisted of detecting what concepts, gender subjectivities, and gender-based actions were made visible or invisible within the text of the framework. The third step involved the technical aspect of government, or the use of technology, and in this case the technology was the vocabulary related to a certain discourse, and how it was used to regulate gender identity and agency. The fourth step involved identifying the episteme of government, or how knowledge was implemented to rationalize the use of this discourse. This step involved connecting the discourse to the bibliography and the selection of sources to expert authority upon a given knowledge domain, skill, or attitude. As part of this method was the micro-genealogy, in which I focused on the historical formation of the discourses, or regimes of knowledge, that were implemented in the PISA GC Framework. The final step involved the formation of identity, where the resulting gender subjectivity and gender-based agency and action were revealed through an examination of the questions in the student questionnaire, and the cognitive test.

6. RESULTS

The analysis of discourse of the PISA GCF involved problematizing the assessment of global competencies in relation to gender-based identity and action, making educational assessment professionals accountable to the students and the nations that they assess. I questioned who ruled the assessment framework, how and under what condition they were legitimized, and what questions they asked themselves to govern students. I did this not just by describing their practices, but also by describing how these practices were rationalized and formed according to specific forms of knowledge.

Research Questions:

1. Through what kind of discourses or regimes of knowledge did PISA GCF regulate gender subjectivities in their assessment of citizenship competencies?
2. In what ways, if any, did the 2018 PISA GCF measure the gender-based political agency and participation of students?

I analyzed three samples: the GCF Text (including the Student Questionnaire and Sample Cognitive Test questions), the Bibliography Sample, and the Author Sample. To answer these questions, I employed two approaches. First, I conducted a quantitative content analysis, using the method of category count, quantifying the frequency of terms to illustrate the visibility of gender concepts in relation to other relevant concepts. I found a series of patterns that revealed important information about the regimes of knowledge. Secondly, I conducted a Foucauldian discourse analysis of the different dimensions of governance of the 2018 PISA GCF Text. As part of the micro-genealogy of gender theory in relation to PISA GCF, I closely examined a selection of texts from the Bibliography Sample. The findings are described in the next two sections.

6.1. Category Count

I first conducted a category count for both the CGF Text and the Bibliography Sample. The results are illustrated in Table 2 and 3, which contain the relative frequency of each category of terms in relation to the total number of relevant terms. Table 2 shows the terms as organized into three categories: culture, race, and gender. Table 3 shows the gender category further split into five groups of terms related to: feminist

identity, female gender identity, LGBTQ identities, trans and intersex identities, and general references to gender and sexuality. Both tables show the distributions of relative frequency in the 2018 GCF Text (OECD, 2019a) and in the Bibliography Sample.

Table 2 illustrates some key differences. Looking at the table, it seems as though the distribution of the frequencies is not the same for the works cited and the GCF text. The relative frequency of terms related to culture seems higher in the GCF text (87.50%) than in the works cited in the bibliography (62.85%). The relative frequency of terms related to gender seems lower in the text (4.17%) than in the works cited (15.60%). Similarly, the relative frequency of terms related to race seems significantly lower in the GCF text (8.33%) than in the works cited (21.54%).

Table 3 illustrates key differences in the types of gender terms used in the GCF Text in comparison to the sample of works cited in the bibliography, which had a seemingly lower relative frequency of terms related to LGBTQ and feminist identities. There was only one mention of the term “transgender” in all the works in the Bibliography Sample, a 6.15% relative frequency of terms related to homosexuality, and less than four percent relative frequency of terms related to feminism. Meanwhile, the category count of the framework text revealed zero mentions of terms related to homosexuality, transgender identity, or feminism. All the terms related to gender in the framework text referred either to the abstract concept of gender and sexuality (10 mentions, or 76.92%), or to heteronormative female identity (3 mentions, or 23.08%). I found that both the framework text and the sample of works cited were highly heteronormative and cisnormative, and that the framework text had zero mention of LGBTQ and feminist terms.

In order to evaluate the significance of these percentages, the finding was corroborated with a Chi-Square goodness-of-fit test in two different ways. The first null hypothesis was that the GCF text would have the same relative frequency distribution as that of the sample of works cited.

H₀₁: the relative frequency distribution of the category counts for The GCF Text will match those of the Bibliography Sample

Appendix Table A1 shows the results of these tests, and their significance. There were two parts to this hypothesis. For the comparison between the GCF Text and the Bibliography Sample in the categories of race, gender, and culture: the relative frequency distribution of the category count for the GCF Text did not match those of the Bibliography Sample, $\chi^2(2, n = 3) = 81.59, p < .00001$.

Table 2*General Category Count of the 2018 GCF Text and Bibliography Samples*

	Gender		Race		Culture		Total Count
	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	
Bibliography Sample	4,018	15.60%	5,548	21.54%	16,185	62.85%	25,751
Framework Text	13	4.17%	26	8.33%	273	87.50%	312

Table 3*Gender Category Count of the 2018 GCF Text and Bibliography Samples*

	Gender & Sexuality		Trans		LGBQ		Female		Feminist		Total Count
	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	
Bibliography Sample	1,691	42.09%	1.00	0.02%	247	6.15%	1,932	48.08%	147	3.66%	4,018
Framework Text	10	76.92%	0	0.00%	0	0.00%	3	23.08%	0	0.00%	13

The null hypothesis was rejected, and this finding could be considered highly significant. For the comparison between the GCF Text and the Bibliography Sample in the gender categories of female, trans, LGBTQ, feminist, and references to “gender and/or sexuality”: the relative frequency distribution of the gender category count for the GCF Text matched those of the Bibliography Sample, $\chi^2(4, n = 5) = 6.72, p > .05$. The null hypothesis was not rejected, because the alternative hypothesis was not significant enough. Based on these results, it could not be said that the GCF text was significantly more heteronormative or cisnormative than the Bibliography Sample, but it could be said that the GCF text is significantly less representative of gender.

Besides comparing the GCF text to the sample of works cited in its bibliography, I also compared both samples to an “unbiased,” or “neutral” distribution, where all categories would have the same frequencies. The second null hypothesis was that terms would be equally distributed among all categories:

H₀₂: the relative frequencies of category counts will be equally distributed, regardless of the sample, source type, discipline, or author

Appendix Table A2 shows the results of these tests. There were several parts to this hypothesis. In all cases, the relative frequency of category counts of each sample were compared to an “unbiased,” or equal distribution among categories. In all the cases presented in the table, the relative frequency distribution of all categories were shown to not be equally distributed, and the null hypothesis was rejected. For instance, the relative frequency distribution of the general category count for the GCF Text were not equally distributed, $\chi^2(2, n = 3) = 412.75, p < .00001$. The null hypothesis was rejected, and the results were significant. The relative frequency distribution of the gender category count for the GCF Text were also not equally distributed, $\chi^2(4, n = 5) = 28.92, p < .00001$. The null hypothesis was rejected, and the results were significant.

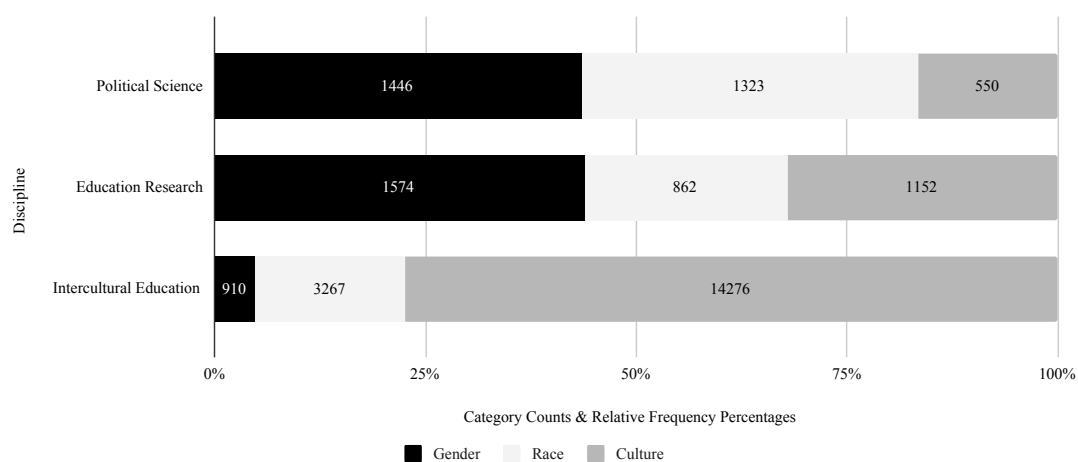
Since all the other Chi-Squares test produced similar results, it suggests that none of the samples could be said to be equally distributed among categories of race, gender, and culture- or even among the sub-categories of gender. There was a bias evident in all the category counts, sometimes towards culture, and sometimes towards gender. In the rest of this section, I will not re-state these findings, but simply point the reader to Table A2, in the Appendix, to refer to the results for the category counts and their significance levels. The discussion section addresses the limitation of this data.

In the next few paragraphs, I illustrate the findings from comparing the works in the Bibliography Sample to each other in order to detect differences in patterns, and possible sources of these differences. I organized the Bibliography sample, according to a variety of different groupings (sometimes excluding work irrelevant to a given grouping), looking for patterns in the data that would reveal bias. I found two groupings that revealed significant patterns. First, I organized the data into groups of disciplines or fields of study: political science, education research, and intercultural studies. Second, I organized the data into groups of source types: sources connected to certain institutions, sources from academic journals, sources from academic books, and sources from dissertations. Finally, besides these groupings, I examined the geographical representation of the authors of both the framework text and cited works to understand the proportion of Western and non-Western authors. The results of these analyses are shown below.

I identified three academic fields that were made visible in the sample of works cited in the bibliography: intercultural studies, political science, and education research. I treated intercultural education studies as a distinctive field because of its interdisciplinary nature, and because of its salience, since it had the highest word count: more than 13,000 higher than political science, and almost 10,000 higher than the two other categories combined. Out of a total of 114 texts referenced in the bibliography, 50% of the sources were from education research (including psychology), 7% were from sociology and political science, and 15% were from intercultural education. I excluded the remaining 28% of publications from this section of the analysis because they were multidisci-

Figure 1

General Category Count of Bibliography Sample: Organized by Discipline

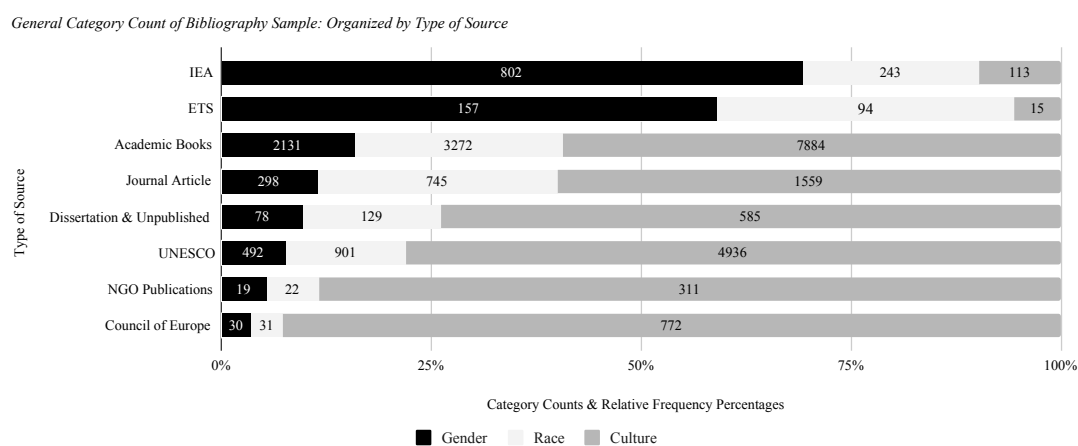


plinary government or NGO publications. I observed that the fields of queer theory, decolonial theory, gender studies, and critical race theory, were omitted from the works cited in the bibliography. I concluded that the field of intercultural studies was significant based on its disproportionately high word count in the sample and based on the numbers of works cited overall; and therefore, I found that the omission of work related to the fields of gender was also significant in this regard.

After conducting the category count analysis arranged by field of study, I found that intercultural studies were significantly different to other disciplines. Figure 1 shows the results of the category count analysis organized by discipline, with the relative frequency percentages shown in colors, and the counts shown numerically (for the relative frequency percentage numbers, refer to Table B1 and B2). These findings seem to indicate that the relative frequency of terms changed according to the field of study to which the work belonged. In comparison to all other fields, the field of intercultural education had the lowest relative frequency of terms related to gender (4.93%), and the highest relative frequency of terms related to culture (77.36%). The field of political science was the only one to have a higher relative frequency of terms related to gender (43.57%) than culture (16.57%). In comparison, education research had a more equally distributed representation of the three categories. The results of the Chi-Square tests that suggest significantly different relative percentages are shown in Table A2. Based on this sample, I found that the fields of study made a difference in the representation of categories in the works cited, and that intercultural education studies had the lowest relative frequency of gender terms, and the highest frequency of culture terms.

The category count analysis arranged by source, illustrated in Figure 2, and Tables B3 and B4, revealed a general bias towards culture over gender.

Figure 2



Over half of the relevant terms in most of the cited works in the sample were related to culture, and less than a fifth were related to gender. The exceptions were IEA and the Educational Testing Service (ETS), which showed a reversed effect, with ETS works showing particular bias towards mentions of gender (59.02%), in comparison to culture (5.64%). Bias towards culture was most evident in sources from the Council of Europe (92.68%), and from NGO publications (88.35%), with a bias against gender made evident in the relative frequency of gender terms: 3.4% and 3.6%, respectively. The bias towards culture was evident for peer-reviewed sources (59.92%), to the detriment of culture (11.45%); yet scholarly texts that were not peer-reviewed, such as dissertations or unpublished work, showed an even greater bias towards culture (73.86%), in relation to gender (9.85%). Although the pattern could be attributed to a low word count, the trend remained even for academic books, which had a bias towards culture (59.34%) in relation to gender (16.04%). In conclusion, most of the works that were associated with institutions (IEA, ETS, NGOs, UNESCO, and the Council of Europe) were more biased in either direction, while academic books and academic journals were slightly more balanced. The results of the Chi-Square tests that suggest significantly different relative percentages are shown in Table A2. There was a general trend towards bias towards either culture or gender, with most types of sources showing a bias towards culture.

After analyzing the proportion of works cited based on the geographical regions, illustrated in Table 4. I found evidence of bias towards the selection of Western sources. The 2018 PISA assessments included 79 countries, including 39 countries that were not members of the OECD (OECD, 2019a). Despite how widespread these assessments

Table 4

Bibliography Sample: Regional Distribution of Sources

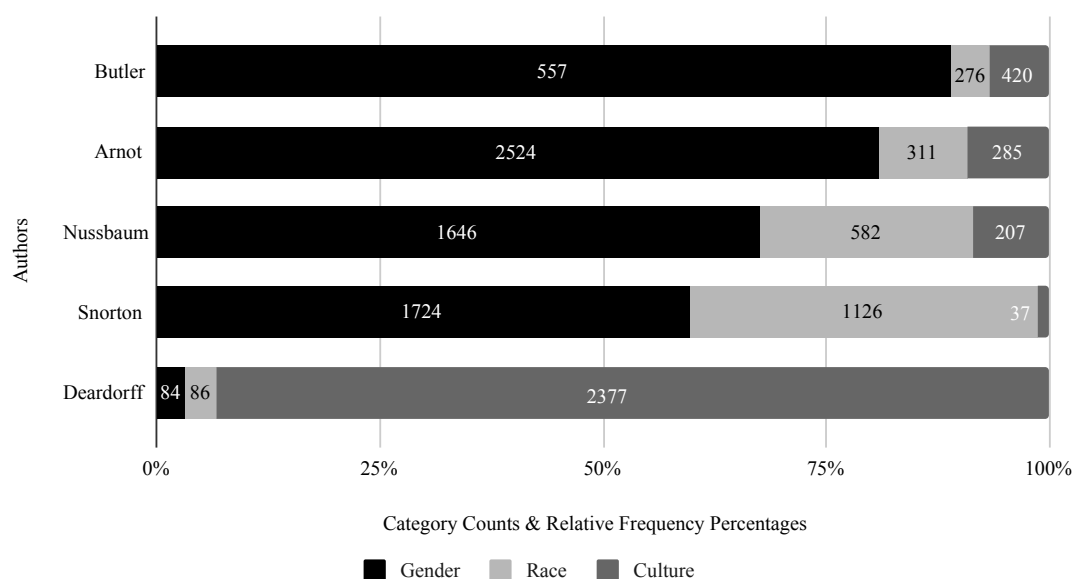
	National Affiliation of Bibliography Sample Authors		Nationality of GCF Text Authors
	Count	Percentage	Count
UK & Canada	73	56.59%	1
UK	21	16.28%	1
Europe	15	11.63%	2
Australia	4	3.10%	0
Asia	2	1.55%	2
Africa	3	2.27%	1
International	14	10.85%	NA

were, and despite claims of international, global, and intercultural expertise, most of the works cited (87.6%) were representative of the West: the US, Canada, Europe, Australia, and the UK. In fact, 10.61% of the sources, which were labeled international, were from organizations like UNESCO and IEA whose leadership is based in the US, Australia, and Europe. Although there were two authors from Asia, and one author from Africa, this representation was not reflected in the works cited. There were very small numbers of sources from countries in Asia (2 sources), Africa (3 sources), Latin America (0 sources), and even Eastern Europe (3 sources). Even the sources that were international belonged to Western organizations whose authors were primarily from the UK, Europe, the US, and Australia, such as UNESCO, IEA, and ETS. Based on this geographical representation, it could not be said that the rationale was to have geographical inclusivity. Ultimately, claims of international expertise were contradicted by the disproportionate representation of Western sources.

Next, I conducted a category count analysis with a micro-genealogy lens, and compared authors cited in the PISA GCF to the authors in the field of gender studies. I found that the field of study to which the authors belonged to seemed to influence their relative mention of gender compared to other concepts. Figure 3, and Table B5, show the extent to which each author discussed gender in comparison to race and culture. The results of the Chi-Square tests that suggest significantly different relative percentages are shown in Table A2. Deardorff rarely mentioned any term not related to culture. In

Figure 3

General Category Count for the Influential Author Sample



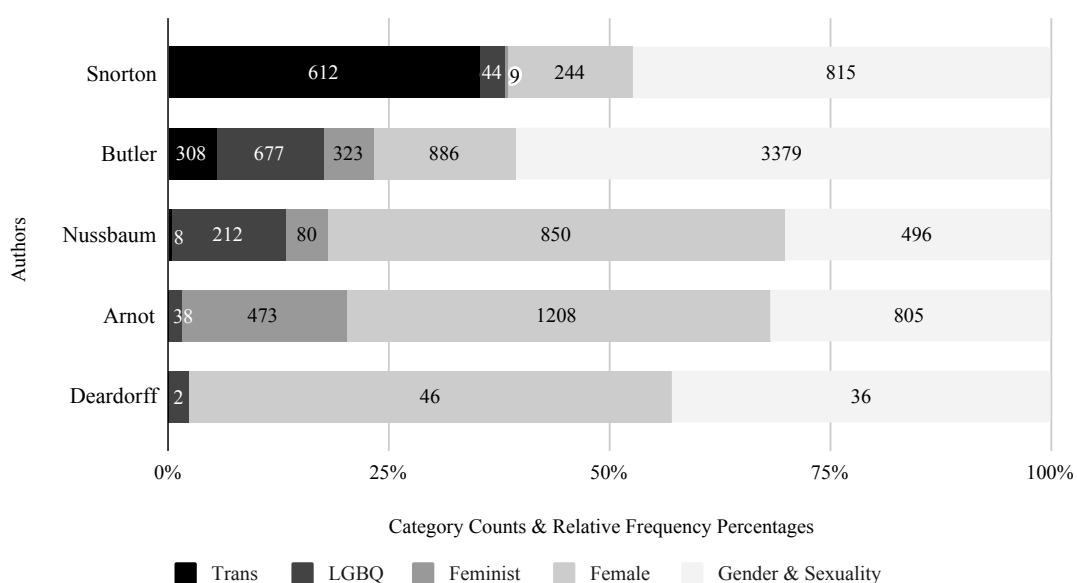
comparison, all other authors dedicated more than half of the relevant terms to gender, with very little mention of culture. Overall, it seems clear that Deardorff, influential to intercultural studies, was biased to discuss culture over gender.

