

**The Influence of Culture and Socioeconomic Status on the Mandatory  
Reporting of Female Genital Mutilation (FGM) by Health and Social  
Care Professionals and Teachers**

**Fatoumata Jatta  
June 2019**

**Research submitted in partial fulfilment of the requirements for the degree of  
Doctor in Clinical Psychology (DClinPsy), Royal Holloway, University of London**

## **Acknowledgements**

Thank you to my supervisors Dr Charles Efferson, Dr Lih-Mei Liao, and Dr Jane Vosper. It has been a pleasure to work with you. You have all been incredibly generous in sharing your expertise. I also appreciate the patience and kindness that you showed towards me. It has made writing this thesis a positive experience, and for that I am grateful.

Thank you to my family; my parents and siblings, and my closest friends. I feel very lucky to have your love and support.

Thank you to all the study participants, and everyone who gave up their time to help with piloting for the study.

To the women and girls, who like me, have been affected by FGM, while all of our stories are different, I hope that in writing this thesis, I have been able to contribute something of value to the understanding of our collective story.

## Table of Contents

ACKNOWLEDGEMENTS.....	2
TABLE OF CONTENTS.....	3
EXECUTIVE SUMMARY.....	5
Systematic Review.....	6
Empirical Study.....	13
Integration, Impact, and Dissemination.....	18
CULTURAL COMPETENCE: A SYSTEMATIC REVIEW ON THE ROLE OF CULTURAL FACTORS IN PROFESSIONALS' DECISIONS ABOUT CHILD MALTREATMENT.....	20
Abstract.....	21
Introduction.....	23
Method.....	36
Results.....	40
Discussion.....	64
Conclusion.....	71
THE INFLUENCE OF CULTURE AND SOCIOECONOMIC STATUS ON THE MANDATORY REPORTING OF FEMALE GENITAL MUTILATION (FGM) BY HEALTH AND SOCIAL CARE PROFESSIONALS AND TEACHERS.....	73
Abstract.....	74
Introduction.....	76
Method.....	91
Results.....	103
Discussion.....	115
Conclusion.....	125
INTEGRATION, IMPACT, AND DISSEMINATION SUMMARY.....	128
Integration.....	129
Impact.....	138
Dissemination.....	144
REFERENCES.....	148
APPENDICES.....	160
Appendix A.....	161

Appendix B.....	164
Appendix C.....	168
Appendix D.....	169
Appendix E.....	176
Appendix F.....	186
Appendix G.....	190
Appendix H.....	191
Appendix I.....	193
Appendix J.....	194

LIST OF TABLES .....	
Table 1.....	43
Table 2.....	62
Table 3.....	96
Table 4.....	96
Table 5.....	108
Table 6.....	111
Table 7.....	114
LIST OF FIGURES .....	
Figure 1.....	42

## **Executive Summary**

# **Cultural Competence: A Systematic Review on the Role of Cultural Factors in Professionals' Decisions about Child Maltreatment**

## **Introduction**

Maltreatment of children, incorporating various types of abuse and neglect, occurs when a caretaker or responsible person harms, threatens to harm, or fails to provide adequate care for a child. It is one of the most powerful risk factors for concurrent and subsequent psychopathology, later health morbidity, and compromised development. Internationally, it is estimated that 40 million children experience abuse annually (World Health Organization, 2014).

Definitions of child abuse and child neglect often fail to meet research needs due to their lack of comparability, reliability, and universally understood delineations. Moreover, despite numerous legislation, there remains very little guidance on how to interpret and implement statutory terms, or what levels of concern should be reported. While a broad definition of child maltreatment allows courts enough flexibility to respond to real-life situations, this stance also negates social and cultural differences, making any cross-cultural comparisons especially problematic. Cultural practices might be simultaneously normative and also harmful to the child. Female genital mutilation (FGM) is an example of a cultural practice that is normative to some groups, and also potentially harmful to affected children. Navigating this confusion poses a challenge for professionals.

The systematic review aimed to provide a comprehensive amalgamation of the existing literature on the role of culture in Western-based professionals' decisions about child maltreatment. It aimed to answer the following question: how do cultural factors influence professionals' decisions about child maltreatment?

## **Method**

The review was conducted according to guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. Online databases (PsycINFO, PubMed, and Web of Science) were searched for relevant studies. Included studies were peer-reviewed articles published in English that: 1) collected original data from participants who were health and social care professionals or teachers working in Western countries, and 2) measured and reported results on cultural factors (including race and ethnicity) that influence professionals' decision-making and/or reporting behaviour concerning child maltreatment. Data on participants, study characteristics, and key findings were extracted and summarised. Study methodology was assessed using a quality appraisal tool for cross-sectional studies.

## **Results**

Sixteen studies were included in the review; all were cross-sectional, and 13 employed vignette methods. Thirteen of the studies (76.5%) sampled participants from the US. Five studies recruited via random sampling of the mailing lists of relevant professional bodies, three of which were nationwide. Response rates for

professional participants in these studies ranged from 23.5% to 60% (mean = 43.56%). The remaining 10 studies used convenience sampling methods in a variety of settings, including workshops, a conference, and online. Response rates for professional participants in these studies ranged from 69% to 100% (mean = 92.06%).

Eleven studies sampled practicing health and social care professionals including psychologists, psychiatrists, doctors, and nurses. Three of these sampled social workers only, and two included other groups involved in child protection such as police, judges, and community and faith members. Four studies sampled teachers only. The sample sizes for professionals varied considerably, ranging from 6939 participants to 35 (median 184.50, IQR 338.25). Of the 16 studies, five sampled fewer than 100 professionals, six sampled between 100 and 300, three between 300 and 600, one had between 1000 to 2000 professionals, and the largest sampled just under 7000.

Seven studies did not report the race or ethnicity of the sample. Of the nine that did, four had a majority sample of 85% to 95% White participants; followed by small proportions of African-American and Hispanic participants. Three had majority samples of around 40% African-American, followed by substantial proportions of White (33 – 38%) and Asian (20 – 24%) participants. Only one had an equal proportion of White, Black, and Hispanic participants. Four studies did not report the gender of the sample. Of the 12 studies that did, nine had a majority of female

participants (64% – 96%), and three had about equal proportions of males and females.

The indirect vignette measurement approach was the most commonly used. Physical abuse, particularly corporal punishment, was the most commonly studied type of abuse. While no particular patterns emerged by country and study population, concerning measurement approach, studies involving teachers were the most likely to use a validated direct measure of cultural factors. The Educators and Child Abuse Questionnaire (ECAQ) was used in three of the five studies that sampled teachers, followed by the Corporal Punishment Scale (CPS) which was used in two studies by the same author.

Nine of the 16 studies were of moderate quality, five studies were of moderate to high quality, while one was of low to moderate quality, and one was of low quality. As all studies were cross-sectional, causal inference is limited. There were a number of major methodological limitations to the studies including the use of non-validated measures, small sample sizes, low response rates, non-representative samples, and social desirability effects. Twelve studies did not control for main confounders such as professionals' gender, SES, history of abuse, and age. Nine studies used non-randomised samples, while four failed to randomise the order of vignettes. Thirteen of the 16 studies were conducted in the US, limiting generalisability.

## **Discussion**

Sixteen studies were identified. While review search terms included several variants for culture, and included the terms nationality, ethnicity, and race, the majority of studies (81.25%) were concerned with race/ethnicity variables of the case, professional, or both. Only one study was concerned with country of residence, and two investigated faith-related factors.

Of the included studies, 62.50% found evidence of cultural bias concerning child maltreatment decisions among professionals of different disciplines, however the extent and nature of this is unclear. With regards to case variables, some vignette studies found that certain professionals judged vignettes with a Black family as less severe and less likely to be reported than identical vignettes with a White family.

However, racial bias was also found in the opposite direction. Some studies found a variety of professionals more likely to judge cases of physical and sexual abuse involving Black families, as abuse, as more serious, and as requiring a report. Moreover, in studies where an effect was found, race was sometimes conflated with other case characteristics such as SES, and only small to moderate amounts of variance were assigned to the studies' variables, leading to inconclusive study results.

Regarding professional variables, while some studies have found no relationship between the race/ethnicity of professionals and their decision-making, others did

find a link. For example, compared to Asian and White preservice teachers, African-American teachers were found to be more accepting of corporal punishment, and to have experienced it more often in their own childhoods, but did not differ from the other racial groups in terms of their ratings of abusiveness.

Studies investigating religion/faith-related variables found evidence of a wide range of practices, spanning several racial and ethnic groups, that might fall into this category, including excessive corporal punishment, medical neglect, ridding-evil practices, and also FGM and honour based violence. The majority of professionals reported that they had not received any specific training about this form of child abuse, suggesting a continued need for multi-agency education and training around faith-related child abuse.

Findings of this review highlight an ongoing need to recognise and mitigate cultural bias among health and social care professionals and teachers. As Western nations continue to grow in cultural diversity, clinicians will be increasingly required to be sensitive to culturally related issues. The field of “cultural competence” has emerged to help adapt services to meet culturally unique needs.

The review highlighted the lack of studies about professionals’ potential cultural bias outside of the US, and towards non-Black/African-American families. Further research is required to examine and compare cultural bias in countries outside of the US and outside of the West, and to expand the assessment of professionals’

potential cultural bias among groups other than White females. The review also raised questions about how cultural biases may interact with related constructs such as gender, SES, and religion. Review limitations, such as possible publication bias and language bias, were also discussed.

## **Conclusion**

This area remains under-researched. This review highlights the need to continue developing practitioners' cultural competence so that professionals are trained and supported to recognise, acknowledge, and where appropriate, mitigate cultural biases. Continued research is needed to determine knowledge gaps, and requirements for training and resources.

# **Empirical Study: The Influence of Culture and Socioeconomic Status on the Mandatory Reporting of Female Genital Mutilation (FGM) by Health and Social Care Professionals and Teachers**

## **Introduction**

Female genital mutilation (FGM) refers to all procedures which alter or cause injury to the female genital organs for non-therapeutic purposes. It is estimated that over 125 million girls and women worldwide are currently living with the consequences of FGM, and that according to current trends another 30 million are at risk of being cut over the next decade (United Nations Children's Fund, 2013). Prevalence data for FGM in the UK and elsewhere in the West is difficult to ascertain. Estimates from Home Office migration figures suggest that in 2011, 137,000 FGM survivors lived in England and Wales. Furthermore, 60,000 girls aged under 14 years were born to mothers who had undergone the procedure (deemed a risk factor for FGM).

International law recognises FGM as one of the most obvious and severe forms of violence against girls and women, and places an obligation on governments to take steps to prevent it. In the UK, it has been unlawful since 1985. The 2015 Serious Crime Act introduced a mandatory reporting duty requiring health and social care professionals and teachers to report *known* cases of FGM (i.e. where a girl discloses she has undergone FGM) in children.

The current figures available suggest that there remains a gap in the reporting of FGM. Despite the assumed prevalence of FGM in the UK, the Metropolitan Police was involved with as few as 145 “incidents of concern” relating to FGM between 2008 and 2011. Reasons for the discrepancy between the assumed prevalence and reporting figures are unclear. One idea is that as society becomes increasingly culturally diverse, clinicians are required to be sensitive to culturally related issues. They might thus face the challenge of determining the appropriate intervention where a client has committed what is considered a crime in the UK and an acceptable practice in another country. In the process, by trying to approach FGM in a culturally sensitive manner, there is a danger that they might avoid necessary interventions for fear of being considered racist. The introduction of mandatory reporting aimed to break down professionals’ concerns regarding cultural sensitivity.