The field of study to which the authors belonged also seemed to influence how inclusive they were of LGBTQ identities. The most significant difference in gender subjectivity was the relative heteronormativity and cishnormativity of the texts. Most texts had a relatively higher reference to heteronormative terms. The results of the Chi-Square tests that suggest significantly different relative percentages are shown in Table A2. As shown in Figure 4, and Table B6, Arnot, a feminist author, and Deardorff notably did not mention trans people in any of their publications. Only authors from queer and trans theory, Butler and Snorton respectively, mentioned trans people in a significant way. Butler had overall more general terms referring to gender and sexuality (88.9%), reflective of the abstract nature of her work, and yet, 5.53% of the gender terms referred specifically to trans people. Snorton, a trans theory author, had a consistent representation of trans people (35.5%) throughout all her major works. Overall, except for trans and queer theory authors, most authors produced cishnormative texts.

I also found that, based on this sample, more updated works tended to be more gender inclusive. Texts from different time periods were selected to reflect the differences in discourse throughout time. Deardorff's earliest publication had no mention of

Figure 4

Gender Category Count for the Influential Author Sample



any other category besides culture. Nussbaum and Butler increased their mention of terms related to trans identity in their later publications. Nussbaum did not mention trans people in her earliest publications, Butler, on the other hand, did mention trans people from the beginning. Therefore, one could assume that the time period from which an author's work is published influences their mention the gender subjectivity terms they used.

Overall, I found that in comparison to the Bibliography Sample, the framework text contained more terms related to culture than terms related to gender or race. While both the GCF and the cited works were highly heteronormative and cisnormative, the framework text had zero mention of LGBTQ and feminist terms. The field of intercultural studies had a disproportionately high word count in the sample, while works in the fields of gender, race, and inequality were omitted. The fields of study made a significant difference in the representation of categories in the works cited, and intercultural studies had the lowest relative frequency of gender terms, and the highest frequency of culture terms. I found that Western sources were disproportionately represented in the Bibliography Sample. After examining certain pivotal authors in feminist, queer, and transgender studies in comparison to two authors influential in the framework, I found that the field of study to which the authors belonged seemed to influence the relative heteronormativity and cisnormativity of the texts. Most texts had a relatively higher reference to heteronormative terms, and except for trans and queer theory authors, most texts were cisnormative. The work of feminist and intercultural studies authors seemed to be the most cisnormative, and the currentness of their work seemed to influence gender subjectivity as well.

6.2. Foucauldian Discourse Analysis

I organized the results of my analysis according to the different dimensions of discourse I examined, and in each dimension, I addressed the gender subjectivity and the political agency of students in relation to gender-based political action. I started by looking at the government as a regime of practice and knowledge. This involved exploring how the framework defined problems, the necessary competencies for the citizen to provide solutions to these, and how these were evaluated. Next, I explored the three dimensions of a regime of knowledge. The first dimension I explored was the fields of visibility, or how gender-based subjectivities and actions were made visible or invisible.

In the second dimension, I explored the technical aspect of government, or how vocabulary was used as a technology that regulated gender-based identity and action. In the third dimension, I explored the rationale of governance, or the overall strategy and reasoning that ruled these practices, as well as a micro-genealogy, in which I explored the historical dimension of discourse. Finally, I looked at the formation of identity, and this involved examining the sample questions asked and how they used the practices, techniques, and rationales to create a gendered subjectivity of students.

6.2.1. Regime of Practices and Knowledge

In this section, I analyzed the rationalization of practices to identify the knowledge domains that were formed in relation to a specific knowledge discourse in the 2018 GCF Text (OECD, 2019a). First, I explored the diagnosis of existing ills. Second, I identified the set of objectives or competencies that the ideal citizen was supposed to obtain. Finally, I described the schemata of evaluation used to control these concepts. These steps helped me identify the practices through which a regime of knowledge was established.

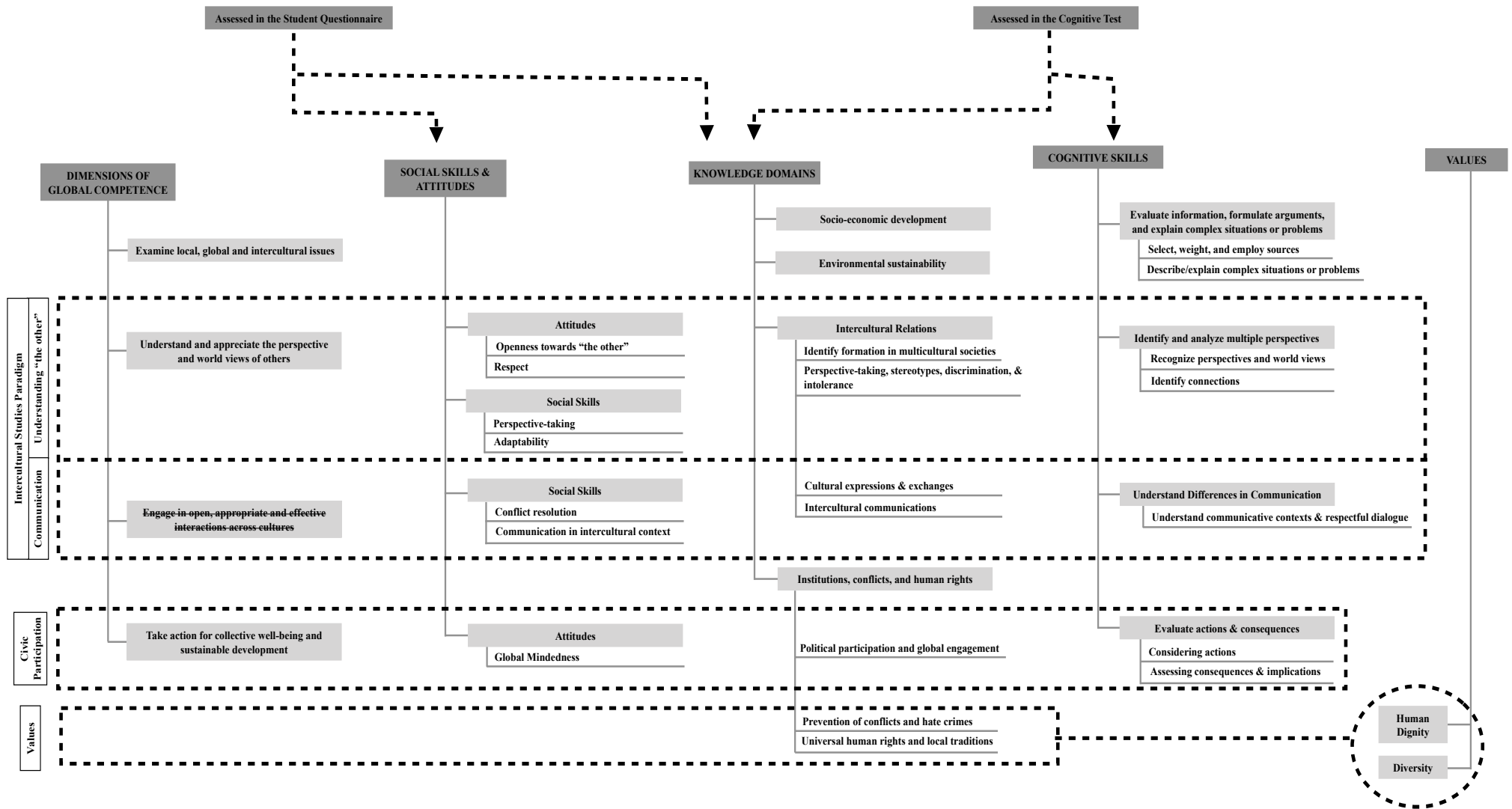
After analyzing the section of the text that described world issues targeted by the assessment, I found evidence that a stated concern for inequality was contradicted by a universalist concern for diversity. The authors mentioned that education could “[end] racism and discrimination” and help meet the Sustainable Development Goal, target 4.7, of promoting the values of gender equality, human rights, and non-violence. However, they contradicted their stated purpose by focusing on intercultural skills as a solution. They encouraged students to challenge “cultural biases and stereotypes,” and adopt “intercultural sensitivity and respect” as well as an “appreciation for diverse peoples, languages, and cultures.” (OECD, 2019a, pp. 166-67) Underlying this contradiction was the assumed connection between inequality and culture that was academically unsupported, and insufficiently questioned in the text. Moreover, they placed students outside the manifestations of inequality, centering the perspective of privileged subjects who tolerate the diverse cultural “others,” while displacing the experiences of marginalized students. The globally competent student was thus depicted neither as oppressor nor oppressed, but universal, existing outside the struggles of systemic inequality, and able to resolve it with intercultural skills. Based on this contradiction, I concluded that the societal ill described was not systemic inequality, but intolerance for diversity.

The authors focused on the negative impact of cultural diversity and the need to restrain it. They described world problems as encounters between people of different cultures who disrupt harmony, or “ethno-cultural conflicts,” “indiscriminate violence in the name of religious or ethnic affiliation,” and the “high influx of immigrants” (pp. 166-67). These descriptions created a blanket portrayal of ethnic affiliation as negative. They also erased power disparities that direct the flow of migrants, inform their encounters, and implicate certain ethnic groups in the marginalization of others. The authors described the problem of cultural diversity in terms of its excess, framing it as something to be restrained, and dismissing the power disparities inherent in the conflicts it creates, which served to minimize the presence of inequality.

Ultimately, my analysis indicated that the authors framed world problems as failures to assimilate. The authors specified that student citizens were supposed to address the lack of harmony and insufficient cultural assimilation. Successful students were described not in the terms of social justice, but in economic terms, as employees who “easily adapt” to diverse environments and adopt “appropriate” behavior. The authors also emphasized detachment from ethnic and gender-based affiliation encouraging students to instead adopt “complex forms of belonging and citizenship,” based on “the nation, the region, the city, the neighborhood, the school.” (pp. 166-67) I interpreted these statements to direct those whose ethnic or gender-based affiliations are systematically attacked to discard their identities, to which they might cling to for survival, and adopt an abstract collective cultural identity in order to conform to a uniform community.

Rather than promoting the inclusivity of “diverse” groups, the priority was given to cultural assimilation, without accounting for its potentially oppressive nature. This formulation of world problems as insufficient assimilation contradicted the stated purpose of reducing inequality and discrimination, and the concern with diversity. I found that the paradigm of intercultural studies was salient as a regime of knowledge established through objectives that promoted cultural assimilation. Through many dimensions, characteristics, and skills, the authors constructed a set of objectives for students to become globally competent. Global competence had four dimensions, a set of four knowledge domains, and two sets of skills, described as either social or cognitive. I illustrated these categorizations in Figure 5, as I understood them to be connected based on the 2018 GCF Text. According to this schema of concepts, the three most relevant groups I identified, marked with dotted lines, were skills related to intercultural studies, civic participation, and values.

Learning Objectives of the 2018 PISA Global Competencies Framework



I found that most skills and knowledge domains emerged from the paradigm of intercultural studies. As shown in the diagram, most of these skills could be described as either objectives for the understanding of, or communication with, a cultural “other.” These skills encouraged assimilation through “adaptability,” and the creation of harmony through “conflict management” and “respectful dialogue.” Cultural assimilation was thus the ultimate objective, guided by the regime of intercultural knowledge.

My analysis indicated that global competence was organized in a complicated schemata for evaluation to increase the authority of the intercultural regime of knowledge. As illustrated in Figure 1, the categorization of concepts determined the evaluation methods: the questionnaire involved social skills, while the cognitive test involved cognitive skills. The authors labeled intercultural skills as both cognitive and social without sufficiently differentiating them. For example, perspective-taking was formulated in four different ways (i.e., under knowledge domain, dimension of global competence, social skill, and cognitive skill), and I found no material differences between these formulations other than phrasing. If so, theoretically, any student characteristic could be labeled as a cognitive skill. The considerations of reliability were also inconsistent. The authors placed cognitive skills on a comparative scale because they considered them more objective than self-reported social skills (pp. 183-197); and yet, the authors excluded communication due to the difficulty of testing it, and acknowledged that perspective-taking was not supported academically (p. 183). Thus, intercultural skills were acknowledged to be unreliable while paradoxically also considered to be more objective when labeled as cognitive. Given these inconsistent considerations of reliability and the insufficient distinctions, I concluded that labeling was used to increase the perceived reliability of intercultural skills.

The bias towards the intercultural regime of knowledge was again evident in the bias towards intercultural skills in the questionnaire. The decisions made to include intercultural skills and exclude values from the assessment was inconsistent with the reasoning given. The values of human dignity and diversity were determined to be “unmeasurable” and out of scope of the questionnaire (pp. 169-179). This was even though four pages were dedicated to the discussion of values, and they seemed to be included in the assessments as knowledge domains under the labels of “prevention of hate crimes” and “universal human rights and local traditions.” Although the authors acknowledged a similar difficulty with measuring “socio-emotional skills and attitudes,” which included

mostly intercultural skills, they used previously validated scales (p.195) to include them in the questionnaire. Thus, the reasoning was not consistent with these two different student characteristics of values and social skills, demonstrating a biased and inconsistent treatment of skills that disproportionately favored the intercultural regime of knowledge.

Concepts related to political participation were not described or measured in relation to gender. Political participation was mentioned as a knowledge sub-domain, as the cognitive skill of evaluating actions, and as the social attitude of global mindedness. In the cognitive test, both the knowledge sub-domain of political participation, and the cognitive skills of evaluating actions, were measured through scenarios about real experiences of student community-level political participation, or their evaluation of actions related to political issues. Global mindedness was measured through items in the questionnaire with concepts like world citizenship, responsibility for “others in the world,” “inter-connectedness”, and “global self-efficacy” (OECD, 2019a, p. 197). In all cases, the measurements were overly general and unspecific to gender-based agency or action, such as gender-based voting, campaigning, or activism. Moreover, despite the repetition of similar terms, the descriptions of participation concepts did not target actual political participation, but student reflection about their actions. Overall, the description of political participation was overly general, and not related to gender-based agency.

Through an analysis of the rationalization of techniques, I found that the framework was formulated in relation to an intercultural regime of knowledge. The authors of the framework diagnosed social ills based on a universal discourse that cast conflicts as results of excessive diversity. They set knowledge objectives in terms of intercultural skills that prioritized cultural assimilation, and they created a complicated schemata of evaluation that prioritized intercultural skills over other concepts, in a manner that was inconsistent with their reasoning. There was evident bias towards an intercultural regime of knowledge that encouraged universality and cultural assimilation.

6.2.2. Field of Visibility

In this section, I explored the field of visibility, and its relation to the diagram of power. Using the results of my category count, I analyzed how the bibliography was used to make certain terms visible or invisible. Based on the findings of my category count, I conducted a closer discourse analysis of the GCF Text. As indicated in the category count, the bibliography was not inclusive, as it made gender studies invisible.

Since the academic perspectives that were made visible were only the ones that were included in the bibliography, my analysis indicated that the bibliography was used as a tool, or diagram of power, that revealed the field of visibility. In this section I explored whether the omission of authors and fields of study from the bibliography resulted in the omission of gender-based terms from the 2018 GCF Text. I conducted discourse analysis to examine whether certain gender subjectivities and certain gender-based actions were made invisible.

In terms of the different representations of gender, I found that the GCF Text made visible only a Western, white, heterosexual, and cisgendered point of view. Since this framework classified students in terms of “male” and “female,” and since there was no mention of other gender identities such as “queer,” “trans,” “feminist,” or “gender non-conforming,” I concluded that these identities were not included. The only allusion to the existence of LGBTQ people was in a long list of behaviors that are viewed with prejudice, among which “intolerance towards sexual inclinations” (OECD, 2019a, p. 213) was mentioned. I did not consider the reduction of queer, gay, and lesbian identities to nonspecific sexual proclivities to be sufficiently representative, and even found it to be otherizing and offensive. Thus, I found that the text both minimized and otherized LGBTQ identities, and promoted a heteronormative and cisnormative gender subjectivity.

I also found that the portrayal of girlhood was insufficient and tokenizing. Accurately analyzing the gender subjectivity of gender-based groups, such as women and girls, proved to be challenging, given their absence from most of the text. The term “woman” and “girl” only appeared once each: the former was dismissed as an overly simplistic marker of identity (p. 171), and the latter was dismissed in favor of an identity based on profession and citizenship (p. 187). Both mentions of womanhood were used to encourage a universal identity based on complexity. Another finding from discourse analysis that was not evident from the category count was the repeated use of feminine pronouns. In the discussion of cognitive skills, the subject was repeatedly described as “the student,” yet female pronouns, such as, “She understands,” “she listens” (OECD, 2019a, p. 188) were also used. This led to awkward phrasing such as “The student views herself.” Although the female gender pronoun was used, it was not about a discussion of girlhood, but about its dismissal and the importance of creating distance from gender-based subjectivity. I therefore found the use of female gender pronouns to

be tokenizing and insufficiently representative of female subjectivity, and that female subjectivity was only included to promote a universal subjectivity.

I found that gender-based political action was discouraged, and action was formulated as value-neutral interpersonal interactions. The only mention of agency was in the social attitude of global mindedness. This attitude involved students forming “A critical awareness of” different views, being open to “reflecting on and changing their vision as they learn of different perspectives,” and creating space for different ways of living without eliminating differences (p.176). This description of global mindedness limited the understanding of agency to the act of reflecting on interpersonal communication, and to an attitude of neutrality in perspective-taking free of ethical values. In fact, this description mirrored the same trend throughout the text, where many skills and attitudes were simply different formulations of the concepts of perspective-taking and communication. I concluded that political action was limited to the interpersonal interactions, and not formulated around gender-based agency, but as encouraging neutrality.

6.2.3. Technical Aspect of Government

In this section, I examined the technical aspect of government, or the use of technology, which in this case was the vocabulary related to a certain discourse, and how it was used to regulate gender identity and agency. First, I examined how the vocabulary accomplished authority over student gender subjectivity, keeping in mind that values did not emerge naturally but were established by the authors. Second, I examined how the vocabulary accomplished authority over student political action.

The term “complex” was often used to replace gender-based identity with a universal identity and to avoid mentioning the many manifestations and material consequences of inequality. It was used 41 times in reference to people, knowledge, and situations, many more times than gender-based terms, and it sometimes worked as a placeholder for identity. It was also used to sanitize world problems. Systemic issues that could be simplistically described as the corporate destruction of natural resources, or worker exploitation, were instead called “complex and controversial global issues” (OECD, 2019a, pp. 170), “complex dynamics of globalization” (pp. 167), or “complex systems and policies” about the environment (pp. 171). This wording minimized the material consequences of inequality, and significantly impacted the framework. For example, the reason given to eliminate values from the assessments was simply that their

evaluation was “complex and therefore beyond the scope” (pp. 178) of the framework, without citing academic research to support this stance. The term was thus also used to dismiss the need for academic research. The use of the term “complex” therefore significantly shaped the assessment because it was used as a placeholder for inequality, marginalized subjects, and academic research.

In the description of the knowledge domain of intercultural relations, the authors used the term “complex” to reject gender-based and race-based identity perspectives in favor of universalized identities based on culture:

“As they engage in learning about other cultures and individual differences, students start to recognise multiple, complex identities and avoid categorising people through single markers of identity (e.g. black, white, man, woman, poor, rich). Students can acquire knowledge in this domain by reflecting on their own cultural identity and that of their peers, by analysing common stereotypes towards people in their community, or by studying illustrative cases of conflict or successful integration between cultural groups.” (Pg. 171).

In this, the only mention of the term “woman,” the authors rejected the label as simplistic, and portrayed “complex” universal communities based on culture as better, erasing communities centered around womanhood and race. The authors also reduced conflict to stereotypes, or misunderstandings between cultural groups imagined on an equal playing field, and given the goal of assimilation. Later in this page, the authors described conflict on a group-level, “between countries, ethnic, or social groups” (pp. 171), imagining a uniform cultural surface-level identity rather than recognizing the gender-based groups within them. In describing conflict as a failure to integrate, the authors placed the onus on the misunderstood to explain themselves rather than on oppressors to eliminate stereotypes, which obfuscated the origin of power, and downplayed stereotypes as accidental misunderstandings. The term complex thus served to dismiss gender-based identities in favor of complex, or abstract, culture-based identities, and in doing so, misrepresented conflicts as failed assimilation between equal groups, and erased the violent nature of inequality.

The only mention of the term girl was also rejected in favor of individualized complexity. According to the categorization of performances in the cognitive test, the advanced female student has a “complex” identity based on her being “a girl, a daughter, a farmer, and a citizen,” while less advanced students understand people through simplistic one-dimensional categories (p. 187). From a feminist standpoint, it is controversial to claim that a young woman’s identity should be attached to her employment,

the state, or to her father, or that being part of a community connected by gender is a problem. The authors indicated that the scenarios in the cognitive tests should “address the multiple, complex identities held by individuals so that they do not perpetuation [sic] the “single story” identity” (p. 212). In both instances, the idea that a more universal complex identity is better for female students than, for instance, a feminist community, was not given an academic foundation. Moreover, the grammar mistakes indicated the poor consideration that such statements were given. The term “complex” was thus used to dismiss gender-based identity.

The term “other” also served as a placeholder for gender, and used more frequently in the text, reinforcing the portrayal of a universal student identity. The word “other” in reference to people was mentioned 59 times in the sections discussing skills, values, and attitudes. It created a uniform identity based on difference from the universal student, sometimes described as a girl, and it erased race and gender identity and privilege, minimizing the importance of this distinction, and therefore minimizing the dynamics of inequality. This erasure was evident when the authors wrote that the students should be open towards “cultural otherness,” and “willing to suspend their own cultural values,” but prevent themselves from making the other exotic by placing them “on an equal footing” (p. 175). The underlying insinuations made were that inequality can be wished away, or that students who experience “cultural otherness” as racism or sexism should be understanding and suspend values that might protect them against these forces. The term “other” was thus used to obfuscate the dynamics of inequality and reduce the relevance of gender identities.

Since gender was rarely mentioned, to illustrate more clearly the workings of this method of erasure I took a closer look at the use of cultural terms to erase race-based identity. In the section on values of human dignity and diversity they cited an example of how not to teach these values by describing a lesson about the American Civil War, a war over slavery and racism. The teacher was described as disciplining “minority students more severely” (p. 179). The authors avoided identifying the race of the teacher or the students, and avoided labeling this glaring incident of racism. Simultaneously, they acknowledged that teachers should not avoid racism in their classrooms.

“Teachers can be instrumental in replacing stereotypes of minority and disadvantaged students with more positive ones. However, teachers often find it difficult to engage in open discussions about diversity and discrimination. Part of the problem is a lack of experience with peo-

ple who are different, and the assumption that conversations about discrimination and ethics will always be contentious. Consequently, teachers may concentrate only on “safe” topics about cultural diversity, such as cross-group similarities, ethnic customs, cuisines, costumes and celebrations, while neglecting more troubling issues such as inequities, injustices and oppression” (p. 179).

This passage marked a key contradiction in the framework, where the authors used cultural terms to avoid mentioning race, yet they discouraged teachers from doing the same. The context illustrated here was particularly racially charged, yet the authors avoided mentioning the racial identities of the teachers or students. The underlying meaning was that white teachers were uncomfortable teaching Black and Brown students about racial discrimination out of fear students would get angry (or “contentious”), or because they would say racist things to each other. The reluctance to identify the race of the students or teachers was evident given the repeated ways that the authors chose to describe student identities in somewhat approximate terms; such as, “cultural diversity,” “multi-ethnic,” “multicultural,” “people from different ethnic groups,” “minority students,” “stereotypes of minority and disadvantaged students,” “people who are different,” and “minority groups.” The tactics that the authors described as being too “safe,” like “cross-group similarities, ethnic customs, cuisines, customs and celebrations,” (p. 179) all alluded to a focus on culture-based collective identities, not race-based identities. Although these examples don’t relate to gender, they demonstrate more clearly the purpose of using cultural vocabulary as a method of erasure. This passage illustrated how culture-based terms were used to replace and erase discussions of race and racism.

The use of vocabulary not only restrained identity, but also restrained the possible actions of students, as cultural terms worked specifically to promote neutrality. Reducing human political actions to the dimension of cultural interactions, and framing conflict as a misunderstanding neutralized active civic participation in politics into internal or micro-level processes of reflection and interpersonal communication. The political actions of students were limited to “appropriate” (p. 166) “effective” (p. 173) and “respectful” (p. 169) communication and behavior, and an encouragement for students to “reflect on [an] issue from multiple cultural perspectives, keeping in mind the interest of all parties involved” (p. 169). Passages like this demonstrate how the text focused on genderless cultural perspectives and prioritized equanimity and neutrality towards “all”

interests, without giving sufficient thought to antisocial interests such as sexism or racism. The centering of imagined conflicts around culture detached these forces from systemic inequality and promoted the value of neutrality.

The authors also used vocabulary related to culture to erase gender-based discrimination and regulate actions related to it. The terms sexism, homophobia, gender-based violence, or discrimination against women, were not mentioned in the framework, and authors always referred to conflict in terms of culture. However, their concern with these forces is noticeable in their description of the context within the multicultural classroom, which they wrote is composed of students with other “gender, religion, socio-economic background[s]” (p. 190). To address these differences, they identified the two intercultural cognitive skills of perspective-taking and communication. Another section that described the application of these skills in the cognitive assessment clarified that these skills were meant to address:

“Common expressions of prejudice and oversimplification [which] include: gender or socioeconomic-based stereotyping about what students can achieve in different subjects; gender or racial biases while selecting applicants for a job; perceptions about certain groups’ predispositions to violence and crime; stereotypes about indigenous cultures; intolerance towards sexual inclinations; religious stereotypes” (213).

This passage enables intercultural skills to successfully address inequity by depicting conflict in individualistic rather than systemic terms, such as: “stereotypes,” “prejudice,” “oversimplification,” “biases,” “perceptions,” and “intolerance.” Although cultural groups are alluded to when describing conflict, these examples described instances of sexism, homophobia, and racism. In describing them as cultural matters between equal groups, the authors were able to reduce systemic inequality to their cultural realm of expertise, eliminating the need to cite research to verify their claims. They claimed that gender-based employment discrimination could be eliminated through perspective-taking, and that the genocide and racism that indigenous groups have experienced for centuries, reduced to stereotypes in their description, could be fixed through better communication. None of these claims were supported with academic research. The authors were explicit about their belief that “intercultural” skills were important because without them, “misunderstanding degenerates into negative stereotypes, discrimination and violent conflict” (p. 170). Thus, the fight of marginalized groups for social

recognition and their fight against racist or sexist state violence was described, essentially, as a big misunderstanding. Using terms related to culture, the authors avoided acknowledging the existence of systemic racism, sexism, and homophobia.

In this section I examined how the terms “complex” and “other,” as well as vocabulary related to culture were used to accomplish authority over student gender subjectivity. I found that these terms erased gender-based identity, universalized the individual, encouraged assimilation, and minimized the material impact of inequality. I also examined how the authors used cultural vocabulary to regulate political action. I found that they did so by describing inequality as intercultural conflict, or misunderstandings to be solved by the intercultural communication skills of communication and perspective-taking, rather than active political participation. Cultural discourse thus restrained which race-based or gender-based identities and actions were possible to imagine.

6.2.4. Governmentality as a Rational Activity

In this section I explored the strategies and reasoning used to organize the practices that regulated gender identity and action. The episteme of government, or how knowledge is implemented to rationalize governance, in this case was the reasoning behind the transfer of knowledge onto the 2018 GCF Framework Text from its bibliography in a way that prioritized a certain regime of knowledge. First, I analyzed the logic behind the selection of sources cited, and second, in order to expose the logic behind the selection of information, I traced the transfer of discourses about gender-based action and identity from a selection of sources from the Bibliography Sample⁴ onto the GCF Text. This was to detect the existence of bias in the selection of sources and information within them. Lastly, adding a historical dimension, I conducted a “micro-genealogy” of certain authors representative of significant regimes of knowledge, connecting them to broader debates in gender studies.

Many of the sources were chosen to support the concepts of intercultural competencies, even if they were outdated. For instance, the authors declared that world problems since the Cold War have been a matter of ethno-cultural conflicts (OECD, 2019a, p. 166). The works they cited were published in 1995, 2007, and 1998, each having ei-

⁴ Whenever not specifically addressing the Bibliography Sample, all other mentions of the bibliography refer to the complete list of works cited in all versions of the GCF Text.

ther a ten- or twenty-year lapse from the time of the 2018 GCF Text publication. Regardless of how often the article by Brubaker and Latin (1998) might have been cited, or the fact that it was published in a highly ranked journal, it was still an inappropriate source of information on current world issues given it was missing the last two decades. The authors also based their categorization of intercultural cognitive skills on Spiro et al. (1988), a publication outdated by three decades. These articles were not chosen because they were updated, but because they supported the intercultural regime of knowledge.

Authors often referenced themselves, or authors within their networks. Martyn Barrett and Veronica Boix Mansilla participated in the second round of the PISA 2018 GCF (OECD, 2019b). Barrett cited himself and his peer, also a professor from the UK, who collaborated with him in the Council of Europe (Huber & Reynolds, 2014, Council of Europe, 2016; Byram, 2008). Fernando Reimers (2017), who worked with Boix Mansilla at Harvard, was also cited. Thus, network connections sometimes might have influenced source selection. The only stand-alone publication of Boix Mansilla (2016) was not peer-reviewed, was cited 26 times on Google Scholar, compared to a mean of 398, and it had a 3rd quartile of the Scimago ranking, compared to most cited peer-reviewed works usually ranked on the first. The Boix Mansilla source was thus not selected based on academic merit. While one of the two versions of the 2018 GCF Text (OECD, 2019a) only cited a text by Boix Mansilla under a different author, the more polished version added her name, as well as two other works (OECD, 2019b). If the second round introduced these changes, her participation significantly influenced source selection. I therefore found that network connections and participation in the creation of the GCF were sometimes a rationale for source selection, and participation in the framework was sometimes a stronger rationale than academic merit.

Some of the intercultural studies works were chosen even when they were unpublished, or when the authors had questionable level of expertise. In the only section of the framework that mentioned agency, the authors cited a student dissertation from Hansen (2010), which contained an unpublished scale for a concept of global mindedness, and it wasn't the only dissertation cited (Doscher, 2012; Williams-Gualandi, 2015). The citation of Williams-Gualandi (2015) was particularly problematic because it was unpublished at the time GCF publication, and the author had not graduated from her doctoral program in education. The reasoning for choosing these texts could be that they were relevant to intercultural studies, or that their authors had connections to the

authors of the GCF Text; however, it could not be said that these works were chosen because of their influence on the field, or the expertise of their authors. Expertise did not seem to be a rationale for citing some texts related to intercultural studies.

In conclusion, many of the sources were not chosen because they were the most cited, most academically rigorous, or most recent. The authors sometimes tended to choose their own work, work from their peers or institutions, or work that supported the intercultural construct, sometimes with merit, and sometimes despite how outdated or poorly supported they were academically. Besides this bias I found in the filtering of works cited, there was also a filtering of the information within those works. The next section presents the results of the analysis of gender information contained in the cited texts, illustrating what information from those sources was included or left out of the framework. Table C1 illustrates the patterns of gender subjectivity and gender-based participation and agency constructs in this smaller selection of works from the Bibliography Sample. Based on the results from the category count, an article was chosen from each type of source, based on their significance to the framework.

The Oxfam (2015) curriculum was more inclusive of gender-based identities and discrimination than the framework. It mentioned “cultural, gender and racial stereotypes” (p. 13), racism and sexism, and the “impacts of historical processes (e.g. slavery and colonization) on people’s identities, cultures, and power today”; although it did not mention political participation, it encouraged students to address the root causes and effect of gender inequality at local, national, and global levels (p. 16). Ultimately, however, it did not mention political participation, had many more mentions of culture relative to gender or race, and encouraged gender equality through religious education. In contrast to the 2018 GCF, this portrayal went beyond culture, acknowledged identities based on gender and race, and the systemic and historical nature of inequality in a context-specific and thorough manner. However, it contained bias towards culture and heteronormativity, and its encouragement of gender equality through institutions that have historically had a conflict of interest with LGBTQ issues. Overall, despite a more comprehensive consideration for gender identities and systemic inequality than the GCF Text, the Oxfam curriculum was heteronormative and biased towards culture.

The academic book by Fernando Reimers (2017) had the highest level of gender-based representation of all the works cited, much higher than the GCF Text. Students were encouraged to investigate the “educational, career, and social freedoms” of a particular gender identity (“Women, Men, Lesbian/Gay, Transgender”) in different

world regions, and reflect on inequity, “gender and sex, identity,” the social construction of gender (pp. 231-238), and the different sources of their own varied identities (p. 141). Students were also asked to educate their communities (pp. 203-205) about women’s rights movements (p. 197) in relation to their current situation in global and local contexts (pp. 196, 201). This curriculum was inclusive of LGBTQ rights to gender self-determination, and comprehensively addressed gender-based identity and inequity in a context-specific way, inclusive of the individual, interpersonal, and macro dimensions. The lessons were also inclusive of some forms of formal gender-based political participation. Overall, this book was significantly more representative of gender than the GCF Text, as it addressed most of the requirements I identified for a comprehensive curriculum for gender-based identity, agency, and civic participation.

The curriculum from UNESCO (2015) created a heteronormative narrative that promoted tolerance of gay people, and like the GCF Text, it also promoted the assimilation of marginalized groups. It described students in terms of “girls and boys,” directed them to “critically question gendered roles” (p. 16), understand differences in “gender, sexuality,” among others, mentioned “sexual minorities” (p. 27), and asked students to consider gender and sexuality as collective identities (p. 43). So, the curriculum was cisnormative and heteronormative, but its understanding of gender was social and collective, not just individual and biological. UNESCO placed an emphasis on tolerance and assimilation. Students were asked to “consider common factors that transcend difference” (p. 37). The purpose was to recognize, overcome, and diminish the appearance of difference. GBV was one of the targets of conflict management (p. 39), and the elimination of homophobia, racism (p. 38), and promotion of gender equality in the classroom (p. 53). While gender inequality was recognized, civic participation was limited to conflict management and the elimination of discrimination on an interpersonal level. Overall, this UNESCO text went slightly beyond the GCF Text in recognizing heteronormative gender-based collective identity, but like the GCF Text, it also promoted assimilation and civic participation at an interpersonal level.

The journal article by Boix Mansilla (2016) was like the GCF Text in its minimization of the significance of identity and political action. Although many of the students were refugee girls, their gender, race, and nationality were not considered relevant enough to mention. The author conflated student experiences with forced displacement and volunteering abroad, and championed the “disposition to understand multiple perspectives” (pp. 2-3) of a student from one of the world's poorest countries. One student

called another student racist after he used a racist slur, and another student, when shown a picture of a woman reading among garbage, refused to find “beauty” in it; both were asked to employ perspective-taking to be more neutral (p. 5). Identity markers, which were erased, might have further justified student perspectives in these cases when perspective-taking was used to minimize their perception of injustice in favor of a neutral perspective, or where student experiences with privilege and oppression were equated. Like the GCF Text, Boix Mansilla did not address political action, prioritizing neutrality over anti-discrimination; also, she more directly used perspective-taking to minimize the importance of identity and the impact of inequality.

Like the GCF Text, the article from the Council of Europe (Huber & Reynolds, 2014) mentioned gender in universalist terms, with a focus on assimilation, and interpersonal political participation. This article includes two authors mentioned previously: Martyn Barret and Michael Byram. The authors created a dichotomy between positive social identities (i.e., gender, sports teams) and negative, “abscribed identities,” imposed by others based on “visible characteristics such as ethnicity or gender” that produce negative socialization (pp. 13-15). The imagined dichotomy of positive and negative identities erased communities built through shared experiences of marginalization, prioritizing social adaptation. They had a contradictory definition of intercultural competence, writing that it “addresses the root of... discrimination, all forms of racism,” enabling citizens to live “harmoniously within culturally diverse societies,” but also that it does not address structural inequality nor discrimination (pp. 7, 11). So, it was paradoxically described as both eradicating and not fully addressing inequality, while prioritizing harmony in diversity. They wrote that actions involved “appropriately, effectively and respectfully” handling disagreements, and also challenging prejudices by encouraging “positive attitudes” towards the positive contributions of marginalized people (p. 21). Like the GCF, this text limited political actions to the interpersonal domain, prioritizing those related to conflict elimination and assimilation rather than social justice.

Overall, I found that the information within these sources was usually more representative of gender than the GCF Text. Even an NGO source, which the category count showed had a greater bias towards culture, had more comprehensive consideration for gender identities and systemic inequality. The academic book I analyzed had the most significant representation of gender, and although it addressed most of the requirements I identified for a gender-based curriculum, these aspects were not reflected

on the GCF. The exceptions were the article by Boix Mansilla and the Council of Europe document partly authored by Barret, which both mirrored the universalist identity and neutral positioning of the GCF Text more closely. Ultimately, I found bias in the selection of both sources and information within those sources that ultimately benefitted the intercultural regime of knowledge and the view of the authors. Next, I present the results of the micro-genealogy I conducted of the two authors whom I considered to be representative of the capabilities approach and intercultural studies (which influenced the GCF Text), and then connect them to broader debates in gender studies. I contextualize these authors within the broader debates about gender subjectivity in their fields, looking specifically for bias against queer and trans identities.