Research suggests that cultural factors (Terao et al., 2001, p. 160) and social factors such as SES (Lopez, 1989) are some of the many variables influencing a professional’s decision to report child abuse. Thus far there has been no research to identify factors affecting reporting and non-reporting of FGM. The study therefore aimed to determine whether cultural sensitivity and SES influence UK professionals’ decision to report FGM, and to explore the relationships between demographic variables and reporting behaviour. It was hypothesised that: 1) professionals will be less likely to report FGM where a family is less integrated into UK culture, and 2) where the family is of higher SES.

## Method

The study utilised an online survey delivered through Qualtrics. An analogue vignette study was chosen, with a 2 x 2 factorial design, with two independent variables: cultural integration and SES. Both had two levels each: high and low. A sample of health and social care professionals and teachers were recruited to complete the online survey.

The first section of the survey required participants to complete a demographic questionnaire that asked about their gender, age, race, and parents' country of birth. They were also asked to indicate their profession, how long they had been qualified or if they were a trainee, and whether they had received any FGM training.

The second section of the survey, used an indirect questioning method whereby four vignettes were presented within a list experiment. The list experiment is an indirect questioning method meant to reduce social desirability bias. Each vignette was presented in turn in a randomly assigned order with a list of four non-sensitive yes/no items for the control group, and with a list of four non-sensitive yes/no items **plus** the FGM-sensitive yes/no item ('are you concerned enough to report this family directly to the police, rather than continuing to monitor them within your service?') for the treatment group. Respondents were asked how many of the list of questions they agreed with. The idea being that as long as the entire list does not apply to them, they are assured that we will not know their answer to the sensitive question. Responses to the lists (specifically, the number of yes responses) constituted the

dependent variable. The difference in the total number of items between control and treatment group identified the proportion of respondents that agree with the sensitive item.

In the third section, respondents were again presented with the same four vignettes (in the same randomly assigned order) and asked to directly answer the FGM-sensitive yes/no item for each. The number of yes responses constituted the dependent variable.

There were 226 attempts at completing the online survey, and a total of 205 health and social care professionals and teachers completed it. Participants were primarily female (79.51%) and White (77.07%). Almost half were psychologists. The majority were aged 26-30, 31-35, and 36-40. Almost half of professionals had been qualified for 5 or more years (42.44%), and the majority of participants (77.07%) had received either no FGM training at all (37.07%) or 1-2 hours only (40%).

Responses to the list experiment that followed the vignettes were analysed using multivariate regression. We further examined responses where the FGM-sensitive item was asked directly, and investigated the relationship between the FGM-sensitive items and the respondents' characteristics. Missing data were dealt with by using both the complete cases method (CC) and the inverse probability weighting method (IPW).

## **Results and Discussion**

Whether asked indirectly or directly, the majority of professionals sampled said that they would report a *known* case of FGM in just under 80% of instances. Moreover, direct questioning suggested that professionals were slightly less likely to report to the police when the family was described as being both highly integrated into UK culture, and of high SES. In the absence of other UK FGM reporting statistics, results tentatively suggest that for the most part the mandatory reporting duty appears to have been successful in breaking down professionals' concerns about needing to be cultural sensitive towards practising families. It is also possible that perceived affluence coupled with high acculturation might dissuade professionals from reporting for fear of causing more harm to the child. Results also indicated that the majority of professionals sampled had received either no FGM training at all, or very little (1-2 hours). The potential reasons for and implications of the findings are discussed. The study's strengths and limitations are outlined, and directions for future research are considered.

## **Conclusion**

Professionals need to be supported to develop skills in cultural awareness so that they are aware of potential biases in this area.

## **Integration, Impact, and Dissemination Summary**

### **Integration**

While FGM might be cultural practice that is only relatively recently being understood as child abuse, it is one of several practices that is normative to some groups, but also potentially harmful to children within those groups. It was hoped that the systematic review would provide a comprehensive overview of the role of cultural factors in professionals' decisions about child maltreatment. The empirical study could then narrow down and provide a detailed understanding of the role of cultural sensitivity (and SES) on professionals' reporting of a currently salient cultural practice, namely FGM. The systematic review therefore provided an overview and clear rationale for the empirical article, and assisted its development in a number of ways that are outlined. Challenges, and what was learnt during the process of integration, are also discussed.

### **Impact**

Potentially, the research findings could have a far-reaching impact, as they relate to the practice of all professionals working with children and families. Findings of the review could help develop cultural competency aspects of professionals' training. Cultural competency training provision would benefit from developing a more rigorous evidence base. In the current climate of ongoing migration, policy-makers would also benefit.

The findings of both the systematic review and the empirical study may be of interest to service users, particularly those impacted by FGM, but also any who are from a minority background. Both papers could help them to make sense of how professionals might relate to their culture and cultural practices. Organisations that offer information and support about FGM may also benefit from findings. Finally, the personal impact on the principal investigator (FJ) is also discussed.

### **Dissemination**

Firstly, dissemination will be via the provision of a summary of the research to interested study participants. Secondly, the findings of the empirical study have been disseminated locally via a presentation to staff and students at Royal Holloway University, and will be disseminated during a Continuing Professional Development session at one of the recruitment sites. Thirdly, the empirical study and the review will be submitted to peer-reviewed journals, and application will be made to present the findings at relevant mental health conferences. Fourthly, findings will be incorporated into the FGM training that the principal investigator (FJ) delivers to various professional bodies. Feedback will be sought after each workshop to ensure that impact is maximised, and also evidenced. Finally, a short summary of the research will be posted on social media sites (LinkedIn, Facebook, and Twitter). This is in order to communicate findings to a wide audience which could potentially lead to further opportunities to disseminate the research generated in these studies.

**Cultural Competence: A Systematic Review on the Role of Cultural Factors in Professionals' Decisions about Child Maltreatment**

## **Abstract**

Child maltreatment is a serious worldwide problem. However, there remains insufficient guidance on how to interpret and implement child protection legislation. Current definitions of child maltreatment do not properly account for cultural differences, and the extent of professionals' cultural sensitivity or bias in decisions about child maltreatment remains unclear. Despite ongoing migration leading to increasingly ethnically diverse populations, Western-based professionals' cultural bias has not been systematically investigated. We aimed to conduct a systematic review of the literature to address this gap.

The review was conducted according to guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. Peer-reviewed articles published in English which used quantitative methods were reviewed. We sought evidence about how cultural factors might influence professionals' decisions about child maltreatment.

Of 390 unique articles, 16 met inclusion criteria. The methodological quality was mostly moderate. Study participants included social workers, teachers and psychologists. Studies were mostly conducted in the United States and Canada, with one in the United Kingdom and another conducted in Sweden and Croatia. All studies were cross-sectional, with the majority employing indirect vignette methods. 81.25% of studies measured race/ethnicity variables, while two investigated faith-related factors, and one was concerned with country of residence. 62.50% of studies

found evidence of cultural bias concerning child maltreatment decisions among professionals of different disciplines, regarding both case and professional variables. However, the extent and nature is unclear. A recent study suggests that professionals lack confidence in their ability to identify and respond to this form of abuse.

This area remains under-researched. This review highlights the need to continue developing practitioners' cultural competence so that professionals are trained and supported to recognise, acknowledge, and where appropriate, mitigate cultural biases. Continued research is needed to determine knowledge gaps, and requirements for training and resources.

## **Introduction**

Maltreatment of children, incorporating various types of abuse and neglect, occurs when a caretaker or responsible person harms, threatens to harm, or fails to provide adequate care for a child (Kemoli & Mavindu, 2014, p. 256). It is one of the most powerful risk factors for concurrent and subsequent psychopathology, later health morbidity, and compromised development (Zeanah & Humphreys, 2018, p. 637). As such, it is a serious worldwide problem, occurring in every community, at all levels of society (Hoyano & Keenan, 2007). The World Health Organisation (WHO) in association with the United Nations Children's Fund (UNICEF) has called for maltreatment to be recognised as a global public health concern (Moody, Cannings-John, Hood, Kemp, & Robling, 2018, p. 2).

## **Classification of Child Maltreatment**

WHO defines child maltreatment as “all forms of physical and/or emotional ill-treatment, sexual abuse, neglect or negligent treatment or commercial or other exploitation, resulting in actual or potential harm to the child's health, survival, development or dignity in the context of a relationship of responsibility, trust or power”. It also acknowledges that the four categories may coexist in one child (World Health Organization, 1999, p. 29).

The Longitudinal Studies of Child Abuse and Neglect (LONGSCAN) Modified Maltreatment Classification System (MMCS) (English, 1997) also defines and classifies types of maltreatment into five categories; physical abuse, sexual abuse,

physical neglect involving failure to provide, physical neglect involving lack of supervision, and emotional abuse.

In the MMCS (English, 1997), physical abuse is categorised as involving the infliction of physical injury upon a child due to a deliberate attempt to hurt that child, or serious discipline or physical punishment inappropriate to the child's age.

Sexual abuse occurs when there is sexual contact or inappropriate sexual exposure, or an attempt at either, occurring between a caregiver or other responsible adult and a child, for purposes of the caregiver's sexual gratification or financial benefit.

Physical neglect, or failure to provide, occurs when an able caregiver or responsible adult fails to exercise a minimum degree of care in meeting the child's basic needs for food, clothing, hygiene, shelter, supervision, medical and dental care, and support. Physical neglect can also involve failing to provide supervision appropriate to the child's emotional and developmental needs.

Finally, emotional abuse is a broad category encompassing the persistent or extreme impeding of a child's basic emotional needs. This includes parental acts that are harmful because they are insensitive to the child's developmental level, including psychological safety and security, acceptance and self-esteem, and age-appropriate autonomy. Examples include frequent yelling, bullying, rejecting the child as punishment, and exposing the child to violence or demeaning acts.

## **Problems with the Definition of Child Maltreatment**

In 1981, Besharov (p. 384) noted that alarmingly, definitions of child abuse and child neglect often failed to meet research needs due to their lack of comparability, reliability, and universally understood delineations. He demonstrated that due to the absence of a widely-accepted and clear definition of child maltreatment, researchers have had to develop and use their own idiosyncratic definitional measures and variables, resulting in as almost as many definitions as number of research projects (Besharov, 1981, p. 384). He also argued that while the courts might benefit from a case-by-case decision-making process, this was detrimental to good research. Where researchers cannot specifically describe the types of child rearing situations that should be labelled as abuse and neglect, data coders are forced to make subjective assessments of each case in a study, resulting in unpredictability of their decisions and therefore undermining studies' measurement reliability (Besharov, 1981, p. 385).

Moreover, because child maltreatment has a wide variety of expressions and causes, it is problematic to assume that the populations considered in child maltreatment studies are representative of all maltreating families, and then generalise findings from these studies to all forms of child maltreatment (Besharov, 1981, p. 386). For example, with regards to the definition of neglect, reasons why a parent might fail to adequately feed or clothe their child may vary considerably, including to deliberately harm or inflict cruelty on their child, to punish their child, but also as a result of

abject poverty. Such different circumstances would require very different treatment or intervention responses from professionals.

Unfortunately the alarm sounded by Besharov (1981) cannot yet be quietened. Though there have been numerous campaigns that have provided the basis for legal reform in the sphere of child maltreatment, these legal initiatives have been described as ad hoc responses to specific problems, rather than part of a coherent and integrated programme of reform across national criminal and civil systems, much less internationally (Hoyano & Keenan, 2007). Despite numerous legislation and literature that identifies conditions, injuries, and behaviour that may warrant concern for possible child abuse, there remains very little guidance on how to interpret and implement statutory terms such as “significant harm” and “reasonable suspicion”, or what levels of concern should be reported (Levi & Crowell, 2011, p. 321.)

### **Culture and Definitions of Child Maltreatment**

Leading social scientists Richerson and Boyd (2005, p. 5) define culture as “information capable of affecting individuals’ behavior that they acquire from other members of their species through teaching, imitation, and other forms of social transmission.” Cultural knowledge therefore encompasses information and skills that an individual could not have developed in a lifetime, and evolutionary anthropologists suggest that this cumulative knowledge, along with genetic variations, has allowed for human adaptation in diverse environments (Richerson &

Boyd, 2005). As the definition suggests, culture is intricately linked to other identity markers such as race, ethnicity, language, and religion, with researchers still sometimes using the terms race, ethnicity, and culture interchangeably (Pfeffer, 1998, p. 1381).