After reviewing the history of Nussbaum's discussions of gender-based agency, I found evidence of explicit dismissiveness of queer and trans identity. In her criticism of Butler, Nussbaum (1999) employed cisnormative discourse, using terms like "gays and lesbians" instead of queer, using the problematic term "hermaphrodite" instead of intersex, and referencing "queer theory" without distinguishing sexuality from gender. Nussbaum mirrored the discourse of the LFSC, which accused queer feminism of not being emancipatory because it did not cater to cis women. She accused Butler's of practicing elitist "self-involved feminism" that ignored the "material conditions of *real* women" [Emphasis mine]. Whenever she minimized queer issues, she emphasized their irrelevance to cis women, writing statements like, "hungry women are not fed by this." This false dichotomy between real and imagined women was paired with a mocking attitude towards queer identity. She equated it to imagining humans having "the bodies of birds or dinosaurs," writing that a "woman dressed mannishly" wasn't new, and that lesbian drag reinforced "existing stereotypes and hierarchies of male-female society." Her statements erased trans identity, and problematized drag as patriarchal. Nussbaum's omission of queer identity was therefore embedded within a history of cisnormative feminist discourse.

Although Nussbaum's understanding of agency was cisnormative, it was gender-based, and yet, the GCF authors stripped her agency construct of any references to gender. The 2018 GCF (OECD, 2019a) relied on Nussbaum's agency construct of human capabilities, yet they misattributed this contribution to her 1997 book, "Cultivating Humanity," even though her first book about capabilities, titled, "Women and Human Development: The Capabilities Approach" wasn't published until 2000. Nussbaum's formulation of capabilities is famously distinguished by her divergence from Amartya

Sen's universalist construct, and her specification of gender-based capabilities (Nussbaum, 2011, p.18). The 2018 GCF did not just misappropriate her work, but also stripped her work on agency and citizenship education of any reference to gender. Nussbaum's 1997 publication contained chapters about Black history, women's studies, and sexuality; it also called for gender and sexuality in world citizenship education (pp. 295, 297), and mentioned the transitional aspect of sexual identity related to "coming out" (p. 250). Although it did not include trans identity nor gender-based political participation, her vast references to identity-based world citizenship was repeatedly contradicted in the 2018 GCF Text. This filtering-out of gender constructs of Nussbaum's work unveiled a universalizing rationale of governance that affected agency constructs.

Furthermore, the way Nussbaum (1997) centered her own experience and dismissed her privilege could be described as racist and oppressive. She rued her inability to play with a black girl as a child, listed the number of Black academic she worked with, demanded that gay students be more tolerant of conservatives (p. 251), and concluded her chapter on race by comparing Black academics to "snarling dogs"⁵ (p. 185). In referencing Nussbaum's text without commenting on gender and race, the authors thus could be said to have sanitized her work, removing both its good and bad elements, and reducing it to generalized commentary about universal identity.

There was evidence of Western bias in the sections of the framework text that cited Darla Deardorff. She was cited in sentences that made wide-sweeping claims that were either academically unsubstantiated or were supported by authors from her personal network. For example, in a section about a concept broadly defined as "African," the authors wrote: "concepts similar to Ubuntu found in different cultures around the world, including in indigenous cultures in the Andes and in Malaysia" (OECD, 2019a, p. 177-78). Given the omission of the names of these communities, and the citation of an unpublished conference report, this claim was difficult to verify. I found the equivalence of indigenous cultures in three continents to be tokenizing and indicative of bias. Besides the two sources from Africa that were cited in the framework, only one other source discussed Africa, a book chapter about "Ubuntu" from Deardorff's own handbook on intercultural education (Nwosu, 2009). Therefore, I found that sections where Deardorff was cited were indicative of a bias, both in terms of personal network connections and Western ethnocentrism.

⁵ This was in reference to a quotation from W.E.B Dubois that was inappropriate for her to use due to her positionality as a white woman

After reviewing the history of Deardorff's discussion of gender in her work, I found evidence of overall bias towards culture and against gender and race-based identities. In one of her many publications that failed to mention gender, Deardorff acknowledged the insufficient inclusion of indigenous voices, and admitted that intercultural competencies could become "tools of neocolonization" (2015, p. 3-4). Yet, as was shown in the category count, Deardorff consistently underperformed compared to authors in other fields in her mentions of gender, and race, and this bias extended to the entire field. Deardorff (2006) in her search for consensus among scholars on the measurement and definition of intercultural competence, mentioned cultural terms 342 times, and the term "other" in relation to people 31 times, but did not have a single mention of terms related to race, gender, or discrimination. Although Deardorff indicated she was aware of a Western bias in her own work, there were not enough strategies in place for the 2018 GCF Text to decrease bias or include authors from "diverse" fields of studies, such as indigenous, critical race theory, or gender studies. The rationale of intercultural authors like Deardorff is therefore reflective of a rationale in the GCF that prioritizes the intercultural regime of knowledge and does not address universalist Western bias.

In this section, I found that the authors employed several strategies to organize the practices of source and information selection in order to regulate gender identity and action. After analyzing the rationale behind source selection, I found that many sources did not seem to be chosen based on their influence, rigor, or currentness, but instead based on their support of the intercultural regime of knowledge, or connections to the framework authors. After tracing the transition of discourses from a sample of key works to the framework, and analyzing the rationale behind the selection of information, I identified a rationality of governance that stripped student subjectivity of gender, universalized it, and neutralized political action. Finally, the micro-genealogy seemed to expose the awareness of universalist neocolonial bias of Deardorff, a significant intercultural studies author of the 2018 GCF. It also revealed that the work of Nussbaum, another author whose work seemed influential, showed evidence of bias against trans and queer identity; yet even her cisnormative discussion of identity-based citizenship was sanitized and stripped of discourses related to gender, race, or sexuality. The rationale of governance that guided the selection of sources and information seemed to support an intercultural regime of knowledge, and to be biased towards universalist identity and neutral political action.

6.2.5. Formation of Identity

In this section, I reveal the resulting gender subjectivity and gender-based agency and action of the 2018 GCF through an examination of the questions in the Student Questionnaire Questions (OECD, 2019b) and the Cognitive Test Questions (OECD, 2019a). In this section I explored how the rationale of governance was applied to regulate gender-based identity, agency, and action in the assessment questions, and how, in order to be considered globally competent, students were made to identify with certain subjectivities. I analyzed specifically the sample questions of the cognitive exam within the framework text, which were inclusive of two scenarios in the cognitive exam. I also analyzed the questionnaire questions that mentioned gender.

The first relevant sample scenario of the cognitive test sample questions described a student complaining about a player who walked away from a game after enduring an hour-long barrage of racial slurs, and how another student realized they were also using offensive terms to refer to him. The authors interpreted this situation as an intercultural problem, evaluating students under a schemata of evaluation that examined certain knowledge objectives, mainly the intercultural “knowledge sub-domain” of perspective-taking, and the “cognitive skill” of evaluating actions and consequences. I found that this scenario made visible a white European male subjectivity, and actions within an interpersonal rather than an institutional dimension, through the technical use of universalist vocabulary, and under a rationale that stripped subjectivity and actions of gender and race-based agency. (OECD, 2019a, pp. 210-211).

This was evident from the context of the first question, the subjectivity students were expected to identify with, and the realm of actions they were expected to consider. Students were expected to be familiar with the experience of xenophobic European boys, regardless of their own marginalized experience, and to infer that the sports was men’s football, based on previous knowledge about the number of players and its international nature. The pronouns required the student to place themselves in a position of racial privilege, with phrases like, “one of your friends” and “you.” The authors used terms that minimized the appearance of structural racism, like “racial insults” in place of racist slurs, reducing marginalization to impoliteness, and “foreign-born,” which placed the emphasis on nationality rather than race. The vocabulary framed the problem as interpersonal, not institutional, referring to disembodied “regulations,” and problematizing the actions of the victimized player rather than the institution, like: “the foreign-

born player decided,” “forcing” other players to play outnumbered, and “destabilizing” them. The identity-based agency of students was thus regulated through vocabulary that focused on a universalist, Western, and privileged male subjectivity and context, and on limiting actions to an interpersonal rather than institutional dimension. (OECD, 2019a, pp. 210-11).

This was also evident from the context of the second question, its subjectivity, and the realm of actions it allowed. The authors consistently used vocabulary to obfuscate racial identity. Students were expected to infer the nationality of the player labeled “the Brain” solely based on him being the “captain” of the national team. The answer key did not address the inherent racism implicated in these nicknames, focusing instead on the “national” and “foreign” dichotomy. Implying that the player’s race was irrelevant to him being described as “the Animal,” not “smart” or “hardworking,” and with “natural gifts,” invited student to recognize the language of eugenics, or the basest form of biological racism, as irrelevant to race. Rather than encourage students to take political action against racist media portrayals, students were asked to reflect on the consequence of the “insults;” however, the answer key merely translated racist stereotypes, rather than requiring student to seriously consider, for instance, the emotional trauma resulting from racist slurs. Stripping the students and the players of racial identities, and placing the focus on their nationalities while discussing glaring incidents of institutional racism in terms of interpersonal relations, as well as narrowing political action to internal introspection, resulted in limiting race-based agency and identity. (OECD, 2019a, p. 211).

The second scenario of the cognitive test was even more relevant to gender, and more revelatory of how the authors regulated student gender-based subjectivities and action. It described a girl's YouTube music video of her singing Michael Jackson in Quechua, as well as other initiatives to revive disappearing languages, like a search engine in Quechua, and legislation to protect Australian Aboriginal languages (pp. 210-11). The authors again interpreted this situation as an intercultural problem, but this time they evaluated students under multiple cognitive skills and knowledge domains: under the intercultural “knowledge sub-domain” of identity formation in multi-cultural societies, under the institutional sub-domain of political participation and global engagement, and under two “cognitive skills” (one more or less described as perspective-taking, and the other being complex situations or problems). Although they made visi-

ble an indigenous female subjectivity, they focused on interpersonal rather than an institutional dimension though the use of vocabulary, and under a rationale that stripped subjectivity and action of historical and structural gender and race-based constructs. (OECD, 2019a, p. 211).

The answer key to the first question addressed perspective-taking, encouraging students to take on the subjectivity of an indigenous girl, and interpret her actions within an ahistorical context. They used the feminine pronoun, but since her gender was not made relevant to the problem, I found it to be tokenizing, and like the previous scenario, certain vocabulary was used to problematize indigenous people rather the systems that oppress them. They continuously referred to “her native language,” or “the indigenous language,” rather than acknowledging indigenous people, or identifying her as indigenous; they even wrote she was “an activist” but did not call her an indigenous activist. Quechua was described unfavorably as “unhip” and “backwards,” as essentially dead or needing to be “revived,” and her efforts were described as a fight against “uniformity,” insinuating that indigenous people are conformists. Clearly, the context provided for Quechua did not address the systemic colonization and racism that has resulted in the marginalization of indigenous communities in Peru, a country that was not mentioned. Lastly, despite describing her as an activist, students were asked to exercise perspective-taking rather than consider political participation. Ultimately, the first question was a missed opportunity to tackle both gender and indigenous-based political identity, agency, and action. (OECD, 2019a, p. 211).

In the second question of this scenario, the authors conflated indigenous experiences, avoided discussions of political participation, and omitted mentions of colonization. They assumed that indigenous people in Latin America and Aboriginal people in Australia have common experiences, while avoiding a discussion of colonization. In providing context, the authors omitted the well-documented history of genocide in Australia, focusing instead on the positive actions taken by the Australian government. I interpreted several phrases as blaming indigenous people for “disappearing” languages: from “young people” who think their language is “not cool,” to the “lack of” indigenous and Aboriginal teachers, written grammar, or dictionaries. Whenever indigenous people weren't actively blamed, their languages were described as almost magically “disappearing” by themselves, rather than being eradicated through a historical process of colonization. They did not ask students to think about the systemic racism that has created

negative images of Aboriginal and indigenous people, nor encourage students to participate in helping pass legislation. In the second question, the authors used vocabulary to minimize the impact of colonization, and in their description of political action, minimized identity-based agency. (OECD, 2019a, pp. 210-11).

The sample scenarios given for the cognitive test revealed how the authors formed their own identity in relation to students, in order to exercise authority over their identity and political agency. The authors assumed authority over students by describing instances of racism as intercultural or interpersonal matters, extending their expertise of intercultural studies to exercise authority over the evaluation of student perception of mostly racial discrimination. They expected students to interpret situations within an interpersonal dimension, to limit political action by projecting their perspective inwards, to reflection, and to take a position that privileged the Western or male perspective. They enforced this through their grading of a rubric, whose answer key indicated whether the students were sufficiently globally competent. In these questions they evaluated students based on their ability to identify with a Western male perspective, or an ahistorical Western perception of indigenous people.

In the questionnaire (OECD, 2019b, Annex C, p. 50-61), there were two questions that tackled gender equality. The authors made technical use of vocabulary and their evaluation framework to frame gender equality as an intercultural problem. The first question about gender was categorized under the label “awareness of global issues” (OECD, 2019b, p. 51). As was shown in Figure 1, there was no specific evaluation criteria for examining “awareness,” other than the general description in the framework of global issues as affecting “all individuals, regardless of their nation or social group” (OECD, 2019a, p. 70). In this question, students were asked about the extent to which they were informed about several topics, such as “equality between men and women in different parts of the world” (OECD, 2019b, p. 51). The second question was categorized under the label “student's engagement (with others) regarding global issues” (p. 55). This phrasing was different from the knowledge subdomain against which it was evaluated, “political participation and global engagement” (OECD, 2019a, p. 190), since it removed the element of political participation. Students were asked whether they were involved in a few activities, among which was the statement “I participate in activities promoting equality between men and women” (OECD, 2019b, p. 55).

Both questions limited action to the interpersonal domain, and subjectivity as universal cisnormativity, while ignoring the gender-based agency of the students. Both

questions specified that inequality existed outside of their community, as a matter of “global issues” that happened in “different parts,” which increased distance between the subjectivity of the student and the genders of the people involved in equality. The terms “men and women” made visible only a universalist, cisnormative and heteronormative subjectivity, erasing People of Color and LGBTQ people, without recognition the systemic or social construction of gender and sexism. The first question addressed “awareness,” an internal process. The second question contained the term “engagement,” not political participation, and students were not asked about “political engagement” but “engagement (with others).” The terms “promoting” and “others” made the activity and people involved (and therefore the dimension of action between micro and macro) unclear, but, given the removal of the term “political participation,” I concluded that this engagement referred to interpersonal communication. I therefore found that gender equality was limited to cis, heterosexual people, and inequality was portrayed as an interpersonal rather than a systemic problem, to be addressed through individual consciousness-raising and interpersonal communication.

In this section, I revealed the resulting gender subjectivity and gender-based agency and action of the 2018 GCF through an examination of the questions in the student questionnaire and the cognitive test. I explored how the rationale of governance was applied to regulate gender-based identity, agency, and action in the assessment questions, and how, in order to be considered globally competent, students were made to identify with certain subjectivities. I analyzed specifically the sample questions of the cognitive exam within the framework text, which were inclusive of two scenarios in the cognitive exam. I also analyzed the questionnaire questions that mentioned gender.

After analyzing the sample scenarios given for the cognitive test, I found that the authors employed the technical use of vocabulary to interpret colonization and racism as an intercultural matter, allowing them to exercise their expertise of intercultural studies to evaluate students’ perception of racism and colonization. They limited action to the reflective and interpersonal dimension, and used contexts and subjectivities that privileged a universalist, Western, or male perspective. In the questionnaire, the questions similarly created a universal, cisnormative, Western subjectivity, and limited action to the interpersonal rather than systemic domain, disregarding any possible gender-based agency in students.

7. DISCUSSION

The main research problem I sought to tackle was the insufficient amount of research that has been conducted to hold governments accountable for the extent to which their civic curricula prepares citizens to fulfill their commitments to gender equality. Much research has shown that large-scale international studies like the 2018 PISA GCF act as a hidden curricula, but it is important for research to go further, and hold PISA accountable for its assessment of gender-based competencies. Overall, the research about gender-based identities and competencies in PISA has been inadequate, and insufficient research has been done on how this curriculum shapes student development of citizenship identity and civic competencies in relation to gender. My goal in this thesis was to illustrate how the process of data production of international assessments like PISA generate discourses about gender-based political agency that might shape policies about civic education for gender equality. Specifically, I first sought to identify the discourses or regimes of knowledge that were used to regulate gender subjectivities in the 2018 PISA GCF. Secondly, I sought to identify the extent to which the 2018 GCF measured the gender-based political agency and political participation of students, if at all.

The results of the quantitative category count indicated the existence of selection bias, where intercultural sources and terminology were disproportionately chosen over sources and information related to gender. Both the GCF text and the sample of cited works were highly heteronormative, cisnormative, and had more terms related to culture than gender-based terms, and unlike some cited texts, the framework omitted LGBTQ and feminist terms. These findings seemed to be related to which fields of study were cited, since intercultural studies had a disproportionately high word count in the sample, while works in the fields of gender, race, and inequality were excluded. Geographically, Western sources had a disproportionately higher representation. I found that these exclusions had a significant impact on the representation of categories in the works cited, since intercultural sources had less gender terms. After conducting a category count of excluded gender studies authors in comparison to those influential to the 2018 GCF, I found that most texts were highly heteronormative and cisnormative, except for trans and queer theory authors. The category count thus indicated that the selection of sources

and information in the 2018 GCF was biased towards Western, cisnormative, heteronormative, and intercultural discourse.

The results of the Foucauldian discourse analysis indicate that the 2018 PISA GCF regulated gender subjectivities through an intercultural regime of knowledge. World problems were described as conflicts caused by diversity, and resolved through knowledge objectives and a schemata of evaluation that prioritized cultural assimilation, neutrality, and intercultural skills over other concepts. The bibliography was used as a tool, or diagram of power, that illustrated bias towards citing intercultural authors over gender studies authors; this resulted in the omission of gender-based identity. Identity was not formulated based on gender, making visible only a Western, white, heterosexual, and cis-gendered point of view, and otherizing LGBTQ identities. The tokenizing use of female pronouns, and the scarce mentions of female identity promoted a universal subjectivity. Political action was formulated as value-neutral interpersonal interactions, and gender-based agency was actively discouraged. Subjectivity was universalized, and political action neutralized using vocabulary, with terms like “complex” and “other” erasing gender-based identity, and intercultural terms promoting universalization and assimilation. These findings indicate that intercultural discourse was used to misrepresent systemic inequality as intercultural conflict to be resolved by neutral universal agents (i.e., students) that use intercultural skills to improve interactions.

The findings also demonstrate that the overall strategy was to rationalize the governance of an intercultural regime through the transfer of knowledge from the cited works onto the framework text in a way that prioritized intercultural studies. I detected bias in the selection of sources, with authors choosing those that either supported the intercultural regime of knowledge, or were connected to GCF authors, despite their lesser influence, rigor, or currentness. After tracing the flow of information from cited sources to the final GCF text, I identified a rationality of governance that stripped student subjectivity of gender, universalized it, and neutralized political action. The micro-genealogy exposed the universalist neocolonial and cisnormative biases of significant authors, whose concepts were further stripped of references to gender, race, or sexuality.

The results further suggest that global competence was evaluated based on identification with a Western male perspective, or an ahistorical Western perception of indigenous people. To accomplish authority over student subjectivity in the cognitive test, and limit action to the reflective and interpersonal dimension, the authors described instances of systemic racism and colonization as interpersonal matters, which allowed

them to evaluate student views of systemic inequality. In the questionnaire, the authors made technical use of vocabulary and their evaluation framework to frame sexism as an intercultural problem, creating a universal, cisnormative, Western subjectivity, and limiting action to the intra and interpersonal, rather than systemic domains.

Overall, I found that the results met my expectations. First, I identified the intercultural regimes of knowledge, and how the bibliography and vocabulary were used to restrict gender subjectivities in the framework to a heteronormative, cisnormative, and universal Western subjectivity. Second, I found that the framework restricted the measurement of gender-based political agency and political participation of students by restricting these concepts to the interpersonal and intercultural, rather than the systemic dimensions, and attaching them to neutrality and harmony rather than to social justice.

My findings that the 2018 PISA GCF neglected gender equality is supported by the definitions of gender-based citizenship education in feminist research. Feminist scholars found that appropriate portrayals of gender equality include aspects like well-being, embodiment, or social relationships, and take the agency, autonomy, and voice of women seriously (Unterhalter, 2017, pp. 7-10). The GCF erased feminism issues and employed a tokenizing use of female gender pronouns. Feminists have found that gendered citizenship education should center local political contexts (Fennell & Arnot, 2008), and spread awareness about race, class, gender, and specific issues of gender equality (Arnot, 2009A). The GCF rarely addressed these issues, and when it included indigenous identity, the context retained a Western perspective. Gender responsive curricula is transformative, deconstructs gender binaries, and encourages intersectionality (Gunnarson, 2009), and some citizenship curricula portray citizenship in explicit terms, with particularity (Erel, 2010; Sant, 2019). The GCF was predominantly white, cisnormative, and portrayed citizenship identity as universal. While some feminists considered identity as flexible and collectively defined (Pulkkinen, 2000, p. 137) the GCF considered identity individualistic, interpersonal, and fixed. Overall, the findings of feminist research support my findings that GCF neglected gender equality.

According to the theoretical framework I implemented in my analysis, the 2018 GCF erased several elements of gender identity, agency, and action. Eteläpelto, Vähäsantanen, Hänn, and Paloniemi, (2013) described identity as having a material, socially constructed, embodied, and temporal (life-course) dimensions. They also described agency as the mediating mechanism between the social and material selves. The 2018 GCF only recognized the material dimension, without addressing the discursive, embodied,

or temporal dimensions. Political agency and action were also not based on gender. While students were asked about gender equality, they were not specifically asked about their own gender identity, nor about their participation in politics on behalf of their gender identity (Holli & Wass, 2010). The GCF did not outline gender-based skills or intention in a manner that was concrete, explicit, or action-specific. It also did not engage formal political action, limiting action to the interpersonal dimension. Students were assumed to have a material or natural gender identity, and were not expected to change their relationship with their gendered bodies or their socially constructed gender identity throughout their lives. The agency that resulted from this universal and static identity was thus not motivated by gender, nor measured according to formal political action.

This portrayal of gender marginalized the identities of LGBTQ students. There was an erasure of the physical, psychic, and socially constructed identities embodied in experiences of *passing* (Snorton, pp. 87, 89) or *coming out*. The insufficient institutional recognition of trans identity means that it is not appropriately legitimized, robbing trans students of their right to self-determination (Jansen, 2016; Zimman, 2019). As shown from the category count, identity concepts were also not determined through an inclusive and open dialogue with academics, activists, or members of civil society who represent feminist, queer, or transgender groups. The 2018 GCF was thus not inclusive of feminist and trans identities, nor the academic communities behind them.

While others might have more generous interpretation of the GCF, such as the use of the term “complex” as an attempt at intersectionality, I maintain that these terms were not used for the sake of inclusivity. Complexity was used in the cognitive test to separate advanced students from basic students: advanced students saw girls as “complex” individuals that “can be at once a girl, a daughter, a farmer, and a citizen”; as opposed to “basic” students who used “relatively fixed... markers of a person’s identity,” and one-dimensional categories of “nationality or religion.” (OECD, 2019a, p. 187) Some could view this as an attempt at intersectionality, or the view that “race, class, gender, sexuality, ethnicity, nation, ability, and age operate not as unitary, mutually exclusive entities, but as reciprocally constructing phenomena” (Collins, 2015, p. 2). While some would argue that intersections of identity are addressed, this is not sufficient. In fact, “intersectionality’s *raison d’être* lies in its attentiveness to power relations and social inequalities” (Collins, 2015, p. 3); and so, without acknowledging oppression, a curriculum cannot be intersectional. My findings that terms like complex were

used to regulate gender identity cannot be reinterpreted as attempts at intersectionality because systemic inequality was inadequately addressed.

While some might argue that the rationale of governance was to promote anti-discrimination, this interpretation is contradicted by the consistent ambivalence towards a concrete discussion of systemic racism. The framework was marked by contradictions. While diagnosing the problems that the framework was to address, the authors claimed to support cultural diversity, yet also championed cultural assimilation, even though diversity is the direct opposite of assimilation. In the description of racism, specifically, the authors contradicted themselves, encouraging teachers to talk about racism, without being willing to make a similar stance. Intercultural bias was evident in the assumption that evaluations of intercultural skills were reliable while the evaluation of values were not, in spite of the existence of the World Values Survey, a large-scale international adult study (WVS Database, 2021). The values of human dignity and diversity were excluded from the questionnaire, yet the virtually indistinguishable knowledge subdomains of prevention of hate crimes and human rights were included, nonetheless. The authors used the word “hate crime” without linking it to its origin in a history of violence against Black people in the US during the Civil Rights Movement (FBI, 2016, May 3). These instances indicate an ambivalence towards discussions of race.

Like a previous study, I found a tokenizing treatment of Black and African identity (Ledger, Thier, Bailey, and Pitts, 2019), but my findings further indicate that the general ambivalence towards race was a matter of colorblind racism. For instance, some might assume the authors were unaware of the link between race and hate crimes; however, they were explicit, writing that hate crimes were sensitive for “a student from a minority group” (OECD, 2019a, p. 190). Critical race theorists consider the use of the term “minority” as a substitute for “black” problematic, and part of a history of racist portrayals of Black people (Kendi, 2017, p. 394). My findings revealed a pattern in the GCF of using similar racist pseudonyms and avoiding race. This was especially egregious when the authors avoided using the “racist” while describing a situation where a teacher disciplined black students during a lesson about the Civil War, a war over slavery. This was a clear perpetuation of the school-to-prison pipeline, the systemic use of the educational system as a venue for racist incarceration (Wald & Losen, 2003). The misrepresentation of glaring examples of systemic racism, and the use of racist pseudonyms is not just inadequate, it is inherently racist.

While some might argue that my analysis of systemic racism and the parallels I drew to systemic sexism, homophobia, and transphobia are not consistent with my research task, this is in fact a misinterpretation of my theoretical lens. The application of intersectionality as a social theory is a heuristic approach that allows researcher to shift from a one-dimensional paradigm of either race or gender, to a multi-faceted approach that captures the processes that connect these identities (Collins, 2019). From this standpoint, I was able to analyze the heteronormative, cisnormative, and universalist discursive practices, and find the intersecting gaps it created, through which certain students “fell through,” allowing the authors to intentionally or accidentally exclude students from the sample. This allows me to conclude that the 2018 GCF almost completely erased the identities of Women of Color and LGBTQ People of Color who exist in the intersection between gender, race, material, and discursive identity.

My finding that the rationale of governance of the GCF was based, not on evidence, but on upholding the intercultural regime, is consistent with how differently the 2018 was described in the publication of the “PISA 2018 Results” (OECD, 2020). It replicated the original 2018 GCF Text (OECD, 2019a) in their introduction, with some updates in response to events in the interim years. Seemingly unable to avoid identity-based movements that contradicted the 2018 framework, the authors reframed them to validate intercultural discourse. They mentioned the “#MeToo movement” by name, but described the #BlackLivesMatter movement as “protests challenging racial discrimination,” writing that both movements shared a “willingness of people to show solidarity and to take action for collective well-being,” (OECD, 2020, p. 54). They added 15 mentions of gender equality in the 2020 results, writing that the 2018 GCF “focused on... gender equality... migration and cultural diversity” (OECD, 2020, p. 54); yet, in the original framework, the value of cultural diversity had been removed and gender equality was mentioned once. These attempts to “fix” the 2018 GCF text, without changing the data, corroborate my findings that the rationale of governance was to uphold discourse, not find truth.

Most importantly, these retroactive changes highlighted the malleability, and relative uselessness, of GCF gender-based agency constructs, since the authors swiftly reformulated agency as needed, without meaningfully changing the data. The original term of “global mindedness” (OECD, 2019a, p. 176) changed to “agency regarding global issues” (OECD, 2020, p. 63). This re-phrasing of agency to make it sound more active did not improve the utility of the data. The extent of their findings were that more

than 50% of students took some action to promote equality between men and women in some countries, and therefore generalized cultural knowledge resulted in more action about a range of global issues (i.e., gender equality). Rather than producing data about the nature or extent of this participation, the study was limited to reaffirming intercultural knowledge. International assessments sometimes focus on establishing unnecessary empirical connections, rather than on the utility of their findings in classroom learning (Hyslop-Margison, Hamalian, and Anderson, 2006, p. 401). If one were to judge data about global competencies based on its utility, one could not conclude that it sought to facilitate classroom learning around gender equality, but to place students on a hierarchy in relation to an intercultural regime of knowledge.

While the discourse changed, it had no effect on the execution of the assessments, so it only served to sanitize identity erasure. The section that included the lesson on the Civil War, and pseudonyms for race (OECD, 2019a, p. 178-179) was removed. I analyzed a section titled, “Perspectives on global competence from different cultures,” which drew parallels between global competencies and tokenistic mentions of “Ubuntu” and nameless “indigenous cultures” (OECD, 2019a, p. 177-78). This section was replaced with the section, “The universal roots of global competence,” a description of languages and religions in an explicit attempt to universalize them, or draw from “older, more universal roots” (OECD, 2020, p. 58-59). The term “complex” was not used to scope out values; instead, the methods required for it were described as too difficult: “a broad repertoire of assessment strategies, ranging from interviews or conversations to observations” (OECD, 2020, p. 64). Lastly, the false attribution of Nussbaum’s concept of capabilities was corrected (OECD, 2020, p. 64). These changes present difficulties for future analysis, making my analysis even more important because it deconstructs the foundational terms of the GCF before they become more sophisticated through more insidious discursive practices.

My results build on existing evidence of bias in the PISA GCF, and my approach to discourse analysis offered a stronger political analysis of discursive institutional power. I offered a better political insight into the methods through which concepts of gender were regulated. The micro-genealogy allowed me to problematize or destabilize the gender histories hidden beneath the intercultural regime of knowledge, by focusing on the institutional exclusion or subjugation of certain texts (Ball, 2013), or in this case, the entire discipline of gender studies. Following good quality Foucauldian discourse analysis, rather than interpreting the meaning of text with critical detachment,

I focused on the methods of institutional control that regulated the discourse (Ball, 2015). I did not interpret textual meaning through key word analysis, as Ledger, Thier, Bailey, and Pitts (2019), did in their study of the OECD's global competence framework. My findings mirrored some of theirs: identifying the higher authority of certain authors, inconsistency of terms, and the Western colonial ideologies and subjectivities that it prioritized, and my findings provided a new insight into the exclusion of gender and race-based identity, and how authority was accomplished through technologies of power.

Ultimately, my thesis countered the claim in the PISA GCF that its concepts contribute to evidence-based policy or that they were purely informed by research. I interpreted the drafting of the framework and construction of its concepts as a struggle within and between communities of "experts" over the inclusion of their work. The authors would metaphorically switch roles between being clients, asking questions, and then acting as detached "expert" consultants answering them. It seems that the term "global competencies" emerged from business literature (Cascio & Boudreau, 2016), yet the failure of authors to establish themselves within this or other academic traditions reduces the ability of researchers to hold them accountable to one. Even after months of re-reading the material, it was difficult for me to parse through the distinction between these concepts, which mostly depended on phrasing, because they were not each tied to or constrained by a body of research. The resulting instability in the terms and their divergence in the questionnaire, the cognitive test, the framework text, and the results document indicated that the terms for global competence were not sourced from academic discipline, but chosen instead based on their malleability.

7.1. Reliability

One of my key concerns in this thesis was meeting the standards for rigorous qualitative research. I found this challenging because the high level of discretion that discourse analysts are afforded (largely due to the influence of postmodern authors such as Foucault) made it difficult to establish a set of criteria for evaluation. There are no set formats or guidelines for the implementation of discourse analysis in qualitative research (Dunne, 2012, p. 130). Because of this, I encountered what Stephanie Taylor (2001) describes as the "crisis of legitimation" that is created when graduate students

and researchers pursue a postmodern approach while being subject to institutional empirical requirements. Students embark on a seemingly futile search for legitimacy and empiricism in a field where the ontological view of reality is that there is no objective truth, and without established procedures for evaluating the knowledge they produced. (pp. 11-12) While other methods would have been less difficult in this regard, I chose to embrace this challenge as an opportunity to grow as a researcher.

Refashioning a “new” methodology did not lessen my commitment to high standards of academic accountability. Some authors have indicated a criteria to guide good quality Foucauldian discourse analysis, which includes: pursuing intellectual freedom while respecting the expectations of a research community; avoiding a dogmatic interpretation of a Foucauldian approach outlined by another researcher, and instead integrating those insights into the application of one’s own approach, describing it clearly, and conducting it in direct conversation with the work of Foucault; and finally, abstaining from prescribing a model for other researchers to follow. (Graham, 2005, p. 6) I believe that I practiced these principles. Although most of my approach emerged from the outline set by another researcher, I did not follow it dogmatically; I adapted it by including a micro-genealogy as well as feminist emancipatory aims. I do not recommend that this approach be replicated exactly by other researchers, and recommend instead that they adjust this approach to their own policy material and social context. I exercised intellectual freedom by embracing the challenge of pursuing an innovative adaptation of Foucauldian discourse analysis, while respecting the qualitative requirements listed in the next few paragraphs.

I examined the reliability of my findings by carefully considering the extent to which they met the expectations of a community of qualitative researchers. Sarah Tracy (2020) described the different expectations of different communities of researchers. Researchers in the positivist paradigm conduct mostly quantitative research, and aim for objectivity, statistical generalizations, and reliability (i.e., stability and replicability). In my research, I navigate the intersection between the interpretive, critical, and postmodern paradigms; communities that conduct qualitative research, question objectivity, and emphasize in-depth information over generalization. I made myself accountable to Tracy’s eight criteria for qualitative research: social relevance, rich rigor, sincerity, credibility, resonance, significance, ethics, and meaningful coherence. I covered the social relevance of gender-based agency and PISA studies in my problem statement, significant contributions are discussed in the section about the applicability of my findings,

and the ethics are discussed in the section about ethical concerns. In the discussion section, I described how I established meaningful coherence of my thesis by achieving my research purpose, describing how methods and procedures fit those goals, meaningfully interconnecting the literature, research questions, foci, findings, and interpretations. Sincerity was addressed in two parts: in the limitations section, I discuss with sincerity, and in a vulnerable way, my original goals, and my mistakes; and in the ethics section, I employ self-reflexivity to explore my positionality. So far, I have addressed some of the expectation for good qualitative research. (Tracy, 2020)

In these paragraphs, I discuss with more detail the rich-rigor, credibility, and resonance of my thesis. Studies with rich-rigor exercise a sufficient, abundant, appropriate, and complex employment of theoretical concepts, data collection and analysis (Tracy, 2020). I clearly defined theoretical concepts in both the literature review and the theory section. My collection of data and choice of data samples was intricate and thoughtful. I did not just use the GCF, but also explored the texts it cited, as well as the work of academics that influenced it. The method section includes an in-depth discussion of the chosen method of discourse analysis in comparison to others, and how data was processed, chosen, and organized in spreadsheets. Overall, I exercised rich-rigor by making choices not on convenience, but based on the effort needed to meet due diligence.

A study with sufficient credibility is characterized by thick description and triangulation (Tracy, 2020). I practiced of thick description in the text analysis by providing contextual information, and exposing tacit knowledge, or focusing not only on what was said, but what was excluded, silenced, or ignored. I provided contextual meaning for certain groups. In the data analysis I provided context for Black students (in describing the school-to-prison pipeline) and indigenous students (in describing the context of colonization); in the theory section, I provided context for trans, queer, and LGB students (in describing experiences of “passing” and “coming out”). I included complete sections of text in findings section in order to be transparent about the context of the quotes. A huge part of my discourse analysis approach was the disclosure of tacit knowledge, or knowledge that floats beneath the surface. I solidified this in the discussion section by connecting quotes not just to the surrounding text, or my interpretations, but with concepts I explored in the theory section and the literature review. I used the quantitative approach of content analysis to triangulate my qualitative data, therefore reducing bias. I believe that the triangulation of data and thick descriptions established sufficient credibility for my findings.

Resonance is established differently in qualitative research than in quantitative research. Rather than employing quantitative generalizability by using large random or representative samples, resonance in qualitative research is established through the way the text influences the reader, and the transferability of the findings (Tracy, 2020). A main challenge in qualitative research is how the researcher can look at “excluded” or “missing” data. The way I systematically illustrated and dissected the insidious racism in the sections about the lesson on the Civil War, for instance, is transferable to other research settings. The way I connected the “missing” subjectivities with the “missing” authors in the bibliography, may allow other researchers to look past the available data in each education policy text. My identification of data in policy texts that is excluded, or made invisible, in a way that is tangible and visible may resonate with other theorists in queer, trans, critical race, and feminist studies that analyze race-based and gender-based identity concepts.

7.2. Limitations

In this section, I discuss in a vulnerable way, my original goals, and my mistakes in all the phases of research. My original goal was to conduct a discourse analysis of both the PISA and the ICCS assessments, and while I collected data for both, I had to narrow my scope due to a lack of time, and complications due to me becoming sick with COVID-19. Another challenge was the writing structure, since my desire to be thorough made my graduate thesis longer than intended.

One of the most challenging aspects of this thesis was the inclusion of so many different concepts and theoretical lenses. My efforts towards increased inclusivity required a broader scope than was perhaps advisable for a master’s thesis, which might have resulted in the insufficient exploration of certain concepts. I attempted as best I could to meet this demand through rigorous academic scholarship, and a methodical and systemic literature review. However, it is possible that some concepts were explained with insufficient clarity. The results were presented in a clear way, following the structure of Mitchel Dean’s governmentality approach. However, there could have been other ways to present the data that would have been more directly related to my research questions, and more specifically focused on gender.

Another challenge was the integration of different methods of discourse and content analysis. A more singular focus on one method might have been more rigorous

and cohesive. Although this limited the level of rigor possible, it allowed for more contextualized, more significant, and more in-depth findings.

Specifically, the quantitative content analysis had several limitations. Not all sources were analyzed for category count; some documents were not accessible online. Moreover, I was not able to extract the total word-count with exactitude, so it was excluded, and I relied on the relative frequencies of terms, rather than the proportion of terms related to the total word count. Although the counts were organized and reviewed several times, there is some space for human error. Also, it is possible that a higher word count at times does not reflect the true nature of the content, since it is possible for a source to discuss the topic of gender without including the specific keyword being searched. Keywords do not always reflect the quality of the content, so the category count findings are more limited.