While a broad definition of child maltreatment allows courts enough flexibility to respond to real-life situations, this stance also negates social and cultural differences, making any cross-cultural comparisons especially problematic. Concepts such as failure to provide, or age-appropriate supervision and autonomy, reveal Western-centric ideas that are unworkable in many other communities. For example, while the definition of neglect includes failing to provide a child with adequate dental care, many people living in rural areas across the world, including West Africa (Varenne, Petersen, & Ouattara, 2004, p. 84), South Asia (Saravanan et al., 2008), and South America (Castilho, Ferreira, & Perini, 2009), struggle to access basic oral health care due to geographical and economic barriers.

Moreover, parents generally have a broad discretion in the ways in which they interact and parent their children (Terao, Borrego, & Urquiza, 2001, p. 162). Issues of culture may therefore play an important role in professionals' understanding of culturally diverse parenting practices, and consequently how they respond to children and families of different cultural backgrounds (Terao et al., 2001, p. 162). Cultural practices might be simultaneously normative and also harmful to the child. For example, some studies suggest that African-Americans and Hispanics in the

United States (US) have higher corporal punishment acceptability in comparison to other ethnic groups (e.g. Pinderhughes, Dodge, Bates, Pettit, & Zelli, 2000; Wissow, 2001). In Europe, smacking remains legal in the United Kingdom (UK) but outlawed in other parts of the continent (Moody et al., 2018, p. 11). Definitions of physical abuse may therefore suffer from cultural preconceptions. Consequently, mandated reporters may find it difficult to decide on cases where harsh physical discipline may be seen as a normative cultural practice due to ethnicity or cultural background (Terao et al. (2001, p. 162).

Female genital mutilation (FGM) is another example of a cultural practice that is normative to some groups, and also harmful to affected children. Having been historically practiced by some communities primarily from Africa, but also Asia and the Middle East, it is described as a cultural practice (Vogt, Efferson, & Fehr, 2017, p. 283). It is understood as a parental practice perpetuated by families' beliefs that it is in the child's best interests (Costello, Quinn, Tatchell, Jordan, & Neophytou, 2015, p. 1264). However, international law has identified FGM as one of the most obvious and severe forms of violence against girls and women, and places an obligation on governments to take steps to prevent it (Christou & Fowles, 2015, p. 349). It is being increasingly thought of as child abuse (Amasanti, Imcha, & Momoh, 2016), and has been unlawful in the UK since 1985. In 2015 a mandatory reporting duty was introduced requiring health and social care professionals and teachers to report *known* cases of FGM in under 18-year-olds. The introduction of mandatory reporting aimed to break down professionals' concerns regarding cultural sensitivity (Mathers

& Rymer, 2015, p. 283). In contrast, the LONGSCAN Modified Maltreatment Classification System (English, 1997) excludes culturally sanctioned physical alterations such as circumcision and ear piercing from its definition of physical abuse (Zeanah & Humphreys, 2018, p. 638). Navigating this confusion poses a challenge for professionals faced with the responsibility of deciding on appropriate interventions where clients have committed what to them might be an acceptable and necessary practice, but is considered a crime in the UK (Terao et. al, 2001, p. 159).

Potential consequences of failing to consider culture in the application of child protection laws include biased reporting, errors in assessing perceived risk, ineffective interventions, and increased out-of-home placements (Rubin, 1992). For example, the rate of African-American children in child maltreatment prevalence rates is about twice that of their representation in the US population as a whole, and it has been suggested that this problematic over-representation of ethnic minorities in child maltreatment figures reflects the over-reporting of ethnic minority groups and the under-reporting of Caucasians due to cultural factors, such as childrearing practices (Ibanez, Borrego, Pemberton, & Terao, 2006, p. 1366). In failing to properly understand how culture might impact child protection, there is therefore a risk that already disadvantaged groups might be further affected by discriminatory clinical practices.

## **Other Factors Influencing Definitions of Child Maltreatment**

Research suggests that there are several variables influencing a professional's decision to report child abuse (e.g. Glachan, 1991; Rubin, 1992; Glough, 1996), with many inconsistencies among professionals of what may or may not be reportable (Rubin, 1992, p. 249; Terao et al., 2001, p. 160). Biases in clinical judgement have been found to be related to variables such as the client's ethnicity, socioeconomic status (SES), age, and gender (Lopez, 1989). For example, Using US Department of Health and Human Services data, Hampton and Newberger (1985) examined the effects of a range of case characteristics on the reporting behaviour of hospital staff and found that cases where the child was Black or relatively less affluent were more likely to be reported. These variables might in turn interact with other variables that have been found to influence professionals' reporting decisions, such as the professional's training background (Tilden et al., 1994), or the perceived severity of the case (Zellman, 1992). Terao et al. (2001, p. 161) suggest that clinically, this might mean that professionals might respond differently to different demographic variables, particularly concerning culturally diverse parenting practices that require professionals to use their own judgment to evaluate whether or not an incident should be reported (Ibanez et al., 2006, p. 1366). In their 2001 paper, Terao et al. grouped these variables into professional, perpetrator, and case characteristics.

Regarding professionals' characteristics, strong predictors of reporting behaviour include the reporters' professional background (Zellman, 1990b), their knowledge of child abuse, amount of training, attitudes towards the case, and the institutional

setting in which the case is considered (Terao et al., 2001, p. 160). Interestingly, a professionals' training background (e.g. medicine vs. social work) has been shown to have greater influence than actual knowledge about child abuse (Tilden et al., 1994). Data has failed to support the hypothesis that professionals' demographic variables such as age, gender, and parenthood might influence their reporting behaviour (Ashton, 1999).

Concerning perpetrator characteristics, Zellman (1992) found that perpetrators of low SES had an increased likelihood of being reported. Other strong predictors of making a report included the reporter's perception of the perpetrator as being lazy and angry, whether a previous report of abuse had been made (Zellman, 1992), and whether there was intent to harm the child (Portwood, 1999).

Finally, with case characteristics, studies have found that the two variables likely to significantly influence professionals' judgement about reporting include the perceived seriousness of the incident (Ashton, 1999; Zellman, 1992) and whether sufficient evidence was available (Zellman, 1990a). Frequency of the abuse, (Portwood, 1999), history of previous abuse, and denial of the abuse have also been found to be strong predictors of reporting behaviour (Zellman 1992; Zellman & Faller, 1999). Moreover, professionals have been found to be more likely to report cases that involve younger children (Ashton, 1999; Tang, 1998), that are sexual in nature (Portwood, 1999), and that involve actual physical or psychological harm to the child (Portwood, 1999).

## **Prevalence**

Unsurprisingly, estimating the prevalence of child maltreatment is a major challenge. There are few worldwide studies that have comprehensively addressed the issue (May-Chahal & Cawson, 2005, p. 970). Data collected from public sector institutions such as social services or the police form only a portion of the true numbers, as there are many more cases that go undetected, unrecorded, or unreported (Moody et al., 2018, p. 2). Moreover, ideas of what constitutes maltreatment can vary and may be influenced by intergenerational changes in attitudes and cross-cultural differences, amongst other things. For example, as discussed above, while smacking remains legal in the UK, unlike in many other parts of Europe, even in the UK it is much less socially acceptable than it was in the 1980s. Studies have also shown that children might experience severe lack of care, physical violence, or sexual assault and not rate themselves as abused (May-Chahal & Cawson, 2005, p. 982).

Nevertheless, attempts have been made to establish child maltreatment prevalence rates. Internationally it is estimated that 40 million children experience abuse annually (World Health Organization, 2014). Sexual abuse is the most researched category in the area (May-Chahal & Cawson, 2005, p. 981). A recent systematic review (Moody et al., 2018) aimed to expand on previous findings by including worldwide prevalence rates of physical, emotional abuse, and neglect. It looked at 337 studies in which participants, whether adult (18+) or children, self-reported lifetime child maltreatment before the age of 18 years. For sexual abuse, half of the

study samples (171 of 337) were found in North America. It found a median prevalence rate for sexual abuse of 20.4% in North America (13.2% to 33.6%) and 14.3% in Europe (7.8% to 28.0%). 28.8% (17.0% to 40.2%) was found in Australian girls. Boys generally had lower rates ranging from 14.1% in North America (4.3% to 21.0%), and 6.2% in Europe (4.8% to 15.2%) (Moody et al., 2018, p. 4). However, it should be noted that gender of the participant may influence reporting as evidence suggests that men may be less likely to reveal a history of maltreatment. Perhaps definitions of maltreatment do not adequately capture the experiences of males, specifically for sexual abuse. Moreover, men might be particularly affected by fears of being labelled as weak or as homosexual (Moody et al., 2018, p. 9).

Rates of physical abuse were more similar across genders, apart from in Europe, where physical abuse was much higher for boys (27.0%, ranging from 7.0% to 43.0%) than for girls (12.0%, ranging from 6.9% to 23.0%). In North America, where the majority of studies had been undertaken, prevalence rates were similar for boys and girls at 24.3% (14.1% to 32.1%) and 21.7% (14.2% to 33.3%) respectively. Possibly due to a Western-centric definition of physical abuse, rates were found to be very high in some continents, for example, 50.8% (36.0% to 73.8%) and 60.2% (43.0% to 84.9%) for girls and boys respectively in Africa (Moody et al., 2018, p. 4).

Median rates of emotional abuse were nearly double for girls than boys in North America (28.4% vs 13.8% respectively) and Europe (12.9% vs 6.2% respectively), but more similar across genders groups elsewhere. Median rates of neglect were highest

in Africa (girls: 41.8%, boys: 39.1%) and South America (girls: 54.8%, boys: 56.7%) but were based on few studies in total, and as discussed above, may be coloured by Western-centric definitions of neglect. In North America, the continent with the highest number of studies, median rates differed between girls (40.5%) and boys (16.6%).

The review included 18 UK studies showing considerable variation in lifetime prevalence rates of self-reported maltreatment in childhood (Moody et al., 2018, p. 5). Prevalence of sexual abuse ranged from 0.7% to 27.8%. Prevalence of physical abuse ranged from 3.6% to 32.6%. Prevalence of emotional or psychological abuse ranged from 4% to 66.7%, and prevalence of neglect ranged from 5.6% to 77.8%. Finally, the prevalence of unspecified maltreatment ranged from 9.5% to 48.4%.

Concerning prevalence of child abuse in the UK, May-Chahal and Cawson (2005) also conducted a unique study in which 2,869 young adults aged 18 to 24, obtained by random probability sampling throughout the UK, were interviewed face to face by trained interviewers. Maltreatment was defined using a post hoc assessment of a range of experiences while the respondents were aged 16 or under. They found that over 90% of respondents reported that they came from a warm and loving family background, and 83% described themselves as having been “very well cared for”. However, maltreatment (both intra and extra-familial) was experienced by 16% of the sample, suggesting that for some respondents, coming from a warm and loving background and experiences of maltreatment are not mutually exclusive. Serious

maltreatment was experienced by 7% of respondents for physical abuse, 6% for emotional abuse, 6% for absence of care, 5% for absence of supervision, and 11% reported sexual abuse involving contact. Despite the existence of a developed child protection system over the last two decades in the UK, child maltreatment rates remain unacceptably high (May-Chahal & Cawson, 2005, p. 982).

### **Review objectives**

Researchers (e.g. Rubin, 1992; Terao et al., 2001) have called for further research on the impact of cultural bias on the reporting process for cases of child abuse, including the subjectivity involved in interpreting the requirements of reporting laws, and how professionals exercise their discretion in deciding whether to comply and how. Examining how cultural factors might influence the decision-making process may provide guidelines for developing a more clear definition of child maltreatment and limit subjective interpretations of reporting duties, leading to more uniform and multi-disciplinary reporting procedures and behaviours (Ibanez et al., 2006, p. 1366).