Most importantly, the Chi-Square goodness-of-fit test is a flawed method of content analysis that has been discredited by some researchers. Categories are not mutually exclusive in a text. These researchers suggest using the *R* software to conduct a more thorough and mathematically sound analysis. (Bestgen, 2017) I could have learned how to use that software but decided not to because I believe I stretched my current skillset to its limit; however, I encourage other students and researchers to do so. I could have decided to remove all my quantitative data, but I decided against it because while flawed, and arguably extreme ways, the data still adds some value. Some of the patterns, such as the higher quantitative presence of cultural terms in comparison to gender or race terms in both the GCF Text and intercultural studies texts, cannot be dismissed.

I found the pressure to produce “real” data via quantitative results, as a student whose strength and training are in the interpretive paradigms, to be a slippery slope that pushed me far out of my comfort zone. I embraced this discomfoting challenge, and I hope to be assessed not by the perfection of my data analysis or results, but by the accuracy and transparency of my research narrative. I have done my best to transparently walk the reader through an iterative and murky journey of discourse analysis, in hope that it can intrigue others to shine new light on this avenue of research, and help the field hold itself accountable to young women in their path towards social justice.

7.3. Applicability of Research Results and Challenges for Further Research

A study that provides a significant contribution does so conceptually, practically, heuristically, methodologically, and extends, builds, or critiques disciplinary knowledge (Tracy, 2020). I developed the concept of gender-based agency into a framework for feminist citizenship education in a way that went beyond the existing literature. Future research may benefit from applying this framework and expand on it. I discussed the catalytic validity, or how this research may push others to act, in my problem statement. Based on this research, education policymakers and researchers can push for better quality data from international large-scale assessments.

For researchers, this thesis approached methodology in a new and exciting way. On a practical level, my thesis can serve as a guide for graduate students and other researchers to apply Foucauldian method of discourse analysis creatively. Researchers might benefit from my application of a micro-genealogy to the tracing, not just to the discourses in an academic text, and to the discourses in the works cited in its bibliography, but also to the discourses excluded from the bibliography. Other researchers could further problematize how bibliographies are used as objects that give an author power over the reproduction of discourse, allowing them to regulate, restrict, or expand the subjectivity within these texts. Bibliographies allow certain authors to elevate the power of institutions they are affiliated with, academics they collaborate with, entire disciplines of study, and through self-referencing, their own work. Most importantly, bibliographies also allow these same authors to subjugate, erase, and exclude entire disciplines, institutions, and authors. Other studies might explore these relations of power, or expand on the approaches of investigation by integrating different methods of tracing, not just the language, but the histories of discourse.

The results about gender-based agency and identity could be considered in the formulations or critical analysis of new large-scale international assessments of student citizenship with relation to identity. In the field of gender studies, researchers could further explore how measures of political participation consider feminist motivations and gender identity, and how such measures limit or expand the possibilities of data production in ways that could help marginalized groups gain more visibility and power. Researchers in queer and transgender studies could problematize limited units of analysis

that do not consider the transitional aspect of gender identity, its fluidity, and the different stages of psychic, embodied, or material identity. Similar problematization could be beneficial in the fields of critical race studies, and intersectionality studies.

In the field of education policy, these results can help policymakers and policy researchers question the data they use to design new policies about citizenship education in relation to gender. Increased awareness about the limited ways that large-scale international assessments explore gender-based agency and identity in citizenship can help them demand better quality data and help them push to close the gender data gap. Most importantly, this thesis could be beneficial to the elaboration or updating of large-scale international assessments in a way that takes gender-based citizenship education more seriously, and takes into account a broader pool of academic disciplines. The integration of transgender, queer, feminist, decolonial, intersectional, and critical race theorists into these bibliographies could address some of the ethical issues identified here.

7.4. Conclusion

According to the theoretical framework of gender-based agency that I implemented in this analysis, the 2018 Global Competencies Framework (GCF) erased several elements of gender-based identity, agency, and action. My findings indicate that bibliographies and vocabulary can be used in large-scale international assessments to push an intercultural regime of knowledge and restrict gender subjectivities to heteronormative, cisnormative, and universal Western identity. In the case of the 2018 GCF, it resulted in the erasure of indigenous, Black, female and LGBTQ identities, and even in colorblind racism. This formulation of global competencies results in the restricted measurement of gender-based political agency, the limitation of student understanding of political participation to interpersonal relations, and the encouragement of students to be neutral rather than to act for social justice.

8. REFERENCES

- Abowitz, K. K., & Harnish, J. (2006). Contemporary Discourses of Citizenship. *Review of Educational Research*, 76(4), 653–690. <http://www.jstor.org/stable/4124417>
- Ahmed, R., & Hyndman-Rizk, N. (2020). The higher education paradox: Towards improving women's empowerment, agency development and labour force participation in Bangladesh. *Gender and Education*, 32(4), 447-465.
doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1080/09540253.2018.1471452>
- Alemán, S. M. (2018). Mapping intersectionality and Latina/o and Chicana/o students slong educational frameworks of power. *Review of Research in Education*, 42(1), 177–202.
<https://doi.org/10.3102/0091732X18763339>
- Alniaçik, A., Göksen, F., & Yüксеker, D. (2019). School to work or school to home? An analysis of women's vocational education in turkey as a path to employment. *Gender and Education*, 31(8), 1040-1056.
doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1080/09540253.2018.1465897>
- Alzeer, G. (2018). Types and dynamics of gendered space: A case of Emirati female learners in a single-gender context. *Gender and Education*, 30(3), 360-378.
doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1080/09540253.2016.1205183>
- Andrews, A. (2014). Women's political engagement in a Mexican sending community: Migration as crisis and the struggle to sustain an alternative. *Gender & Society*, 28(4), 583–608. <https://doi.org/10.1177/0891243214523124>
- Archer, M. S. (2000). *Being human: The problem of agency*. Cambridge University Press.
- Arnot, M. (2009A). A global conscience collective?: Incorporating gender injustices into global citizenship education. *Education, Citizenship and Social Justice*, 4(2), 117–132.
<https://doi.org/10.1177/1746197909103932>
- Arnot, M. (2009B). *Educating the gendered citizen: Sociological engagements with national and global political agendas*. Abingdon, Oxon; New York, NY: Routledge
- Arnot, M., Chege, F. N., & Wawire, V. (2012). Gendered constructions of citizenship: young Kenyans' negotiations of rights discourses. *Comparative Education*, 48(1), 87–102.
<https://doi.org/10.1080/03050068.2011.637765>
- Arnot, M., & Dillabough, J. (2000). *Challenging democracy: International perspectives on gender, education and citizenship*, Routledge Falmer.
- Aronson, B., & Laughter, J. (2016). The theory and practice of culturally relevant education: A synthesis of research across content areas. *Review of Educational Research*, 86(1), 163-206. Retrieved March 2, 2020, from www.jstor.org/stable/24752872

- Arribas-Ayllon, M., & Walkerdine, V. (2017). Foucauldian discourse analysis. In S. W. Rogers & C. Willig (Eds.), *The SAGE handbook of qualitative research in psychology* (Second ed., pp. 91–108). SAGE Publications Ltd.
- Baird, J., Johnson, S., Hopfenbeck, T. N., Isaacs, T., Sprague, T., Stobart, G., & Yu, G. (2016). On the supranational spell of PISA in policy. *Educational Research*, 58(2), 121-138. doi:10.1080/00131881.2016.1165410
- Ball, S. J. (1994) *Education reform: A Critical and post-structural approach*. Buckingham: Open University Press.
- Ball, S. J., & Olmedo, A. (2012). Care of the self, resistance and subjectivity under neoliberal governmentalities. *Critical Studies in Education*, 54(1), 85-96. doi:10.1080/17508487.2013.740678
- Ball, S. J. (2013). *Foucault, power, and education (Routledge Key Ideas in Education)* (1st ed.). Routledge.
- Ball, S. J. (2015). What is policy? 21 years later: Reflections on the possibilities of policy research. *Discourse: Studies in the Cultural Politics of Education*, 36(3), 306-313. doi:10.1080/01596306.2015.1015279
- Berkovich, I., & Benoliel, P. (2019). Understanding OECD representations of teachers and teaching: A visual discourse analysis of covers in OECD documents. *Globalisation, Societies and Education*, 17(2), 132-146. doi:10.1080/14767724.2018.1525281
- Baroutsis, A., & Lingard, B. (2017). Counting and comparing school performance: An analysis of media coverage of PISA in Australia, 2000–2014. *Journal of Education Policy*, 32(4), 432-449. doi:10.1080/02680939.2016.1252856
- Bestgen, Y. (2017). Getting rid of the Chi-square and Log-likelihood tests for analysing vocabulary differences between corpora. *Quaderns de Filologia: Estudis Lingüístics* 22(12), 33-56. doi:10.7203/qf.22.11299
- Bhanji, N. (2013). Homing desires, (trans)sexual citizenship and racialized bodies. In Stryker, S., & Aizura, A. (Eds.) *The Transgender Studies Reader 2*. (pp. 512-526). Routledge
- Bhog, D., & Ghose, M. (2014). Writing gender in: Reflections on writing middle-school political science textbooks in India. *Gender & Development*, 22(1), 49-62. doi:10.1080/13552074.2014.891830
- Biseth, H. (2009). Democracy and education in a multicultural Scandinavia: What mandate is designated to educators? *Intercultural Education*, 20 (3), 243–254. <https://doi.org/10.1080/14675980903138590>
- Blaskóá, Z., Dinis da Costaa, P., Vera-Toscanob, E. (2019). Non-cognitive civic outcomes: How can education contribute? European evidence from the ICCS 2016 study. *International Journal of Educational Research*, 98, 366-378. <https://doi.org/10.1016/j.ijer.2019.07.005>

- Boix Mansilla, V. (2016), "How to be a global thinker", *Educational Leadership*, Vol.74/4, pp. 10-16.
- Bonal, X., & Tarabini, A. (2013). The role of PISA in shaping hegemonic educational discourses, policies and practices: The case of Spain. *Research in Comparative and International Education*, 8(3), 335–341. <https://doi.org/10.2304/rcie.2013.8.3.335>
- Bondy, J. M. (2016). Latina youth, education, and citizenship: A feminist transnational analysis. *Theory & Research in Social Education*, 44(2), 212-243. DOI: 10.1080/00933104.2016.1170644
- Boryczka, J. (2009). Whose responsibility? the politics of sex education policy in the United States. *Politics & Gender*, 5(2), 185-210.
doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1017/S1743923X09000154>
- Brubaker, R., & Laitin, D. (1998). Ethnic and nationalist violence. *Annual Review of Sociology*, 24(1), 423–452. <https://doi.org/10.1146/annurev.soc.24.1.423>
- Brunila, K., & Kallioniemi, A. (2018). Equality work in teacher education in Finland. *Policy Futures in Education*, 16(5), 539–552. <https://doi.org/10.1177/1478210317725674>
- Brunila, K., & Rossi, L. (2018). Identity politics, the ethos of vulnerability, and education. *Educational Philosophy and Theory*, 50(3), 287-298. doi:10.1080/00131857.2017.1343115
- Butler, J. (1999). *Gender trouble: Feminism and the subversion of identity*. New York: Routledge.
- Butler, J. (2020). *The force of nonviolence: An ethico-political bind*. New York: Verso.
- Byram, M. (2008), *From foreign language education to education for intercultural citizenship: essays and reflections*. Bristol, Blue Ridge Summit: Multilingual Matters.
<https://doi.org/10.21832/9781847690807>
- Cahill, C., Alberto Quijada Cerecer, D., Reyna Rivarola, A.,R., Hernández Zamudio, J., & Alvarez Gutiérrez, L. (2019). 'Caution, we have power': Resisting the 'school-to-sweatshop pipeline' through participatory artistic praxes and critical care. *Gender and Education*, 31(5), 576-589.
doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1080/09540253.2019.1582207>
- Carlson, M., & Kanci, T. (2017). The nationalised and gendered citizen in a global world: Examples from textbooks, policy and steering documents in Turkey and Sweden. *Gender and Education*, 29(3), 313–331. <https://doi.org/10.1080/09540253.2016.1143917>
- Carrara, S., Nascimento, M., Duque, A., & Tramontano, L. (2016). Diversity in school: A Brazilian educational policy against homophobia. *Journal of LGBT Youth*, 13(1-2), 161-172. doi:10.1080/19361653.2016.1160269
- Carvalho, L. M. (2012). The Fabrications and Travels of a Knowledge-Policy Instrument. *European Educational Research Journal*, 11(2), 172–188.
<https://doi.org/10.2304/eej.2012.11.2.172>

- Cascio, W. F., & Boudreau, J. W. (2016). The search for global competence: From international HR to talent management. *Journal of World Business*, 51(1), 103-114.
doi:<https://doi.org/10.1016/j.jwb.2015.10.002>
- Cheema, J. (2014). A Review of Missing Data Handling Methods in Education Research. *Review of Educational Research*, 84(4), 487-508. Retrieved March 3, 2020, from www.jstor.org/stable/24434247
- Chan-Tiberghien, J. (2004). Towards a 'global educational justice' research paradigm: cognitive justice, decolonizing methodologies and critical pedagogy. *Globalisation, Societies and Education*, 2(2), 191-213. DOI: 10.1080/14767720410001733647
- Choules, K. (2006). Globally privileged citizenship. *Race Ethnicity and Education*, 9(3), 275-293. DOI: 10.1080/13613320600807634
- Clegg, S. (2006). The problem of agency in feminism: a critical realist approach. *Gender and Education*, 18(3), 309-324. <https://doi.org/10.1080/09540250600667892>
- Collins, P. H. (2015). Intersectionality's Definitional Dilemmas. *Annual Review of Sociology*, 41(1), 1-20. <https://doi.org/10.1146/annurev-soc-073014-112142>
- Cornu, C. (2016). Preventing and addressing homophobic and transphobic bullying in education: A human rights-based approach using the United Nations convention on the rights of the child. *Journal of LGBT Youth*, 13(1-2), 6-17.
doi:10.1080/19361653.2015.1087932
- Costa, A. and B. Kallick (2013), *Dispositions: Reframing teaching and learning*, Corwin Press, Thousand Oaks, CA.
- Council of Europe (2016), *Competences for democratic culture: Living together as equals in culturally diverse democratic societies*, Council of Europe Publishing, Strasbourg, <http://book.coe.int>.
- Cox, C., & Meckes, L. (2016). International large-scale assessment studies and educational policy-making in Chile: Contexts and dimensions of influence. *Research Papers in Education*, 31(5), 502-515. doi:10.1080/02671522.2016.1225349
- Creswell, J. W., & Creswell, J. W. (2007). *Qualitative inquiry & research design: choosing among five approaches* (2nd ed.). Sage Publications.
- Crossouard, B., Dunne, M., & Durrani, N. (2019). Silencing youth sexuality in Senegal: Intersections of medicine and morality. *Gender and Education*, 31(2), 153-170.
doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1080/09540253.2017.1296115>
- Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education*, 10(3), 241-266. <https://doi.org/10.1177/1028315306287002>

- David, M. E. (2015). Gender & Education Association: a case study in feminist education? *Gender & Education*, 27(7), 928–946. <https://doi-org.turing.library.northwestern.edu/10.1080/09540253.2015.1096923>
- Dean, M. (2009). *Governmentality: Power and rule in modern society*. Sage Publications, Inc.
- Dennis, J. P., & Vander Wal, J. S. (2009). The Cognitive Flexibility Inventory: Instrument Development and Estimates of Reliability and Validity. *Cognitive Therapy and Research*, 34(3), 241–253. <https://doi.org/10.1007/s10608-009-9276-4>
- Denzin, N. K., & Lincoln, Y. S. (2018). *The SAGE handbook of qualitative research (5th ed.)*. Thousand Oaks: Sage Publications.
- Doscher, S. (2012), *The development of rubrics to measure undergraduate students' global awareness and global perspective: A validity study*, Florida International University, <http://dx.doi.org/10.25148/etd.FI12050125>.
- Dunne, C. (2012) How a discourse-based approach changed a project, in *Applying theory to educational research: An introductory approach with case studies* (eds Adams, J., Cochrane, M., and Dunne, L.), John Wiley and Sons Ltd., West Sussex, (pp. 123-137)
- Dussel, I. (2010). Foucault and education. In M. W. Apple, S. J. Ball, & L. A. Gandin (Eds). *The Routledge international handbook of the sociology of education* (pp. 27-36). Routledge.
- Engel, L. C., Rutkowski, D., & Thompson, G. (2019). Toward an international measure of global competence? A critical look at the PISA 2018 framework. *Globalisation, Societies and Education*, 17(2), 117-131. doi:10.1080/14767724.2019.1642183
- Ennis, R., Millman, J. and Tomko, N. (1985). *Cornell Critical Thinking Tests*. Midwest Publications.
- Ennis, R. and Weir, E. (1985), *The Ennis–Weir Critical Thinking Essay Test*, Midwest Publications.
- Ertl, H. (2006). Educational standards and the changing discourse on education: The reception and consequences of the PISA study in Germany. *Oxford Review of Education*, 32(5), 619-634. doi:10.1080/03054980600976320
- Erel, U. (2010). Feminist citizenship: Activating politics and theory. *Women's Studies Quarterly*, 38(1), 289-294. doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1353/wsqr.0.0228>
- Eteläpelto, A., Vähäsantanen, K., Hänn, P. & Paloniemi, S. (2013). What is agency? Conceptualizing professional agency at work. *Educational Research Review*, 10, 45-65.
- FBI. (2016, May 3). Hate crimes. FBI. Retrieved March 2, 2021, from <https://www.fbi.gov/investigate/civil-rights/hate-crimes>.
- Fairclough, N. (2001). Critical discourse analysis as a method in social scientific research. In *Methods of Critical Discourse Analysis* (pp. 121–138). London: SAGE Publications.

- Fairclough, N. (2003). Introduction. In *Analysing discourse: Textual analysis for social research* (pp. 1–18). Routledge.
- Fairclough, N. (2014). Introduction: critical language study. In *Language and Power* (3rd ed., pp. 1–16). Routledge.
- Fejes, A. (2009). Fabricating the European Citizen. In Peters, M., Besley, A.C., Olssen, M., Maurer, S., and Weber, S. (Eds.). *Governmentality Studies in Education*. (pp. 515-526). Rotterdam, The Netherlands: Sense Publishers.
- Fennell, S., & Arnot, M. (2008). Decentring hegemonic gender theory: the implications for educational research. *Compare: A Journal of Comparative Education*, 38(5), 525–538.
<https://doi-org.turing.library.northwestern.edu/10.1080/03057920802351283>
- Foucault, M. (1965). *Madness and Civilization: A History of Insanity in the Age of Reason*. London: Tavistock.
- Foucault, M. (1970). *The Order of Things: An Archeology of the Human Sciences*. New York: Pantheon Press.
- Foucault, M. (1977). *Discipline and Punish: The Birth of the Prison*. London: Penguin.
- Foucault, M. (1978). *The History of Sexuality: An Introduction, Volume I*. New York: Vintage House.
- Foucault, M. (1981a). The order of discourse. In R. Young (Ed.) (1981), *Untying the Text: A Post-Structural Anthology*. Boston: Routledge and Kegan Paul, pp. 48–78.
- Foucault, M. (1981b). Questions of method. *Ideology and Consciousness*, 8: 3–14.
- Foucault, M. (1985). *History of Sexuality: The Use of Pleasure, Volume 2*. Harmondsworth: Penguin.
- Foucault, M. (1988a). Technologies of the self. In L.H. Martin, H. Gutman and P.H. Hutton (Eds), *Technologies of the Self: A Seminar with Michel Foucault*. Amherst: University of Massachusetts Press, pp. 16–49.
- Foucault, M. (1988b). The political technology of individuals. In L.H. Martin, H. Gutman and P.H. Hutton (Eds), *Technologies of the Self: A Seminar with Michel Foucault*. (pp. 145–162). University of Massachusetts Press.
- Foucault, M. (1988c). *The Care of the Self: The History of Sexuality, Volume 3*. Vintage Books.
- Foucault, M. (1994). The birth of biopolitics. In P. Rabinow (Ed.), *Michel Foucault Ethics, Subjectivity and Truth: The Essential Works of Michel Foucault 1954–1984 Volume I*. New York: The New York Press.
- Foucault, M. (2007) *Security, Territory, Population: Lectures at the Collège de France, 1977-1978*, trans. G. Burchell. Palgrave Macmillan.
- Foulds, K. (2014). Buzzwords at play: gender, education, and political participation in Kenya. *Gender and Education*, 26(6), 653–671.
<https://doi.org/10.1080/09540253.2014.933190>

- Flyvbjerg, B. (2001). *Making social science matter*. Cambridge, UK: Cambridge University Press.
- Francis, D. A. (2021). 'Oh my word; for us African gays it's another story.' revealing the intersections between race, same sex-sexuality and schooling in South Africa, *Race Ethnicity and Education*, 24(1), 1-17. DOI: 10.1080/13613324.2019.1679752
- Freire, P. (2018). *Pedagogy of the oppressed* (50th anniversary ed.). New York: Bloomsbury Academic.
- Gessen, M., Shteir, R., & Mishra, P. (2020, February 09). Judith Butler Wants Us to Reshape Our Rage. Retrieved from <https://www.newyorker.com/culture/the-new-yorker-interview/judith-butler-wants-us-to-reshape-our-rage>
- Fiani, C. N., & Han, H. J. (2018). Navigating identity: Experiences of binary and non-binary transgender and gender non-conforming (TGNC) adults. *The International Journal of Transgenderism*, 20(2-3), 181–194. <https://doi.org/10.1080/15532739.2018.1426074>
- Galman, S., & Mallozzi, C. (2012). She's not there: Women and gender as disappearing foci in U.S. Research on the elementary school teacher, 1995—present. *Review of Educational Research*, 82(3), 243-295. Retrieved March 5, 2020, from www.jstor.org/stable/23260045
- Garratt, D., & Piper, H. (2010). Heterotopian cosmopolitan citizenship education? *Education, Citizenship and Social Justice*, 5(1), 43–55. <https://doi.org/10.1177/1746197909353568>
- Gee, J. P. (2011). *How to do discourse analysis: A toolkit*. New York, NY: Routledge.
- Glas, S., Spierings, N., & Scheepers, P. (2018). Re-Understanding religion and support for gender equality in Arab countries. *Gender & Society*, 32(5), 686–712. <https://doi.org/10.1177/0891243218783670>
- Godfrey, E.B., Grayman, J.K. (2014). Teaching citizens: The role of open classroom climate in fostering critical consciousness among youth. *Journal of Youth Adolescence* 43(1), 1801–1817. <https://doi-org.ezproxy.jyu.fi/10.1007/s10964-013-0084-5>
- Gordon, T., Holland, J., Lahelma, E., & Thomson, R. (2005). Imagining gendered adulthood: Anxiety, ambivalence, avoidance and anticipation. *European Journal of Women's Studies*, 12(1), 83–103. <https://doi.org/10.1177/1350506805048857>
- Gordon, T. (2006). Girls in education: Citizenship, agency and emotions. *Gender and Education*, 18(1), 1–15. <https://doi.org/10.1080/09540250500194880>
- Gordon, T., Holland, J., Lahelma, E., & Thomson, R. (2008). Young female citizens in education: Emotions, resources and agency. *Pedagogy, Culture & Society*, 16(2), 177-191. doi:10.1080/14681360802142146
- Gordon, T., Katri K. & Kirsti L. (toim.): Suomineitonen hei! Kansallisuuden sukupuoli. Tampere: Vastapaino, 197-210.

- Gorur, R. (2011). ANT on the PISA trail: Following the statistical pursuit of certainty. *Educational Philosophy and Theory*, 43, 76-93. doi:10.1111/j.1469-5812.2009.00612.x
- Graham, L. (2005) Discourse analysis and the critical use of Foucault. Paper presented at Australian Association for Research in Education, Annual Conference, Sydney, 27 November–1 December 2005. Available online at <http://www.eprints.qut.edu.au/archive> (accessed 29 November 2007).
- Grant, C., & Sleeter, C. (1986). Race, class, and gender in education research: An argument for integrative analysis. *Review of Educational Research*, 56(2), 195-211. Retrieved March 5, 2020, from www.jstor.org/stable/1170375
- Grey, S., & Morris, P. (2018). PISA: Multiple ‘truths’ and mediatised global governance. *Comparative Education*, 54(2), 109-131. doi:10.1080/03050068.2018.1425243
- Guillard, J. (2016). Is feminism trending? pedagogical approaches to countering (sl)activism. *Gender and Education*, 28(5), 609-626. doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1080/09540253.2015.1123227>
- Gunnarsson, K. (2021). In the middle of things: encountering questions about equality in social studies education. *Gender and Education*, 33(1), 33–49. <https://doi.org/10.1080/09540253.2019.1583321>
- Haavio-Mannila, E. (1979). How women become political actors: Female candidates in Finnish elections. *Scandinavian Political Studies*, 2, 351-372. doi:10.1111/j.1467-9477.1979.tb00228.x
- Halberstam, J. (2016). Transgender in a global frame. In S. Horlacher (Eds.). *Transgender and intersex: Theoretical, practical, and artistic perspectives* (pp. 165-187). Palgrave MacMillan.
- Hansen, R. (2010), *Impact of Study Abroad on Ethnocultural Empathy and Global-Mindedness*, Unpublished Doctoral Dissertation, Ball State University.
- Harden, A. D. (2009). Neoliberalism, knowledge, capitalism and the steered university: The role of OECD and Canadian federal government discourse. In Peters, M., Besley, A.C., Olssen, M., Maurer, S., and Weber, S. (Eds.) *Governmentality Studies in Education* (pp. 515-526). Sense Publishers.
- Heyes, C. (2013). Feminist solidarity after queer theory: The case of transgender. In Stryker, S., & Aizura, A. (Eds.) *The Transgender Studies Reader 2* (pp. 201-212). Routledge
- Hill Collins, P. (2019). *Intersectionality as critical social theory*. Duke University Press.
- Holli, A., & Wass, H. (2010). Gender-based voting in the parliamentary elections of 2007 in Finland. *European Journal of Political Research*, 49(5), 598-630. <https://doi.org/10.1111/j.1475-6765.2009.01910.x>
- Hooghe, M., & Stolle, D. (2004). Good girls go to the polling booth, bad boys go everywhere. *Women & Politics*, 26(3-4), 1-23. doi:10.1300/J014v26n03_01

- Hook, D. (2007). *Foucault, Psychology and the Analytics of Power (Critical Theory and Practice in Psychology and the Human Sciences)* (Revised ed.). Palgrave Macmillan.
- Howard, A., Dickert, P., Owusu, G. & Riley, D. (2018). In service of the Western world: Global citizenship education within a Ghanaian elite context. *British Journal of Educational Studies*, 66(4), 497-514. DOI: 10.1080/00071005.2018.1533100
- Huber, J. and C. Reynolds (eds.) (2014), *Developing intercultural competence through education*, Council of Europe Publishing <http://book.coe.int>.
- Huffer, L. (2010). *Mad for Foucault: Rethinking the foundations of queer theory*. Columbia University Press.
- Hyslop-Margison, E. J., Hamalian, A. & Anderson, G. (2006) A critical examination of empirical research: The case of citizenship education. *Theory & Research in Social Education*, 34(3), 395-407. DOI: 10.1080/00933104.2006.10473314
- IEA (International Association for the Evaluation of Educational Achievement) (2020). About. Accessed on June 2nd, at: <https://www.iea.nl/studies/iea/iccs>
- Illes, J. (2012). Young sexual citizens: reimagining sex education as an essential form of civic engagement. *Sex Education*, 12(5), 613-625. DOI: 10.1080/14681811.2011.634152
- IPU (2020). Monthly Ranking of Women in national parliaments. Parline: the International Parliamentary Union's (IPU) Open Data Platform. Retrieved February 2, 2021, from <https://data.ipu.org/women-ranking?month=10&year=2020>.
- Sulkunen, I. (2007). Suffrage, gender and citizenship in Finland: A comparative perspective. *NORDEUROPAforum*, 17(1), 27–44. doi: 10.18452/7930
- Jansen, S. (2016). Transgender and intersex: Unavoidable essentialism and the normative struggle for recognition. In S. Horlacher (Eds.). *Transgender and intersex: Theoretical, practical, and artistic perspectives* (pp. 115-140). Palgrave MacMillan.
- JMP (2020). *Chi-square goodness of fit test*. Retrieved February 30, 2021, from https://www.jmp.com/en_ch/statistics-knowledge-portal/chi-square-test/chi-square-goodness-of-fit-test.html.
- Joris, M., & Agirdag, O. (2019). In search of good citizenship education: A normative analysis of the international civic and citizenship education study (ICCS). *European Journal of Education*, 54(2), 287-298. doi:10.1111/ejed.12331
- Jupp, J., Berry, T., & Lensmire, T. (2016). Second-wave white teacher identity studies: A review of white teacher identity literatures from 2004 through 2014. *Review of Educational Research*, 86(4), 1151-1191. Retrieved March 2, 2020, from www.jstor.org/stable/44668246
- Kanes, C., Morgan, C., & Tsatsaroni, A. (2014). The PISA mathematics regime: Knowledge structures and practices of the self. *Educational Studies in Mathematics*, 87(2), 145-165. doi:10.1007/s10649-014-9542-6

- Kendi, I. X. (2016). *Stamped from the beginning: the definitive history of racist ideas in America*. Nation Books.
- Kessl, F. (2009). What's the use of studies on governmentality in social work? In Peters, M., Besley, A.C., Olssen, M., Maurer, S., and Weber, S. (Eds.). *Governmentality Studies in Education*. (pp. 415-430). Sense Publishers.
- Kiwan, D. (2008). Citizenship education in England at the cross-roads? Four models of citizenship and their implications for ethnic and religious diversity. *Oxford Review of Education*, 34(1), 39–58. <https://doi.org/10.1080/03054980701584551>
- Kontula, O. (2010). The evolution of sex education and students' sexual knowledge in Finland in the 2000s. *Sex Education*, 10, 373-386. doi:10.1080/14681811.2010.515095
- Lahelma, E. (2002). Isänmaa, armeija ja tasa-arvo. Teoksessa
- Lahelma, E. (2011). Gender awareness in Finnish teacher education: An impossible mission? *Education Inquiry*, 2(2), 263-276. doi:10.3402/edui.v2i2.21979
- Lahelma, E. (2012) Female paths to adulthood in a country of 'genderless gender', *Gender and Education*, 24:1, 1-13, doi: 10.1080/09540253.2010.549111
- Lahelma, E., & Tainio, L. (2019). The long mission towards gender equality in teacher education:: Reflections from a national project in Finland. *Nordic Studies in Education*, 39, 69-84. doi:10.18261/issn.1891-2019-01-06
- Lahelma, E., Örn, E., & Weiner, G. (2020). Reflections on the emergence and development of research on gender and education in the Nordic countries. Unpublished. Retrieved through email on February 4, 2020.
- Lappalainen, S., & Lahelma, E. (2016). Subtle discourses on equality in the Finnish curricula of upper secondary education: Reflections of the imagined society. *Journal of Curriculum Studies*, 48(5), 650-670. doi:10.1080/00220272.2015.1069399
- Lavery, G. (2019). The King's Two Anuses: Trans Feminism and Free Speech. *Differences (Bloomington, Ind.)*, 30(3), 118–151. <https://doi.org/10.1215/10407391-7974030>
- Lazar, M. (2005). Politicizing gender in discourse: Feminist critical discourse analysis as political perspective and praxis. In *Feminist Critical Discourse Analysis: Gender, Power and Ideology in Discourse* (2005th ed., pp. 1–30). Palgrave Macmillan.
- Ledger, S., Thier, M., Bailey, L., and Pitts, C., (2019). OECD's Approach to measuring global competency: Powerful voices shaping education. *Teachers College Record*, 121(8). 1-40. <https://www.tcrecord.org> ID Number: 22705, Date Accessed: 4/2/2020 3:54:13 AM
- Lee, W. O. (2003). Students' concepts and attitudes toward citizenship: the case of Hong Kong. *International Journal of Educational Research*, 39(6), 591-607. <https://doi.org/10.1016/j.ijer.2004.07.007>
- Liesner, A. (2012). Peer pressure: Comments on the European educational reform. *Policy Futures in Education*, 10(3), 297–301. <https://doi.org/10.2304/pfie.2012.10.3.297>

- Liu, O., L. Frankel and K. Roohr (2014), *Assessing critical thinking in higher education: Current state and directions for next-generation assessment*, Educational Testing Service, <http://dx.doi.org/10.1002/ets2.12009>.
- Liu, O., Mao, L., Frankel, L., Xu, L., (2016), Assessing critical thinking in higher education: The HEIghten™ approach and preliminary validity evidence, *Assessment and Evaluation in Higher Education*, 41(5), 677-694.
<http://dx.doi.org/10.1080/02602938.2016.1168358>.
- Love, B. L. (2017). A ratchet lens: Black queer youth, agency, hip hop, and the Black ratchet imagination. *Educational Researcher*, 46(9), 539–547.
<https://doi.org/10.3102/0013189X17736520>
- Maber, E. J. T. (2016). Finding feminism, finding voice? Mobilising community education to build women's participation in Myanmar's political transition. *Gender and Education*, 28(3), 416-430. doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1080/09540253.2016.1167175>
- Manohar, N. N. (2019). Gendered agency in skilled migration: The case of Indian women in the United States. *Gender & Society*, 33(6), 935–960.
<https://doi.org/10.1177/0891243219865544>
- Martin, M. and R. Rubin (1995). A new measure of cognitive flexibility, *Psychological Reports*, 76(2), 623-626, <https://doi.org/10.2466/pr0.1995.76.2.623>
- Mayo, J. (2013). Critical pedagogy enacted in the Gay—Straight Alliance: New possibilities for a third space in teacher development. *Educational Researcher*, 42(5), 266-275. Retrieved December 26, 2020, from <http://www.jstor.org/stable/23462392>
- McGaw, B. (2008). The role of the OECD in international comparative studies of achievement. *Assessment in Education: Principles, Policy & Practice*, 15(3), 223-243.
doi:10.1080/09695940802417384
- McNay, L. (2004). Agency and experience: gender as a lived relation. *The Sociological Review* (Keele), 52(s2), 173–190. <https://doi.org/10.1111/j.1467-954X.2005.00530.x>
- McNeill, T. (2013). Sex education and the promotion of heteronormativity. *Sexualities*, 16(7), 826–846. <https://doi.org/10.1177/1363460713497216>
- Meyer, H. (2014). The OECD as pivot of the emerging global educational accountability regime: How accountable are the accountants? *Teachers College Record*, 116 (9), 1-20.
<https://www.tcrecord.org> ID Number: 17543, Date Accessed: 4/2/2020 5:39:39 AM
- Meyer, T. G. (2020). Difference as privilege: Identity, citizenship and the recontextualisation of human rights in Japan's social studies curriculum. *Critical Studies in Education*, 61(1), 17-36. doi:10.1080/17508487.2017.1352007
- Moeller, K. (2021). The politics of curricular erasure: debates on race, gender, and sexuality in the Brazilian 'common core' curriculum. *Race Ethnicity and Education*, 24(1), 18-38.
DOI: 10.1080/13613324.2020.1798382

- Morgan, C., & Shahjahan, R. A. (2014). The legitimation of OECD's global educational governance: Examining PISA and AHELO test production. *Comparative Education*, 50(2), 192-205. doi:10.1080/03050068.2013.834559
- Namaste, V. (2009). Undoing Theory: The "Transgender Question" and the epistemic violence of Anglo-American feminist theory. *Hypatia*, 24(3), 11-32. Retrieved December 6, 2020, from <http://www.jstor.org/stable/20618162>
- Neumann, K., Fischer, H. E., & Kauertz, A. (2010). From PISA to educational standards: The impact of large-scale assessments on science education in Germany. *International Journal of Science and Mathematics Education*, 8(3), 545-563. doi:10.1007/s10763-010-9206-7
- Nussbaum, M. C. (1997). *Cultivating humanity: a classical defense of reform in liberal education*. Harvard University Press.
- Nussbaum, M. C. (2000). *Women and human development: The capabilities approach*. Cambridge University Press.
- Nussbaum, M. C. (2011). *Creating capabilities: The human development approach*. Belknap Press of Harvard University Press.
- Nwosu, P. (2009), Understanding Africans' conceptualizations of intercultural competence, in Deardorff, D. (ed.), *The SAGE handbook of intercultural competence*, Sage, Thousand Oaks, CA.
- Oberle, M. (2012). Pupils' political knowledge regarding the European Union. *Citizenship, Social and Economics Education*, 11(2), 89–104.
<https://doi.org/10.2304/csee.2012.11.2.89>
- O'Donnell, J. L. (2018). The promise of recognition and the repercussions of government intervention: the transpedagogical vision of popular educators in Buenos Aires, Argentina. *Gender and Education*, 30(8), 1078–1097.
<https://doi.org/10.1080/09540253.2017.1296113>
- OECD Development Cooperation Directorate (2016). *Better Policies for 2030: An OECD Action Plan on the Sustainable Development Goals*. Retrieved from:
<https://www.oecd.org/dac/Better%20Policies%20for%202030.pdf>
- OECD (2019a), PISA 2018 Assessment and Analytical Framework, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/b25efab8-en>. Retrieved from: https://www.oecd-ibrary.org/education/pisa-2018-assessment-and-analytical-framework_b25efab8-en
- OECD (2019b), PISA 2018 Assessment and Analytical Framework, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/b25efab8-en>. Retrieved from: <https://www.oecd.org/pisa/innovation/global-competence/>
- OECD (2020), *PISA 2018 Results (Volume VI): Are students ready to thrive in an interconnected world?*, PISA, OECD Publishing, <https://doi.org/10.1787/d5f68679-en>.