While there are systematic reviews on related topics such as the extent of healthcare provider racial discrimination (Paradies, Truong, & Priest, 2013) and the influence of implicit racial and ethnic bias among health care professionals on health care outcomes (Hall et al., 2015), a major limitation that has been identified in the current research is the use of race to distinguish population groups. This is because race does not capture the full diversity present in all racial groups and is insufficient to infer the cultural preferences of an individual within that group. Researchers have

therefore argued that the next logical step is to further explore participants' cultural characteristics (Kesner et al., 2016, p. 324). The aim of this systematic review is therefore to provide a comprehensive amalgamation of the existing literature on the role of culture in Western-based professionals' decisions about child maltreatment. To the best of our knowledge this is the first systematic review to look at the relationship between the more broadly defined 'culture' and professionals' decision-making. We aimed to answer the following question: how do cultural factors, whether professional, perpetrator, or case, influence professionals' decisions about child maltreatment?

## **Method**

### **Data Sources**

A social sciences librarian was consulted to assist with determining relevant computerised databases and search terms to use. The following databases were searched for studies: PsycINFO (within 'abstract', yielding results from 1971), PubMed (within 'all fields', yielding results from 1982), and Web of Science (within 'topic', yielding results from 1991). Searches were conducted in February 2019. The academic journals/articles filter was used. We used the following search string to search the databases for studies completed prior to 22<sup>nd</sup> February 2019:

("healthcare professional\*" OR "health professional\*" OR "health personnel" OR "health staff" OR "health worker\*" OR clinician\* OR psychologist\* OR therapist\* OR

"social worker\*" OR "social care worker\*" OR "social care professional\*" OR "case worker\*" OR "care coordinator" OR teacher\*)

AND

("report\*" OR "decision-making" OR "decision making")

AND

("child abuse" OR "child neglect" OR "child maltreatment" OR "child welfare" OR "child protection" OR "child safeguarding")

AND

(nation\* OR ethnic\* OR race OR cultur\* OR "cultural competenc\*" OR intercultur\* OR "cross-cultural" OR "cross cultural" OR "transcultural" OR "trans-cultural" OR "cultural diversity" OR "cultural diversities" OR "multicultural" OR "multi-cultural").

### **Study Selection**

A checklist of the inclusion and exclusion criteria was created prior to the search to assist with eligibility assessment. We included studies in the review if they met the following criteria:

- (1) Collected original data from participants who are health and social care professionals or teachers. This is the professional group involved in child protection. Examples include psychologists, doctors, dentists, nurses, midwives, and social workers.

- (2) Measured and reported results on cultural factors (including race and ethnicity) that influence professionals' decision-making and/or reporting behaviour concerning child maltreatment (children aged 18 or below).
- (3) Were published in a peer reviewed journal irrespective of publication year. This was in order to establish a minimum level of study quality.
- (4) Were written in English. This was due to our limited proficiency in other languages. The use of translators for non-English speaking texts would have been beyond the scope of this review.

We excluded:

- (1) Articles or reports that are strictly theoretical or conceptual.
- (2) Qualitative studies; these use inductive or exploratory data analysis methods, while quantitative methods employ confirmatory or deductive analysis methods (Barker, Pistrang, & Elliott, 2002, p. 229). Including qualitative studies would therefore compromise heterogeneity. Reviews, case histories, case studies, and historical, ethical or educational analyses were also excluded. This was in order to further manage heterogeneity (Mulrow, Langhorne, & Grimshaw, 1997), and ensure a focus on empirical data.
- (3) Studies that did not specifically from the outset, measure and report on cultural factors influencing professionals' decisions about child maltreatment, even if these were later found to be relevant, were also excluded in order to manage heterogeneity, and study and measurement quality.

- (4) Studies originating from Asia, Africa, the Middle East, and South America were excluded since understanding of child maltreatment and child abuse legislation in these countries is socially, historically and culturally different from Western countries (Moody et al., 2018, p. 10) and therefore not suitable for this review. Moreover, cultural bias or cultural sensitivity may be a different phenomenon in non-Western countries (Hall et al., 2015, p. e62).
- (5) Similarly, cross-cultural studies comparing the decision-making or reporting behaviour of professionals in Western countries with those in non-Western countries, were also excluded.

### **Data Synthesis and Quality Appraisal**

The quality of each eligible study was assessed using the Public Health Wales Observatory (PHWO) critical appraisal tool for cross-sectional studies (2014) (see Appendix A). The tool is adapted from the Health Evidence Bulletin Wales critical appraisal tool (Weightman, Barker, & Lancaster, 2000), which itself is a modified version of the Critical Appraisal Skills Programme (CASP). The instrument assesses key areas of study quality, including clarity of aims, appropriateness and rigour of study design, execution, and analysis, including consideration of possible bias, and relevance of results. It was chosen due to its applicability to observational studies. Moreover, it has been used effectively in Paradies et al.'s (2013) systematic review about the extent of healthcare providers' racism, which is another review of cross-sectional studies.

While the tool has 11 sections, two items were left unrated. Item 6 was left unmarked due to being the only negatively worded question in the tool, so that allocating a point for a yes answer to this question would skew results. Item 11 was also left unrated as it is intended for use when screening for eligibility, and it was decided that due to the low number of studies in the review, all studies fulfilling selection criteria would be included. The number of “yes” answers on the checklist was translated into points, for a maximum score of 9 points. Half a point was given if a study fulfilled some of the criteria in a section. Studies received a total quality score ranging from one to nine (1 to 2: low, 3: low - moderate, 4 - 6: moderate, 7: moderate - high, and 8 to 9: high). Ambiguities about study quality were resolved by detailed examining of the full text source documents.

## **Results**

### **Search Results and Data Extraction**

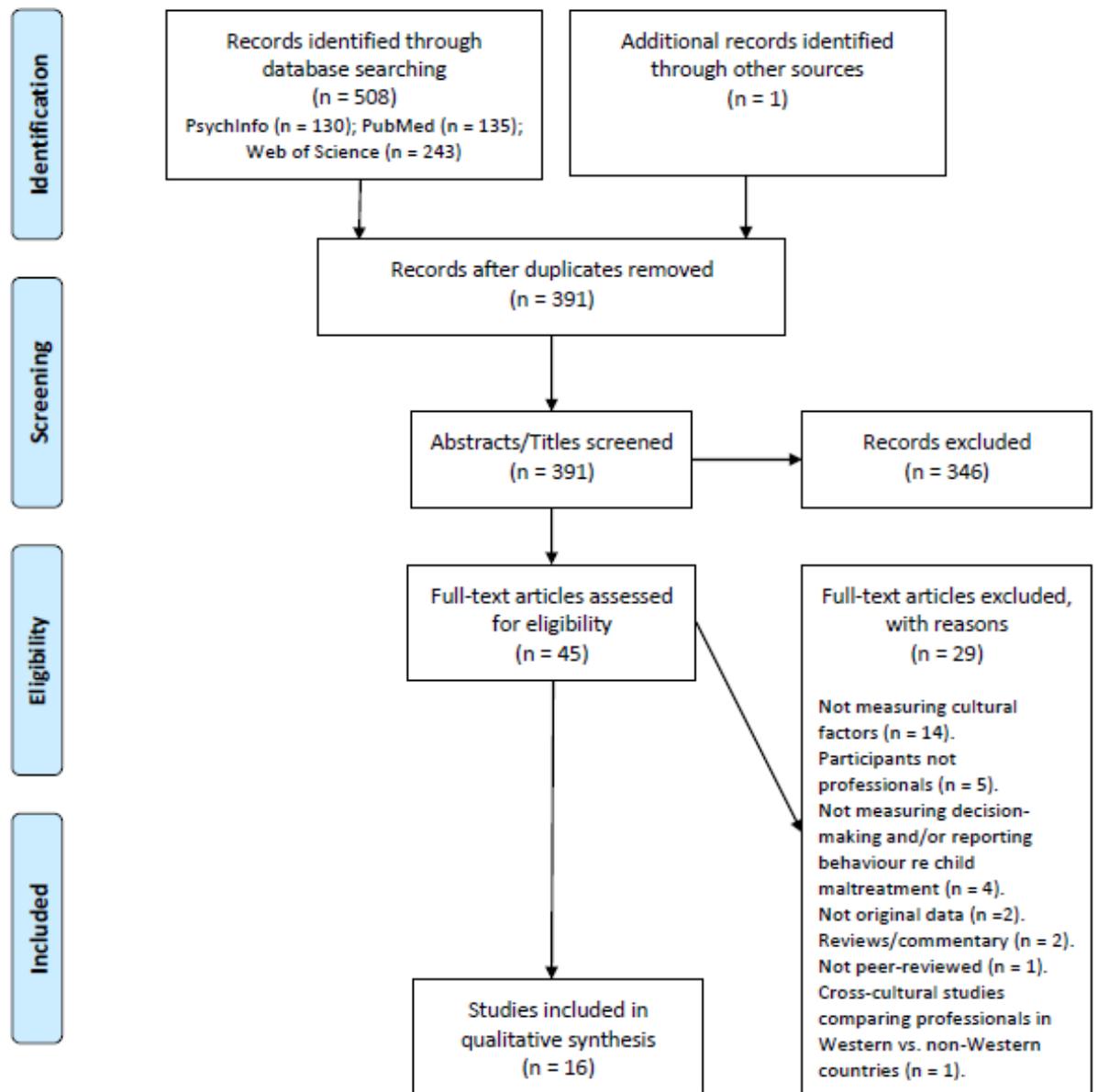
After performing the electronic searches, 508 results were checked in order to remove duplicates, after which 390 abstracts remained. PsychInfo yielded 130, and there were also 135 from the PubMed, and 243 from the Web of science databases (Fig. 1). The researcher (FJ) independently screened the articles to determine eligibility. Studies were included or excluded after reading the title and abstract; however, it was also necessary to examine the full text document of some studies to determine eligibility. One further study was identified during this process. Inter-rater agreement of study selection was assessed through verification of the selection decision by the supervisor CE. CE rated 5 of the 45 (just over 10%) full text articles

assessed for eligibility. Based on the full text documents of all the 5 papers, CE and FJ's agreement according to the selection checklist was 100%.

After completing the inclusion and exclusion process, 16 peer-reviewed journal articles were included in this review for data extraction. Figure 1 shows the process of identifying and including studies. A data extraction document was used to assist with identifying and collecting relevant information from the included studies. Information extracted included the citation, purpose of the study, study method, study location, sample type, and results and findings.

To recap; this review aimed to synthesise the current literature regarding the role of cultural factors in Western-based professionals' decisions about child maltreatment. We aimed to investigate the manner in which cultural factors, whether professional, perpetrator, or case, influence professionals' decisions about child maltreatment. Findings were not suitable for a meta-analysis due to heterogeneous populations, outcome measures, and study design (Egger, Schneider, & Smith, 1998; Liberati, 1995). A summary of the methodological characteristics of the studies is therefore presented, followed by a narrative synthesis of the substantive findings regarding the role of cultural factors in professionals' judgements about child maltreatment and the impact of this on their clinical practice.

Figure 1. PRISMA (Moher, Liberati, Tetzlaff, & Altman, 2009) flow diagram of the systematic review process.



## Qualitative Synthesis of Selected Studies

### Study design characteristics.

A summary of the main characteristics of each study is detailed in Table 1, including study design, country, sample and setting, and racial/ ethnic background of the professional and patient groups (where applicable).