- Okello, W. K., & Turnquest, T. A. (2020). 'Standing in the kitchen': Race, gender, history and the promise of performativity, *International Journal of Qualitative Studies in Education*, DOI: 10.1080/09518398.2020.1828653
- Okkolin, M. (2016). Who decides? Tanzanian women's narratives on educational advancement and agency. *Gender and Education*, 28(7), 886-902.
doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1080/09540253.2015.1124069>
- Olssen, M., Codd, J. A., & O'Neill, A. (2004). *Education policy: Globalization, citizenship and democracy* (First ed.). SAGE Publications Ltd.
- Oxfam (2015), *Education for global citizenship: A guide for schools*.
- Pulkkinen, T. (2000). *The postmodern and political agency*. SoPhi.
- Phoenix, A. (2009). De-colonising practices: negotiating narratives from racialised and gendered experiences of education, *Race Ethnicity and Education*, 12(1), 101-114. DOI: 10.1080/13613320802651053
- Passani, A., & Debicki, M. (2016). Students opinions and attitudes toward LGBT persons and rights: Results of a transnational European project. *Journal of LGBT Youth*, 13(1-2), 67-88. doi:10.1080/19361653.2015.1087927
- Pearce, R. (2020). A methodology for the marginalised: Surviving oppression and traumatic fieldwork in the neoliberal academy. *Sociology*, 54(4), 806–824. <https://doi.org/10.1177/0038038520904918>
- Perez, C. C. (2019). *Invisible women: Data bias in a world designed for men*. Abrams Press.
- Peters, M. (2007). Educational Research: 'Games of Truth' and the Ethics of Subjectivity. *Counterpoints*, 292, 181-191. Retrieved October 6, 2020, from <http://www.jstor.org/stable/42979099>
- Peters, M. A., Besley, A. C., Olssen, M., Maurer, S., & Weber, S. (Eds.). (2009). *Governmentality Studies in Education*. Sense Publishing.
- Pitoňák, M., & Spilková, J. (2016). Homophobic prejudice in Czech youth: A sociodemographic analysis of young people's opinions on homosexuality. *Sexuality Research and Social Policy*, 13(3), 215-229. doi:10.1007/s13178-015-0215-8
- Pons, X. (2012). Going beyond the 'PISA Shock' discourse: An analysis of the cognitive reception of PISA in six European countries, 2001–2008. *European Educational Research Journal*, 11(2), 206–226. <https://doi.org/10.2304/eeerj.2012.11.2.206>
- Pons, X. (2017). Fifteen years of research on PISA effects on education governance: A critical review. *European Journal of Education*, 52(2), 131-144. doi:10.1111/ejed.12213
- Popkewitz, T. S., & Brennan, M. T. (1997). *Foucault's challenge: Discourse, knowledge, and power in education* (First Edition). Teachers College Press.
- Raaum, N. C. (2005). Gender equality and political representation: A Nordic comparison. *West European Politics*, 28(4), 872-897. doi:10.1080/01402380500217003

- Rao, N. (2015). Marriage, violence, and choice: Understanding Dalit women's agency in rural Tamil Nadu. *Gender & Society*, 29(3), 410–433.
<https://doi.org/10.1177/0891243214554798>
- Reimers, F. (2017), *Empowering students to improve the world in sixty lessons. Version 1.0*, CreateSpace Independent Publishing Platform.
- Russell, S. G., Lerch, J. C., & Wotipka, C. M. (2018). The making of a human rights issue: A cross-national analysis of gender-based violence in textbooks, 1950-2011. *Gender & Society*, 32(5), 713–738. <https://doi.org/10.1177/0891243218786686>
- Rutkowski, D. (2015). The OECD and the local: PISA-based test for schools in the USA. *Discourse: Studies in the Cultural Politics of Education*, 36(5), 683-699.
doi:10.1080/01596306.2014.94315
- Sandlin, J., O'Malley, M., & Burdick, J. (2011). Mapping the Complexity of Public Pedagogy Scholarship: 1894–2010. *Review of Educational Research*, 81(3), 338-375. Retrieved March 4, 2020, from www.jstor.org/stable/23014296
- Sabzalian, L. (2019). The tensions between Indigenous sovereignty and multicultural citizenship education: Toward an anticolonial approach to civic education. *Theory & Research in Social Education*, 47(3), 311-346. DOI: 10.1080/00933104.2019.1639572
- Sant, E. (2019). Democratic education: A theoretical review (2006–2017). *Review of Educational Research*, 89(5), 655–696. <https://doi.org/10.3102/0034654319862493>
- Schuelka, M. J. (2013). Excluding students with disabilities from the culture of achievement: The case of the TIMSS, PIRLS, and PISA. *Journal of Education Policy*, 28(2), 216-230. doi:10.1080/02680939.2012.708789
- Seider, S., Tamerat, J., Clark, S. et al. (2017). Investigating adolescents' critical consciousness development through a character framework. *Journal of Youth Adolescence* 46(2), 1162–1178. <https://doi-org.ezproxy.jyu.fi/10.1007/s10964-017-0641-4>
- Sellar, S., & Lingard, B. (2013). The OECD and global governance in education. *Journal of Education Policy*, 28(5), 710-725. doi:10.1080/02680939.2013.779791
- Shoshana, A., & Shchada, S. (2018). Agentive suspension of oppressive structures: Self-concept and emotion-work among Muslim mothers who study at a Jewish university. *Gender and Education*, 30(7), 934-950.
doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1080/09540253.2017.1416069>
- Simpson, A., & Dervin, F. (2019). Global and intercultural competences for whom? by whom? for what purpose? : An example from the Asia society and the OECD. *Compare: A Journal of Comparative and International Education*, 49(4), 672-677.
doi:10.1080/03057925.2019.1586194

- Springer, E. (2020). Bureaucratic tools in (gendered) organizations: Performance metrics and gender advisors in international development. *Gender & Society*, 34(1), 56–80. <https://doi.org/10.1177/0891243219874058>
- Sung, Y., & Kang, M. O. (2012). The cultural politics of national testing and test result release policy in South Korea: A critical discourse analysis. *Asia Pacific Journal of Education*, 32(1), 53-73. doi:10.1080/02188791.2011.621691
- Snorton, C. R. (2009). "A New Hope": The Psychic Life of Passing. *Hypatia*, 24(3), 77-92. Retrieved March 13, 2021, from <http://www.jstor.org/stable/20618165>
- Snorton, C. R. (2017). *Black on both sides: A racial history of trans identity*. Minneapolis: University of Minnesota Press.
- Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1994). Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains. In R. B. Ruddell, M. R. Ruddell, & H. Singer (Eds.), *Theoretical models and processes of reading* (p. 602–615). *International Reading Association*. (Reprinted from the "Tenth Annual Conference of the Cognitive Science Society Proceedings.")
- Stone, S. (2006). The empire strikes back: A posttranssexual manifesto. In S. Stryker & S. Whittle (Eds.) *The transgender studies reader* (pp. 222-235). Routledge.
- Stromquist, N. P. (2005). Comparative and international education: A journey toward equality and equity. *Harvard Educational Review*, 75(1), 89–111. <https://doi.org/10.17763/haer.75.1.4842626v0r385j12>
- Stryker, S. (2008). *Transgender History*. Seal Press.
- Stucky, J. L., Dantas, B. M., Pocahy, F. A., de Queiroz Nogueira, P. H. Nardi, H. C., & Costa, A. B. (2020). Prejudice against gender and sexual diversity in Brazilian public high schools. *Sexuality Research and Social Policy*, 17(3), 429-441. doi:10.1007/s13178-019-00406-z
- Tan, C. (2017). Chinese responses to Shanghai's performance in PISA. *Comparative Education*, 53(2), 209-223. doi:10.1080/03050068.2017.1299845
- Taylor, S. (2001) Locating and conducting discourse analytic research, in *Discourse as Data: A Guide for Analysis* (eds M. Wetherell, S. Taylor and S. Yates), Sage, London, 142–155.
- The London Feminist Salon Collective. (2004). The problematization of agency in postmodern theory: as feminist educational researchers, where do we go from here? *Gender and Education*, 16(1), 25–33. <https://doi.org/10.1080/0954025032000170318>
- To the members of the Pontifical Academy of Social Sciences on the occasion of their Plenary Assembly [28 April - 2 May 2017] (24 April 2017): Francis. (2017, April 24). Re-

trieved from http://www.vatican.va/content/francesco/en/messages/pont-messages/2017/documents/papa-francesco_20170424_messaggio-accademia-scienzesociali.html

- Tolonen, T. (2019). Minority young men's gendered tactics for making space in the city and at school. *Gender and Education*, 31(3), 408-424.
doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1080/09540253.2018.1496228>
- Tormey, R., & Gleeson, J. (2012). The gendering of global citizenship: findings from a large-scale quantitative study on global citizenship education experiences. *Gender and Education*, 24(6), 627–645. <https://doi.org/10.1080/09540253.2011.646960>
- Tracy, S. J. (2020). *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact*. Second Edition. Wiley-Blackwell.
- Tsouroufli, M. (2018). Gendered pedagogic identities and academic professionalism in Greek medical schools. *Gender and Education*, 30(1), 45-58.
doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1080/09540253.2016.1262008>
- Unterhalter, E. (2017). Negative capability? Measuring the unmeasurable in education. *Comparative Education*, 53(1), 1-16. DOI: 10.1080/03050068.2017.1254945
- UNESCO (2015), Global Citizenship Education: Topics and Learning Objectives, UNESCO.
- Usher, E., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research*, 78(4), 751-796. Retrieved March 5, 2020, from www.jstor.org/stable/40071145
- Vaccari, V. and Gardinier, M.P. (2019) Toward one world or many? A comparative analysis of OECD and UNESCO global education policy documents. *International Journal of Development Education and Global Learning*, 11 (1): 68–86. DOI <https://doi.org/10.18546/IJDEGL.11.1.05>
- Väisänen, M., & Roderick, F. (n.d.). Åström, Emma Irene (1847 - 1934). Retrieved November 24, 2019, from <https://kansallisbiografia.fi/english/person/3699>.
- Valtonen, H., & Rautiainen, M. (2012). La educación Finlandesa desde 1850 hasta el presente. *Istor*, 12 (48), 129-160.
- Van Dijk, T. A. (2001). Multidisciplinarity in CDA: a plea for diversity. In *Methods of Critical Discourse Analysis* (pp. 95–119). SAGE Publications.
- Wald, J., & Losen, D. J. (2003). Defining and redirecting a school-to-prison pipeline. *New Directions for Youth Development*, 2003(99), 9–15. <https://doi.org/10.1002/yd.51>
- Waldow, F., Takayama, K., & Sung, Y. (2014). Rethinking the pattern of external policy referencing: Media discourses over the ‘Asian tigers’ PISA success in Australia, Germany and South Korea. *Comparative Education*, 50(3), 302-321.
doi:10.1080/03050068.2013.860704

- Weber, R. (1990). Techniques of content analysis. In *Basic content analysis* (pp. 41-70). SAGE Publications, Inc., <https://www.doi.org/10.4135/9781412983488>
- Wells, K. (2017). Sexual minority teachers as activist-educators for social justice. *Journal of LGBT Youth*, 14(3), 266-295. doi:10.1080/19361653.2017.1324344
- Whittle, S. & Turner, L. (2016). Transgender in a global frame. In S. Horlacher (Eds.). *Transgender and Intersex: Theoretical, Practical, and Artistic Perspectives* (pp. 29-65). Palgrave MacMillan.
- Williams-Gualandi, D. (2015), "Intercultural Understanding: What Are We Looking for and How do we Assess What We Find?", *International and Global Issues for Research*, No. 2015/7, University of Bath, Bath, <http://bathedresearch.weborium.net>.
- Wodak, R., & Corson, P. (2012). Introduction. In *Encyclopedia of Language and Education* (Language Policy and Political Issues in Education ed., Vol. 1, pp. xi–xiv). Springer Publishing.
- Wodak, R. (2015). What CDA is about - a summary of its history, important concepts and its developments. In M. Meyer (Ed.), *Methods of Critical Discourse Studies: Vol. Introducing Qualitative Methods series* (Third ed., pp. 1–13). SAGE Publications Ltd.
- Weaver-Hightower, M. (2003). The "boy turn" in research on gender and education. *Review of Educational Research*, 73(4), 471-498. Retrieved March 5, 2020, from www.jstor.org/stable/3516000
- Webb, R. (2019). 'Being yourself': Everyday ways of doing and being gender in a 'rights-respecting' primary school. *Gender and Education*, 31(2), 258-273. doi:<http://dx.doi.org.ezproxy.jyu.fi/10.1080/09540253.2017.1309010>
- Wilkins, A. (2012). The spectre of neoliberalism: Pedagogy, gender and the construction of learner identities. *Critical Studies in Education*, 53(2), 197-210. doi:10.1080/17508487.2012.672332
- World values survey association. WVS Database. (2021). Retrieved January 6, 2021, from <https://www.worldvaluessurvey.org/WVSContents.jsp>.
- Yang, C. (2016). Encounters between the 'oppressed' and the 'oppressor': rethinking Paulo Freire in anti-racist feminist education in Sweden, *Race Ethnicity and Education*, 19(4), 835-855. DOI: 10.1080/13613324.2014.885421
- Zimman, L. (2019). Trans self-identification and the language of neoliberal selfhood: Agency, power, and the limits of monologic discourse, *International Journal of the Sociology of Language*, 2019(256), 147-175. doi: <https://doi-org.ezproxy.jyu.fi/10.1515/ijsl-2018-2016>
- Zion-Waldoks, T. (2015). Politics of devoted resistance: Agency, feminism, and religion among orthodox Agunah activists in Israel. *Gender & Society*, 29(1), 73–97. <https://doi.org/10.1177/0891243214549353>

Zoido, P., & Schleicher, A. (2016). The Policies that Shaped PISA, and the Policies that PISA Shaped. In *Handbook of Global Policy Series* (pp. 374–385). John Wiley & Sons, Ltd.

9. APPENDICES

9.1. Appendix A

Table A1

Chi Square Test Statistics

Sample Composition	Value of Statistic General Categories ¹	n	df	Value of Statistic Gender Categories	n	df
2018 PISA GCF Text	81.5880***	3	2	6.7176	5	4

* $p \leq .05$, ** $p < .001$, *** $p < .00001$; ¹ values of Pearson Chi-Square Statistic goodness-of-fit test compared to the real distribution of the Bibliography Sample

Table A2

Chi Square Test Statistics

Sample Composition	Value of Statistic General Categories ¹	n	df	Value of Statistic Gender Categories	n	df
Total Bibliography Sample	10,233.4883***	3	2	4,288.0279	5	4
2018 PISA GCF Text	412.75***			28.9231***		
Sources Grouped by Discipline	Value of Statistic General Categories	n	df	Value of Statistic Gender Categories	n	df
Political Science	426.4760***	3	2	939.1936***	5	4
Education Research	214.3612***	3	2	2,276.3113***	5	4
Intercultural Studies	16,550.3434***	3	2	1,133.4835***	5	4
Sources Grouped by Type of Source	Value of Statistic General Categories	n	df	Value of Statistic Gender Categories	n	df
IEA Studies	511.1372***	3	2	1,203.0499***	5	4
ETS Studies	114.1880***	3	2	276.9172***	5	4
Academic Books	4,189.7670***	3	2	1,677.4585***	5	4
Journal Articles	942.5542***	3	2	453.9322***	5	4
Dissertations & Unpublished Works	590.3864***	3	2	117.5128***	5	4
UNESCO	5,719.3304***	3	2	771.8415***	5	4
NGO Publications	479.5284***	3	2	19.1579**	5	4
Council of Europe Publications	1,320.1032***	3	2	72***	5	4
Sample Works from Influential Authors	Value of Statistic General Categories	n	df	Value of Statistic Gender Categories	n	df
Works by Butler	8,714.6840***	3	2	5,964.9140***	5	4
Works by Snorton	1,520.4344***	3	2	1,467.1891***	5	4
Works by Nussbaum	1,373.0768***	3	2	1,452.1896***	5	4
Works by Arnot	3,176.6558***	3	2	2,096.5666***	5	4
Works by Deardorff	4,125.0624***	3	2	119.33***	5	4

* $p \leq .05$, ** $p < .001$, *** $p < .00001$; ¹ values of Pearson Chi-Square Statistic goodness-of-fit test compared to an equal distribution

9.2. Appendix B

Table B1

General Category Count of the Bibliography Sample: Selected Sources Grouped by Discipline

	Gender		Race		Culture		Total Count
	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	
Political Science	1446	43.57%	1323	39.86%	550	16.57%	3319
Education Research	1,574	43.87%	862	24.02%	1,152	32.11%	3,588
Intercultural Studies	910	4.93%	3267	17.70%	14,276	77.36%	18,453

Table B2

Gender Category Count of the Bibliography Sample: Selected Sources Grouped by Discipline

	Gender & Sexuality		Trans		LGBQ		Female		Feminist		Total Count
	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	
Political Science	526	36.38%	0	0%	191	13.21%	600	41.49%	129	8.92%	1,446
Education Research	752	47.78%	1	0.06%	11	0.7%	804	51.08%	6	0.38%	1,574
Intercultural Studies	372	40.88%	0	0%	45	4.95%	481	52.86%	12	1.32%	910

Table B3*General Category Count of the Bibliography Sample: Selected Sources Grouped by Type of Source*

	Gender		Race		Culture		Total Count
	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	
IEA	802	69.26%	243	20.98%	113	9.76%	1,158
ETS	157	59.02%	94	35.34%	15	5.64%	266
Academic Books	2,131	16.04%	3,272	24.63%	7,884	59.34%	13,287
Journal Articles	298	11.45%	745	28.63%	1,559	59.92%	2,602
Dissertations & Un-published Work	19	5.4%	22	6.25%	311	88.35%	352
UNESCO	492	7.77%	901	14.24%	4,936	77.99%	6,329
NGO	19	5.4%	22	6.25%	311	88.55%	352
Council of Europe	30	3.60%	31	3.72%	772	92.68%	833

Table B4*Gender Category Count of the Bibliography Sample: Selected Sources Grouped by Type of Source*

	Gender & Sexuality		Trans		LGBQ		Female		Feminist		Total Count
	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	
IEA	399	49.75%	0	0%	0	0%	403	50.25%	0	0%	802
ETS	53	33.76%	0	0%	0	0%	104	66.24%	0	0%	157
Academic Books	807	37.87%	1	0.05%	235	11.03%	947	44.44%	141	6.62%	2,131
Journal Articles	169	56.71%	0	0%	3	1.01%	125	41.95%	1	0.34%	298
Dissertations & Un-published Work	41	52.56%	0	0%	0	0%	37	47.44%	0	0%	78
UNESCO	182	36.99%	0	0%	3	0.61%	302	61.38%	5	1.02%	492
NGO	10	52.63%	0	0%	6	31.48%	3	15.79%	0	0%	19
Council of Europe	24	80%	0	0%	0	0%	6	20%	0	0%	30

Table B5*General Category Count of the Influential Author Sample*

	Gender		Race		Culture		Total Count
	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	
Snorton	1724	59.72%	1126	39%	37	1.28%	2887
Butler	5573	88.9%	276	4.4%	420	6.7%	6269
Nussbaum	1646	67.6%	582	23.90%	207	8.5%	2435
Arnot	2524	80.9%	311	9.97%	285	9.13%	3120
Deardorff	84	3.3%	86	3.38%	2377	93.33%	2547

Table B6*Gender Category Count of the Influential Author Sample*

	Gender & Sexuality		Trans		LGBQ		Female		Feminist		Total Count
	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	Word Count	Relative Frequency	
Snorton	815	47.27%	612	35.5%	44	2.55%	244	14.15%	9	0.52%	1724
Butler	3379	60.53%	308	5.53%	677	12.15%	886	15.9%	323	5.8%	5573
Nussbaum	496	30.13%	8	0.49%	212	12.88%	850	51.64%	80	4.86%	1646
Arnot	805	31.89%	0	0%	38	1.51%	1208	47.86%	473	18.74%	2524
Deardorff	36	42.86%	0	0%	2	2.38%	46	54.76%	0	0%	84