**Table 1. Characteristics of 16 studies measuring the impact of cultural factors on professionals' decisions about child abuse (CA).**

Study	Country	Study aims, design, and method	Sample/ Setting	Response rate	Professional racial/ ethnic background	Patient/client racial/ethnic background (real or hypothetical)	Gender/ Age
Bottoms et al., (1995).	USA.	To determine nature and outcome of cases of faith-related CA reported to professionals nationwide.  Cross sectional survey: non-validated tool designed for the study.	Clinical psychologists (n = 5998). Psychiatrists (n = 7381). Social workers (n = 5896).  Recruited via random sampling from mailing lists of professional bodies.  Part 1: 6939 valid respondents (2722 CPs, 2083 Ps, and 2134 SWs), of whom 2136 reported encountering at least one ritualistic or religion-related CA case. Part 2: 797 responded to follow-up surveys, of which 720 deemed valid (297 CPs, 200 Ps, 223 SWs).	Part 1: 37%.  Part 2: 37%.	Not reported.	Not reported.	Not reported.

Britner & Mossler, (2002).	USA.	<p>To study impact of child's ethnicity, child's age, and chronicity of abuse on different professionals' placement decisions following instances of CA (physical). And to study how professionals prioritise 18 different pieces of information (e.g. parents' substance abuse).</p> <p>Cross sectional survey: non-validated, randomised vignettes + tool designed for the study.</p>	<p>Total sample (n = 90).</p> <p>Professionals: social workers and child protective services personnel (n = 43), mental health experts/consultants, e.g., clinical psychologists (n= 23), guardians (n = 8), judges (n = 6), court appointed special advocates (n = 10).</p> <p>Original sample (n = 150). Recruited via mail in five Virginia jurisdictions.</p>	60%.	White (83%), non-White (17%).	African-American, White.	70% female, (M = 40.2 years).
Egu & Weis, (2003).	USA.	<p>To study impact of professional's race, child's race, and severity of abuse on professionals' recognition and reporting of abuse (physical).</p> <p>Cross sectional survey: non-validated, randomised vignettes, with validated photos of the child, + tool designed for the study.</p>	Teachers (n = 540). Recruited from credential or graduate classes in five universities in Southern California.	"Almost all".	White (n = 180), Black (n = 180), Hispanic (n = 180).	White, Black, Hispanic.	63.5% female, (M = 33 years).

Hansen et al., (1997).	USA.	To study impact of case (race, SES, age of victim, type of maltreatment) and professional (gender, training, experience of reporting, personal history of maltreatment) characteristics on professionals' reporting of CA (neglect, sexual, physical, and psychological).  Cross sectional survey: non-validated, randomised vignettes + tool designed for the study.	Psychologists (n = 125). Recruited via random sampling drawn from mailing lists of Nebraska, Iowa, Kansas, and Missouri licensing boards.  Social workers (n = 85). Recruited via random sampling drawn from mailing lists of Nebraska licensing boards.	Psychologists 43.4%.  Social workers 40.5%.	Psychologists: White non-Hispanic (95.2%), Native-American (1.6%), African-American (.8%), Asian-American (.8%), Hispanic (.8%), "Other" (.8%).  Social workers: White (91.8%), African-American (3.5%), Native-American (2.35%), Hispanic (2.35%).	African-American or White	Psychologists 55.2% female, (M = 46.2 years).  Social workers 52.9% female, (M = 46.6 years).
Jackson & Nuttall, (1994).	USA.	To study impact of victim and perpetrator race on professionals' judgments about sexual abuse.  Cross sectional survey: non-validated, non-randomised vignettes + tool designed for the study.	Social workers (n = 172).  Recruited via stratified random sampling drawn from the national directory, equally distributed by gender.	40%.	Not collected.	White, Minority.	Original sample: 50/50 male/female.  45% of females responded vs. 33% males.  Aged 30 to 77 years (M = 48.3 years).
Kenny, (2001).	USA.	To study impact of professionals' variables (ethnicity age, gender, years teaching, education, level of training) on professionals' reporting behaviour.  Cross sectional survey: validated, non-randomised vignettes + validated Teachers and Child Abuse Questionnaire + Educators and Child Abuse Questionnaire (ECAQ).	Teachers (n = 197).  Recruited from teacher education workshops in 40 schools in an urban southeast district with over 300 schools.	90%.	44% White, 36% Hispanic, 11% African-American, 5% Afro-Caribbean, .5% Asian, 4.5% "other."	Not applicable.	80% female.  Aged 21 to 66 years, (M = 37 years).

Kesner et al., (2016).	USA.	<p>To study impact of professionals' race on attitudes towards corporal punishment and reporting behaviour.</p> <p>Cross sectional survey: non-validated, non-randomised vignettes + Corporal Punishment Scale (CPS) + Educators and Child Abuse Questionnaire (ECAQ).</p>	<p>Preservice teachers (n = 105).</p> <p>Recruited from a child development class for elementary teacher education majors.</p>	100%	43% African-American, 33% White, 24% Asian.	Not applicable.	95% female. Aged 20 to 45 years (M = 23.5 years).
Kesner & Stenhouse, (2018).	USA.	<p>To study the effect of race and culture on preservice teachers' perceptions of corporal punishment and its effect on mandated reporting.</p> <p>Cross sectional survey: non-validated, non-randomised vignettes + Demographics and Hollingshead Four-Factor Index of Socioeconomic Status (HFFISS) + Corporal Punishment Scale (CPS) + Educators and Child Abuse Questionnaire (ECAQ) + non-validated Corporal Punishment Experiences (CPE) designed for the study.</p>	<p>Preservice teachers (n = 51).</p> <p>Recruited from a developmental psychology class at a large urban university.</p>	96%.	42% African-American, 38% Anglo-White, 20% Asian.	Not applicable.	96% female. Aged 20-35 years (M = 22.8 years).

Levi & Crowell, (2011).	USA.	To study impact of professionals' characteristics (race, age, gender, professional training, seniority, or prior education) on the extent to which a standard interpretation of reasonable suspicion exists among professionals.  Cross sectional survey: piloted, validated tools (Differential Diagnosis Scale (DDS) and Estimated Probability Scale (EPS)).	Doctors (n = 81). Original sample (n = 117).  Recruited at an annual conference.	69%.	85% White, 3% African-American, 1% Hispanic, 11% Other.	Not applicable.	69% female. Aged 31-75 years (M = 47 years).
Maiter, (2004).	Canada.	To study whether there are differences in definitions of CA (corporal and psychological discipline, and inappropriate supervision) between professionals and South-Asian parents.  Cross sectional survey: piloted, randomised vignettes + non-validated tool designed for the study.	Social workers (n = 14) recruited from a large urban child protection agency.  South-Asian Canadian parents (n = 21). Recruited via flyers and word of mouth.	Not reported.	37.5% Jewish, 21.4% Canadian, 14.3% South-Asian.	37.9% India, 27.6% Sri Lanka, 24.1% Pakistan.	Professionals 85.7% female. Aged 24 to 52 years (mean not reported).  South-Asian Canadian parents 62% female (M = 41.2 years, range not reported).
Mandel et al., (1995).	USA.	To study impact of child's race, age, and SES on professionals' judgments about placement decisions.  Cross sectional survey: validated, randomised (via condition) vignettes + validated tool designed for the study.	Police officers (n = 47); social workers (n = 34).  Recruited from a Milwaukee workshop on the investigation of CA and neglect.	Not reported.	Not collected.	Black, White.	Police officers (M = 39 years) were 78.7% male.  Social workers (M = 37 years) were 82.4% female.

Oakley et al., (2017).	UK.	To explore frontline practitioner and community group awareness and understanding of CA linked to faith or belief (CALFB), and identify support and training requirements.  Cross sectional survey: non-validated tool designed for the study.	Total sample (n = 1361). Recruited using an online survey distributed via membership organisations of the National Working Group for Action on CALFB and social media.  Professionals: social work (n = 91), teaching (n = 156), counselling (n = 79), police (n = 318), medicine (n = 60), 'other' (n = 219). Faith (n = 771) and community organisations (n = 143).	94%.	Not collected.	Not applicable.	Not collected.
Pećnik & Brunnberg (2005).	Sweden and Croatia.	To study impact of professionals' characteristics (country of residence, age, gender, parenthood status, and professional experience) on responses to alleged CA (neglect, corporal punishment, and physical abuse).  Cross sectional survey: validated, randomised vignettes + non-validated tool designed for the study.	Social workers attending meetings or training sessions at social welfare centres.  Croatian (n = 87), from 19 centres throughout the country.  Swedish (n = 72), from 15 centres in 11 cities and villages.	Croatian 88%.  Swedish 99%.	Not reported.	Not applicable.	Total sample 90.6% female.  Croatian sample 93% female, aged 27 - 62 (M = 41.8).  Swedish sample 87% female, aged 24 - 64 (M = 42).
Sieracki et al., (2015).	USA.	To study impact of race, SES of the foster care environment, and System of Care services on social workers' placement decisions.  Cross sectional survey: validated, randomised (via condition) vignettes + The Child and Adolescent Needs and Strengths (CANS) + non-validated tool designed for the study.	Social workers (n= 229). Original sample (n = 1000).  Recruited via random sampling drawn from mailing list of the Illinois National Association of Social Workers.	23.5%.	European-American (85.4%), African-American (5.3%), Biracial/Multiracial (3.1%), Latino (2.7%), Asian-American (1.8%), Native American (.4), not reported (.8%).	African-American, White.	86.8% female. Aged 24 - 80 (M= 50.6).

Turbett, & O'Toole, (1983).	USA.	<p>To study impact of case characteristics (ethnicity, SES, and severity) on responses to CA (physical abuse) and compare teachers' responses in this study to nurses' and doctors' responses from previous studies.</p> <p>Cross sectional survey: non-validated, randomised vignettes + non-validated tool designed for the study.</p>	<p>Teachers (n = 91). Recruited from one elementary and one junior high school.</p> <p>Nurses (n = 178). Recruited from 19 hospitals.</p> <p>Doctors (n = 76). Recruited from 3 hospitals.</p>	<p>Teachers 95%.</p> <p>Nurses' and doctors' rates not reported.</p>	Not reported.	Black, White.	Not reported.
Zellman, G. (1992).	USA.	<p>To study impact of case and personal characteristics (including race/ethnicity, gender, SES, severity, and risk) on professionals' judgments and reporting intentions about CA (physical and sexual abuse, and neglect).</p> <p>Cross sectional survey: non-validated vignettes (computer program used to assign a unique combination appropriate to each respondent's profession) + non-validated tool designed for the study.</p>	<p>Total sample (n = 1196). Recruited via nationwide stratified random sampling (from professional organisation directories and commercial lists of schools and child care centres), yielding 15 states (in which reports must be made to the police), stratified by size and affluence.</p> <p>Professionals: school principals (n = 267), paediatricians (n = 257), psychologists (n = 210), social workers (n = 230), and general and family practitioners (n = 104), child psychiatrists (n = 103), and child care providers (n = 126). Psychologists and social workers who did not see children were excluded, eliminating c. 20%.</p>	59% (38% for general and family practitioners; at least 55% for other groups).	Not collected.	White, Black, Hispanic.	Not collected.

This review is about the role of cultural factors in Western-based professionals' decisions about child maltreatment. Of the 16 included studies, all were cross-sectional, 13 of which utilised vignette methods. Thirteen of the studies (76.5%) sampled participants from the US, (Bottoms, Shaver, Goodman, & Qin, 1995; Britner, & Mossler, 2002; Egu, & Weis, 2003; Hansen et al., 1997; Jackson & Nuttall, 1994; Kenny, 2001; Kesner, Kwon, & Lim, 2016; Kesner & Stenhouse, 2018; Levi & Crowell, 2011; Mandel, Lehman, & Yuille, 1995; Sieracki, Fuller, Leon, Jhe Bai, & Bryant, 2015; Turbett, & O'Toole, 1983; Zellman, 1992). Of the remaining three studies, one sampled participants from Canada (Maiter, 2004), one from Sweden and Croatia (Pećnik & Brunnberg, 2005), and one from the UK (Oakley, Kinmond, Humphreys, & Dioumd, 2017).

Five studies recruited via random sampling of the mailing lists of relevant professional bodies such as the American Psychological Association (Bottoms et al., 1995; Hansen et al., 1997; Jackson & Nuttall, 1994; Sieracki et al., 2015; Zellman, 1992). While three of these studies were nationwide (Bottoms et al., 1995; Jackson & Nuttall, 1994; Zellman, 1992), only one stratified by participant gender (Jackson & Nuttall, 1994), and another by size and affluence of US state (Zellman, 1992). One study also utilised professional bodies' mailing lists, but at least part of the sample was non-random (i.e. in two out of the five localities, social work supervisors allowed social workers time for survey completion, and they collected and returned the completed questionnaires, leading to a higher response rate among social workers) (Britner & Mossler, 2002). Of the non-nationwide studies, two recruited from single

US states, namely Virginia (Britner & Mossler, 2002), and Illinois (Sieracki et al., 2015), and one from four states in the US (Hansen et al., 1997). Response rates for professional participants in these studies ranged from 23.5% to 60% (mean = 43.56%).

The remaining 10 studies used convenience sampling methods in a variety of settings, including workshops (e.g. Kenny, 2001; Pećnik & Brunnberg, 2005), a conference (Levi et al., 2011), and online (Oakley et al., 2017). Response rates for professional participants in these studies ranged from 69% to 100% (mean = 92.06%). The one study that used parent participants did not report response rates. One other study also did not report them. Response rates for the total sample of 16 studies ranged from 23.5% to 100% (mean = 71.28%).

### **Study samples.**

Eleven studies sampled practicing health and social care professionals including psychologists, psychiatrists, doctors, and nurses. Three of these sampled social workers only (Jackson & Nuttall, 1994; Maiter, 2004; Sieracki et al., 2015), and two included other groups involved in child protection such as police, judges, and community and faith members (Britner & Mossler, 2002; Oakley et al., 2017). Four studies sampled teachers only (Egu & Weis, 2003; Kenny, 2001; Kesner et al., 2016; Kesner & Stenhouse, 2018), two of which sampled preservice teachers, included in the review because they complete supervised teaching and therefore have a statutory duty to report child maltreatment (Kesner et al., 2016; Kesner &

Stenhouse, 2018). Finally, one study sampled both teachers and health care staff (Turbett & O'Toole, 1983).

The sample sizes for professionals varied considerably, ranging from 6939 participants (from an original sample of 19,275) to 35 (median 184.50, IQR 338.25). Of the 16 studies, five sampled fewer than 100 professionals, six sampled between 100 and 300, three between 300 and 600, one had between 1000 to 2000 professionals, and the largest sampled just under 7000.

Seven studies did not report the race or ethnicity of the sample. Of the nine that did, four had a majority sample of between 85% to 95% White participants (Britner & Mossler, 2002; Hansen et al., 1997; Levi et al., 2011; Sieracki et al., 2015), followed by small proportions of African-American and Hispanic participants. Three had majority samples of around 40% African-American (Kenny, 2001; Kesner et al., 2016; Kesner & Stenhouse, 2018), followed by substantial proportions of White (33 – 38%) and Asian (20 – 24%) participants. Only one had an equal proportion of White, Black, and Hispanic participants (Egu & Weis, 2003). Samples also included small proportions of Native American (0.4 – 3.5%) and multiracial (3.1%) professionals.

Four studies did not report the gender of the sample. Of the 12 studies that did, nine had a majority of female participants (64% – 96%), and three had about equal proportions of males and females (Hansen et al., 1997; Jackson & Nuttall, 1994; Mandel et al., 1995).

### **Measurement of cultural factors.**

Direct measures of cultural factors occur when the assessed feature is asked about specifically, while indirect measures require extrapolation from collected data (De Houwer & Moors, 2010).

#### ***Direct measures.***

A few studies directly measured cultural factors. Two used self-completed surveys specifically designed to ask about religion/faith-related child abuse. Bottoms et al. (1995) collected data on practitioners' experiences of faith-related cases, including frequency, nature, and outcome. Oakley et al. (2017) assessed the extent of professional and community group awareness and understanding of faith-related cases, as well as their training needs.

Levi and Crowell (2011) measured associations between race and professionals' understanding of 'reasonable suspicion'. Using the Differential Diagnosis Scale (DDS) to assess professionals' thresholds for what constitutes 'reasonable suspicion', and the Estimated Probability Scale (EPS) to ask them to correlate 'reasonable suspicion' with a numerical probability that abuse occurred.

#### ***Indirect measures.***

Vignettes, in the form of hypothetical written scenarios, are indirect measures that can infer cultural bias in diagnosis and recommended treatment (Paradies et al., 2013). Thirteen studies in this review utilised this method. Ten studies used identical

vignettes in which the race/ethnicity of the child/family was changed from African-American/Black, to White, or Hispanic, in order to measure the impact of case characteristics on professionals' attitudes and decisions (Britner, & Mossler, 2002; Egu & Weis, 2003; Hansen et al., 1997; Jackson & Nuttall, 1994; Maiter, 2004; Mandel et al., 1995; Pećnik & Brunnberg, 2005; Sieracki et al., 2015; Turbett, & O'Toole, 1983; Zellman, 1992).

***Studies utilising both direct and indirect measures.***

In conjunction with vignettes, Kenny (2001) used the Teachers and Child Abuse Questionnaire to measure teachers' competence in identifying and assessing child abuse, knowledge of reporting procedures, and number of actual cases encountered. The study also used the 12-item Educators and Child Abuse Questionnaire (ECAQ) to assess teachers' knowledge and competence in identifying child maltreatment, their knowledge of reporting procedures, and their overall attitudes towards corporal punishment. Kesner et al. (2016) and Kesner and Stenhouse (2018) investigated the impact of teachers' ethnicity on their interpretation of vignettes. Alongside vignettes, they also used the ECAQ, as well as the 29-item Corporal Punishment Scale (CPS) to assess attitudes toward corporal discipline specifically, and the Demographics and Hollingshead Four-Factor Index of Socioeconomic Status (HFFISS) to assess teachers' SES. Kesner and Stenhouse (2018) also used the non-validated Corporal Punishment Experiences (CPE) to assess preservice teachers' own experiences of corporal punishment. Finally, Sieracki et al., 2015 used a portion of

the Child and Adolescent Needs and Strengths (CANS) to determine social workers' assessment of the needs and strengths of the hypothetical child in the vignette.

### **Role of cultural factors in professionals' decisions about child abuse.**

#### ***Faith-related cases.***

Two studies investigated faith-related variables. Bottoms et al. (1995) collected data on 1652 cases of child abuse. They found that 94% of abuse was perpetrated by religious professionals, mostly Catholic, fundamentalist, or Protestant. 48% of cases involved ridding-evil, and 23% of neglect cases involved sexual abuse. They found that clinicians' belief in allegations was high at 1.96 on a scale of 0 not true to 2 true, with no significant differences found for the three case types (medical neglect, ridding evil, and religious authority). Social services were most likely to investigate ridding-evil or medical neglect cases, and more likely to investigate child (59%) versus adult cases (8%), while cases of medical neglect were unlikely to be prosecuted.

Oakley et al. (2017) found that while 75% of professionals had heard of the term child abuse linked to faith or belief (CALFB), only 33% were confident they would be able to identify indicators of this form of abuse, only 52% were confident they knew how to respond, and 74% had received no specific training on CALFB. Witchcraft and spirit possession were the most prevalent answer to the request for a definition of CALFB (n = 109). However, issues such as FGM (n = 22), honour based violence (n = 15), excessive physical punishment (n = 10), and medical neglect (n = 4) were also

included. Professionals disagreed about whether CALFB is a specific form of child abuse, requiring the need to include considerations of faith and belief, and requested specific training to address definitions and suggest appropriate responses. Many also noted the importance of engaging with faith or belief both at a personal and community level. Similarly, Maiter (2004) found that when judging vignettes about corporal and psychological discipline, and inappropriate supervision, more social workers than South-Asian parents recommended that families should seek help from community and religious leaders.

#### ***Case and perpetrator variables.***

Of the 13 vignette studies, 8 found that race/ethnicity influences the decision-making of professionals, while 5 did not. Hansen et al. (1997) found that case characteristics (over professional characteristics such as training and personal history of maltreatment) had the most impact on ratings. Race impacted outcomes across all forms of maltreatment so that vignettes with an African-American family were often rated as being significantly less severe and less likely to be reported than identical vignettes with a White family. Psychologists and social workers gave vignettes with White persons higher suspicion ratings than vignettes with African-American persons for sexual abuse, neglect, psychological maltreatment, and the total (summed) rating. Vignettes with White persons also received significantly higher reporting ratings for physical abuse, psychological maltreatment, and the total rating. Small to moderate amounts of variance (11.1 to 25.9 percent) were accounted for by the multiple regression equations. In contrast, Egu and Weiss (2003) found that

teachers' evaluations of whether a child was physically abused or should be reported as abused, were not affected by whether the child was White, Black or Hispanic.

Mandel et al. (1995) found a similar pattern to Hansen et al. (1997). When judging vignettes, both police and social workers were less likely to agree with the removal of children who were living in a poor, Black (vs. high SES, White) neighbourhood and were older (11-12 years old vs. 6-7yrs). They explain that because race and SES were manipulated in combination, rather than separately, this result may be due to the interaction of race with age, class, or both.

In contrast, Britner et al. (2002) found that professionals' placement decisions following instances on physical abuse were not significantly influenced by the ethnicity (African-American or White) and age of the child, nor chronicity of abuse. Sieracki et al. (2015) also found that when making placement decisions about community or residential care, social workers were not impacted by whether the child was African-American or White, nor by the SES of the foster care environment, or treatment history. This was in contrast to other clinical and environmental factors such as level of antisocial behaviour and monitoring needs.

Earlier studies found a significant effect in the opposite direction to Hansen (1997) and Mandel et al. (1995). Turbett, & O'Toole (1983) found that although ethnicity (and SES) had little or no relationship to teachers' and nurses' recognition and reporting of child abuse, for doctors, cases involving Black (vs. White) families were more likely to be labelled as abuse. Zellman (1992) also found that particularly in

cases of physical and sexual abuse, incidents involving Black and lower SES families were generally judged to be more serious and more likely to be defined as abuse, and the law was regarded as more clearly requiring a report. In such cases, the outcomes of reports were judged to be better for lower-status families, and in every case professionals were more likely to report them. In contrast, Jackson and Nuttall (1994) found that social workers' judgements about hypothetical sexual abuse allegations were not significantly affected by whether the child or perpetrator were White or from a minority group.

#### **Professionals' variables.**

Kesner et al. (2016) found that African-American (and not Asian) preservice teachers exhibited more favourable attitudes toward corporal punishment than Asian (MD = 8.26,  $p < 0.05$ ), and White (MD = 11.73,  $p < 0.01$ ) preservice teachers. However, African-American participants, did not differ from the other racial groups in terms of how they understand child maltreatment and view their role as mandated reporters. With the exception related to the rating of abusiveness on the most severe discipline technique by Asian participants compared to African-American (MD = 0.82,  $p < 0.01$ ) and White participants (MD = 0.75,  $p < 0.01$ ), the groups were in general agreement as to the severity, appropriateness, and effectiveness of the various discipline techniques used in the vignettes.

Similarly, Kesner & Stenhouse (2018) found that African-American teachers were more accepting of the use of physical punishment as a discipline technique

compared to White teachers, rating some moderate and severe forms of physical discipline as more effective and appropriate compared to White participants. In addition, they experienced corporal punishment more often and with greater severity in childhood. However, with one exception, ratings of abusiveness did not differ between the two groups, suggesting that the African-American participants in this study were able to view a discipline technique as effective, and also judge it as abusive. The authors suggested that this might be due to the interaction between African-American professionals' own cultural backgrounds with the more dominant Anglo middle-class normative views of the US. However, an earlier study looking at professionals' actual reporting behaviour (Kenny, 2001), found that none of the Black teachers in the sample had reported abuse and only 11.76% had assisted in abuse reports, compared to 31.51% of White and 27.58% of Hispanic teachers who had reported abuse, and 41.10% of White and 36.21% of Hispanic teachers who had assisted in reports.

Pećnik and Brunnberg (2005) found that Croatian social workers (working in Croatia) were more likely than Swedish social workers (working in Sweden) to consider child protection necessary in the middle stages of the hypothetical case (rather than the beginning and final stages; each stage of the vignette contained increasing amounts of information and information of escalating severity). Croatian social workers were also more likely to define the problem in terms of child abuse and consider compulsory removal of a child to be more appropriate. Country of residence accounted for 3.9% and 5.1% variance of judgements.

Other studies did not find significant main effects for cultural factors concerning professionals' variables. Levi and Crowell (2011) found that while there was significant variability in doctors' ratings of what might constitute 'reasonable suspicion' of abuse and therefore trigger their reporting duties, on both the scales in their study, ratings were not influenced by doctors' race and other variables such as their age, gender, and seniority. Egu and Weiss (2003) found that while teachers' evaluations of whether a child was abused or should be reported as abused were influenced by the severity of the abuse, they were not influenced by teachers' own race, nor interactions with the race of the child in the vignettes. Finally, in a small study of 14 social workers and 21 South-Asian Canadian parents, Maiter (2004) found that practitioners did not differ from parents in their judgements about whether different scenarios constituted child abuse, including judgements about corporal discipline practices.

### **Study quality.**

Table 2 summarises each of the studies' limitations and provides their quality assessment ratings. Study quality was assessed in relation to the relevance and clarity of aims, suitability and rigour of design and analysis, including risk of bias, and applicability of results (PHWO, 2014). Nine of the 16 studies were of moderate quality (Bottoms et al., 1995; Britner, & Mossler, 2002; Hansen et al., 1997; Kenny, 2001; Kesner et al., 2016; Kesner & Stenhouse, 2018; Levi et al., 2011; Mandel et al., 1995; Turbett, & O'Toole, 1983). Five studies were of moderate to high quality (Egu & Weis, 2003; Jackson & Nuttall, 1994; Pećnik & Brunnberg, 2005; Sieracki et al.,

2015; Zellman, 1992). While one was of low to moderate quality (Oakley et al., 2017), and one was of low quality (Maiter, 2004).

As all studies were cross-sectional, causal inference is limited (Paradies et al., 2013). There were a number of major methodological limitations to the studies. Maiter (2004), and Kesner and Stenhouse (2018), had small sample sizes of 35 and 51 respectively. Two studies did not report response rates (Maiter, 2004; Mandel et al., 1995). Of those that did, Sieracki et al. (2015) had a low response rate of 23.5%, whereas Bottoms et al. (1995), Britner and Mossler (2002), and Jackson and Nuttall (1994) had response rates of around 40%, and Hansen et al. (1997) and Zellman (1992) had response rates of circa 60%. Only six studies had response rates of over 90% (Egu & Weis, 2003; Kenny, 2001; Kesner et al., 2016; Kesner & Stenhouse, 2018; Oakley et al., 2017; Pećnik & Brunberg, 2005; Turbett & O'Toole, 1983). Studies also had non-representative samples due to having low proportions (6.5% - 15%) of non-White professionals (Hansen et al., 1997; Levi & Crowell, 2011; Sieracki et al., 2015), low proportions (5 - 20%) of male participants (Kenny, 2001; Kesner et al., 2016; Kesner & Stenhouse, 2018; Sieracki et al., 2015), and recruiting from a single workshop (Mandel et al., 1995). The samples of two studies (Kesner et al., 2016; Kesner & Stenhouse, 2018) also consisted of professionals-in-training recruited from single classes.

**Table 2. Limitations and quality appraisal of the review studies.**

Study	Limitations	Evidence quality (score out of 9)
Bottoms et al., (1995).	Retrospective design. Not all confounding and bias considered (no analysis of race, ethnicity, gender, or other variables such as SES). 37% response rate. Power analysis not done. Non-standardised survey.	6 (Moderate)
Britner & Mossler, (2002).	Partially non-random, self-selected sample. Not all confounding and bias considered (sex and years of education confounded with group membership, e.g. majority of judges and guardians were male, whereas majority of other professional groups were female). 60% response rate. A non-standardised survey was used. Low proportion of judges, guardians, and child advocates.	6 (Moderate)
Egu & Weis, (2003).	Not all confounding and bias considered (do not control for SES or history of abuse). Possible social desirability influence.	7.5 (Moderate - High)
Hansen et al., (1997).	40.5% to 43.4% response rate. Low proportion of non-White professionals. Power analysis not done. Non-standardised survey.	6.5 (Moderate)
Jackson & Nuttall, (1994).	40% response rate. Female gender response bias. Presentation of vignettes not randomised. Power analysis not done. A non-standardised survey was used.	7 (Moderate - High)
Kenny, (2001).	Non-random, self-selected sample. Not all confounding and bias considered (do not control for SES or history of abuse). Responses to vignettes did not consider impact of race. Presentation of vignettes not randomised. Female gender response bias. Power analysis not done. Possible social desirability influence.	6.5 (Moderate)
Kesner et al., (2016).	Non-random, self-selected sample. Non-representative sample (recruited from one class). Not all confounding and bias considered (do not include measures of the participants' actual experiences with corporal punishment). Female gender response bias. Participants are preservice teachers not yet fully in practice. Power analysis not done. Possible social desirability influence.	6 (Moderate)
Kesner & Stenhouse, (2018).	Non-random, self-selected sample. Non-representative sample (recruited from one class). Small sample size. Presentation of vignettes was not randomised. Female gender response bias. Small effect sizes. Participants are preservice teachers not yet fully in practice. Possible social desirability influence. Power analysis not done.	6 (Moderate)
Levi & Crowell, (2011).	Non-random, self-selected sample. Not all confounding and bias considered (do not control for SES or history of abuse). Low proportion of non-White professionals. Power analysis not done. Possible social desirability influence.	6.5 (Moderate)

Maiter, (2004).	Non-random and small sample size. Not all confounding and bias considered (do not control for SES). Significance values not given. Limited statistical analysis. Power analysis not done. Possible social desirability influence.	2.5 (Low)
Mandel et al., (1995).	Non-random, self-selected sample. Non-representative sample (recruited from one workshop). Participants were only given one vignette each. Male response bias for police officers and female response bias for social workers. Not all confounding and bias considered (no analysis of professionals' race, ethnicity, SES, history of abuse.) Power analysis not done. Possible social desirability influence.	6 (Moderate)
Oakley et al., (2017).	Self-selected sample. Not all confounding and bias considered (no analysis of gender, age, race, ethnicity, SES, etc.) No statistical analysis. Non-standardised survey.	4 (Low - Moderate)
Pečnik & Brunnberg (2005).	Non-randomised sample. Power analysis not done.	7.5 (Moderate - High)
Sieracki et al., (2015).	23.5% response rate. Not all confounding and bias considered (do not consider effects of professionals' race, gender, history of abuse). Female gender response bias. Low proportion of non-White professionals. Non-standardised survey. Power analysis not done.	7 (Moderate - High)
Turbett, & O'Toole, (1983).	Non-random, non-representative sample (recruited from two schools). Not all confounding and bias considered (do not report gender, age, history of abuse etc.) Presentation of vignettes not randomised. Power analysis not done. Non-standardised survey. Possible social desirability influence.	5 (Moderate)
Zellman, G. (1992).	Not all confounding and bias considered (do not consider effects of professionals' race, gender, history of abuse). 59% response rate. 38% response rate for general and family practitioners. Non-standardised survey.	7.5 (Moderate - High)

Threats to internal validity due to social desirability effects were found for eight studies (Egu & Weis, 2003; Kenny, 2001; Kesner et al., 2016; Kesner & Stenhouse, 2018; Levi & Crowell, 2011; Maiter, 2004; Mandel et al., 1995; Turbett, & O'Toole, 1983). Twelve studies did not control for main confounders such as professionals' gender, SES, history of abuse, and age (e.g. Bottoms et al., 1995; Britner and Mossler, 2002; Egu & Weis, 2003). Nine studies also used non-randomised samples (e.g. Kenny, 2001; Levi & Crowell, 2011; Sieracki et al., 2015), while four failed to randomise the order of vignettes (e.g. Jackson & Nuttall, 1994; Kenny, 2001). Eight

studies used non-validated measures (e.g. Bottoms et al., 1995; Zellman, 1992), and seven utilised non-validated vignettes (e.g. Egu & Weis, 2003; Jackson & Nuttall, 1994). While one study conducted limited statistical analysis (Maiter, 2004), another had none (Oakley et al., 2017), and 12 failed to utilise power analyses (e.g. Hansen et al., 1997; Kesner & Stenhouse, 2018). Thirteen of the 16 studies were conducted in the US, limiting the generalisability of results to other countries.

### **Discussion**

We believe this to be the first systematic review to critically assess the quality of studies determining the manner in which cultural factors influence Western-based professionals' decision-making about child maltreatment. Sixteen studies were identified. While review search terms included several variants for culture, and included the terms nationality, ethnicity, and race, the majority of studies (81.25%) were concerned with race/ethnicity variables of the case, professional, or both. Only one study was concerned with country of residence (Pećnik & Brunnberg, 2005), and two investigated faith-related factors (Bottoms et al., 1995; Oakley et al., 2017). Just over 62% of the review studies found evidence of cultural bias concerning child maltreatment decisions among professionals of different disciplines. However, the extent and nature of this is unclear.

In considering the results of this review, an important caveat should be noted. Culture is broadly understood to be any information acquired via social transmission that is capable of influencing individual behaviour (Richerson & Boyd, 2005, p. 5),

and is comprised of cultural beliefs, values, attitudes and behaviours shared by a particular group of people (Nieto & Bode, 2011). However, its delineation is not without its problems. It is inextricably linked to social constructs such as race and ethnicity, with researchers still sometimes using the terms interchangeably (Pfeffer, 1998, p. 1381). Moreover, classification of its components is also problematic. For example, race has been described as a social construct created to minimize cultural and ethnic diversity for the purposes of devising a classification system that reduces this diversity into static categories (Kesner et al., 2016, p. 320). However, despite the inadequacies of categorising often distinct groups of people with one racial label, demarcations such as those set by the US Census Bureau, continue to be used by researchers investigating the experiences of various racial and cultural groups (Kesner et al., 2016, p. 319). Researchers have also recognised the inherent challenges in attempting to disentangle cultural factors from 'social' factors such as SES or environmental stressors (Betancourt, Green, Carrillo, & Ananeh-Firemong, 2003, p. 294).

Bearing this caveat in mind, the results of this review show that among all the studies, the indirect vignette measurement approach was the most commonly used. Physical abuse, particularly corporal punishment, was the most commonly studied type of abuse. While no particular patterns emerged by country and study population, concerning measurement approach, studies involving teachers were the most likely to use a validated direct measure of cultural factors. The ECAQ was used

in three of the five studies that sampled teachers, followed by the CPS which was used in two studies by the same author.

Of the studies included in this review, 62.50% found evidence of cultural bias concerning child maltreatment decisions among professionals of different disciplines, however the extent and nature of this is unclear. With regards to case variables, some vignette studies found that certain professionals (psychologists, social workers, and police) judged vignettes with a Black family as less severe and less likely to be reported than identical vignettes with a White family. One hypothesis is that a racial bias existed so that professionals assessed maltreatment to be more extreme or non-normative when presented with a White family, but less extreme and more normative when presented with an African-American family (Hansen et al., 1997, p. 329; Mandel et al., 1995, p. 918). Another possibility is that the White-majority samples (over 95% when reported) realised that race was being evaluated and overcompensated by rating vignettes with African-American persons significantly lower (Hansen et al., 1997, p. 329). However, racial bias has also been found in the opposite direction. Some studies found a variety of professionals more likely to judge cases of physical and sexual abuse involving Black families, as abuse, as more serious, and as requiring a report. Moreover, in studies where an effect was found, race has sometimes been conflated with other case characteristics such as SES, and only small to moderate amounts of variance have been assigned to the studies' variables, leading to inconclusive study results.

Regarding professional variables, while some studies have found no relationship between the race/ethnicity of professionals and their decision-making, others did find a link. Croatian social workers have been found to be more likely than Swedish social workers to judge identical cases of different types of maltreatment as abuse and consider compulsory removal of the child, suggesting that cultural/social welfare system differences between the two countries may lead to distinct clinical outcomes. However, country of residence accounted for only small amounts of variance, with the best predictors being case specific assessments such as visible signs of abuse. Moreover, compared to Asian and White preservice teachers, African-American teachers have been found to be more accepting of corporal punishment, and to have experienced it more often in their own childhoods. Nevertheless, they did not differ from the other racial groups in terms of their ratings of abusiveness, and how they understand child maltreatment and their role as mandated reporters. Although in a study investigating their actual reporting behaviour, it was found that none of the Black teachers had reported abuse, compared to 31.51% of White and 27.58% of Hispanic teachers, possibly indicating a reluctance to report abuse as a result of cultural norms (Kenny, 2001, p. 88).

Studies investigating religion/faith-related variables found evidence of a wide range of practices, spanning several racial and ethnic groups, that might fall into this category, including excessive corporal punishment, medical neglect, ridding-evil practices, and also FGM and honour based violence. They found that while clinicians' belief in all allegations was high, social services were more likely to investigate

ridding-evil or medical neglect cases, while cases of medical neglect were unlikely to be prosecuted. Moreover, a recent UK study suggests that few professionals feel confident in their ability to identify this form of abuse, and only about half feel that they know how to respond. Since the majority of professionals also reported that they have not received any specific training about this form of child abuse, there appears to be a continued need for multi-agency education and training around faith-related child abuse.

Findings of this review have substantial implications for child protection, and highlight an ongoing need to recognise and mitigate cultural bias among health and social care professionals and teachers. As Western nations continue to grow in cultural diversity, clinicians will be increasingly required to be sensitive to culturally related issues (Terao et al., 2001, p. 158), while ensuring that the children and families in such cases are not discriminated against. As with related issues such as provider racial discrimination, a critical preliminary to this task is to cultivate a more rigorous and systematic approach to monitoring cultural sensitivity and bias among professionals, and develop multi-strategy, evidence-based approaches to managing cultural bias (Paradies et al., 2013, p. 383).

The field of “cultural competence” has already emerged to promote a health care system that acknowledges and integrates the importance of culture, develops cultural knowledge, is vigilant of culture-related differences in treatment and outcomes, and adapts services to meet culturally unique needs (Betancourt et al.,

2003, p. 294). In line with existing literature (e.g. Whaley & Davis, 2007; Zurynski, Sureshkumar, Phu, & Elliott, 2015), this review suggests that professionals' education, training, and supervision has a continued need to include cultural competence elements covering a plethora of practices including FGM, corporal punishment, ridding-evil, and refusal of blood transfusions and other medical interventions, and spanning several delineations of populations including Africans from FGM-practicing countries, Catholics, Jehovah's Witnesses, and so on. Moreover, while research suggests that there is an issue to attend to, how to do this remains unclear. Evidence that cultural adaptations result in improved outcomes is currently limited, and in some instances, cultural adaptation may reduce the benefits of a program if essential elements are discarded (Kirmayer, 2012, p. 160). Culturally competent practice must therefore be based on research on what is clinically effective as well as the wider social impact of reforms.

Studies included in this review were mostly conducted in the US which must organise cultural difference in specific ways that reflect American history, demography, and politics (Kirmayer, 2012, p. 160). This review also highlights that research on potential cultural bias towards non-Black/African-American families is limited. Moreover, most of the studies in the review had a majority sample of White professionals. Further research is required to examine and compare cultural bias in countries outside of the US and outside of the West, and to expand the assessment of professionals' cultural bias among other groups. This review also raises questions of how cultural biases may interact with related constructs. The professional samples

used were primarily female. This might resemble the gender characteristics of certain populations of professionals such as social workers (Pećnik & Brunnberg, 2005, p. 137). However, because some of the review studies found that females were more likely to believe and report allegations (e.g. Jackson & Nuttall, 1994; Kenny, 2001), it is possible that the results represented in this review were skewed by the gender imbalance. Future research should try to address this by ensuring more gender balanced samples. Since studies in the review struggled to disentangle the impact of cultural factors from characteristics such as gender, age, SES, national origin, and religion, and since bias can exist on multiple social dimensions, particularly affecting children and families with multiple minority identities (Hall et al., 2015, p. e74), further more nuanced research is needed to better understand the impact of these variables on professionals' decisions.

Vignettes compromise external validity since real-life situations usually contain a lot more information and nuance on which professionals can judge their decisions (Mandel et al., p. 919). However, they allow for more systematic manipulation of variables (Hansen et al., 1997, p. 330). Professionals may also respond differently to vignettes than to real-life clinical encounters (Paradies et al., 2013, p. 383). As only four studies made use of both direct and indirect measurement methods, and both have their limitations, future research could manage this by utilising both methods in the same study (Paradies et al., 2013, p. 383). Self-completed surveys as well as vignettes, are also subject to a number of biases, including social desirability, particularly if participants have an inclination of study aims. Future studies could try

to minimise social desirability effects. For example, computer-based speeded self-report tasks have been used to assess 'gut reactions' (Ranganath, Smith, & Nosek, 2008). De Cao & Lutz (2015, p. 3) have also pioneered the use of list experiments to research attitudes on FGM, asking sensitive questions indirectly in an attempt to elicit more truthful answers from respondents.

A number of limitations to the present review should also be noted. As non-English and grey literature texts such as unpublished, non-peer-reviewed journals, and theses were excluded from the review, it is possible that the review suffers from both publication bias and language bias. In one study meta-analyses of published trials overvalued an effect by 12% compared with those including grey literature (McAuley, Pham, Tugwell, & Moher, 2000). Fortunately, the impact of language bias on review conclusions is thought to be minimal (Wright, Brand, Dunn, & Spindler, 2007, p. 25). Although inter-rater agreement of study selection was assessed through verification by a supervisor, both the screening and quality appraisal of studies were conducted by one reviewer, so it is also possible that this introduced bias (Wright et al., 2007, p. 26).

### **Conclusion**

This is the first systematic review to consider the role cultural factors play in Western-based professionals' decisions about child maltreatment. Despite no date restrictions, only 16 studies were identified between the years 1983 and 2018, suggesting that as vast an area as it is, it remains under-researched. This review

provides some evidence that professionals' cultural bias exists. However, the exact nature of these biases and the manner in which they interact with other individual and group characteristics remains unclear. A key task for child protection researchers is to develop more rigorous, more uniform, and more insightful approaches to monitoring child maltreatment, as well as professionals' cultural sensitivity and bias when responding. This review highlights the need to continue promoting practitioners' cultural competence so that professionals are trained and supported to recognise, acknowledge, and where appropriate, mitigate cultural biases. Migration trends suggest that services and the families they support will be increasingly multicultural and ethnically diverse. Continued research is needed to determine knowledge gaps, and requirements for training and resources.

**The Influence of Culture and Socioeconomic Status on the Mandatory Reporting of Female Genital Mutilation (FGM) by Health and Social Care Professionals and Teachers**

## Abstract

Female Genital Mutilation (FGM) is described as a traditional and cultural practice of a number of countries in Africa, Asia, and the Middle East. It comprises various procedures which alter or injure the external female genital organs for non-therapeutic reasons, potentially resulting in damage to both physical and mental health.

Increasingly considered as child abuse, in 2015, the United Kingdom introduced the FGM mandatory reporting duty, requiring all health and social care professionals and teachers to report *known* cases of FGM (i.e. where a girl discloses she has undergone FGM) in under 18-year-olds; arguably in order to break down professionals' concerns regarding cultural sensitivity.

Research suggests that cultural factors and social factors such as socioeconomic status (SES) are some of the many variables influencing professionals' decisions to report child abuse. This research employs an experimental design using hypothetical case scenarios (hereafter vignettes) to examine how cultural sensitivity and SES may influence professionals' decision to report FGM. Professionals' variables and their relationship to reporting behaviour were controlled for and examined for exploratory purposes. Both direct questioning and indirect questioning techniques (to reduce social desirability bias) were employed.

Results indicated that whether asked indirectly or directly, the majority of professionals sampled said that they would report a *known* case of FGM in just under 80% of instances. Moreover, direct questioning suggested that professionals were slightly less likely to report to the police when the family was described as being both well acculturated and of high SES. Implications for professionals' practice and training, and for future research and policy directions are discussed.

## Introduction

### **FGM: Definition and Prevalence**

Female genital mutilation (FGM), female circumcision, and female genital cutting are terms used to refer to all procedures which alter or cause injury to the female genital organs for non-therapeutic purposes (Dustin, 2010, p. 8). Since FGM is the term used in current United Kingdom (UK) discourse (Dustin, 2010, p. 20), it will be used throughout this review.

FGM has been given the following typology by the World Health Organization (see Earp, 2015, p. 91):

FGM Type 1 or clitoridectomy refers to the partial or total removal of the clitoral glans (the part of the clitoris visible to the naked eye) or prepuce (“hood”). It is the most common form.

FGM Type 2 or excision refers to the partial or total removal of the external clitoral glans and/or hood, and/or the labia minora, with or without removal of the labia majora.

FGM Type 3 or infibulation, the most extreme form of FGM and also one of the rarest, refers to the narrowing of the vaginal opening (leaving a small opening for the

flow of urine and menses) via a seal created from cutting and stitching the labia minora and/or the labia majora, with or without excision of the external clitoris.

Finally, FGM Type 4 refers to all other harmful procedures to the female genitalia, including the Muslim practice of “Sunna” which involves the pricking or nicking of the clitoris while leaving the body of the organ intact.

It is estimated that over 125 million girls and women worldwide are currently living with the consequences of FGM, and that according to current trends another 30 million are at risk of being cut over the next decade (United Nations Children’s Fund, 2013). Prevalence data for FGM in the UK and elsewhere in the West is difficult to ascertain. Estimates from Home Office migration figures suggest that in 2011, 137,000 FGM survivors lived in England and Wales (Macfarlane & Dorkenoo, 2015, p. 5). Furthermore, 60,000 girls aged under 14 years were born to mothers who had undergone the procedure (deemed a risk factor for FGM) (Macfarlane & Dorkenoo, 2015, p. 6).

However, risk figures often build on the assumption that girls of parents who were born in FGM-practising countries are at risk of being subjected to FGM (Johnsdotter, 2019, p. 2). A growing number of studies show that migration appears to be a key catalyst for changes in attitudes and practices regarding circumcision of girls among Africans in Europe (Johnsdotter, 2019, p. 2; e.g. Vogt et al., 2017). Increasingly, a number of researchers and FGM-practitioners are arguing that assumptions about

large-scale illegal FGM activities in Europe may lack substantiation (e.g. Creighton, Zimran, Otoo-Oyortey, & Hodes, 2019; Johnsdotter, 2019). Therefore, it is important to be cautious about these numbers.

### **FGM: Physical and Psychological Consequences**

Although reports vary, it is thought that FGM is mostly performed on girls aged from 1 week to 17 years, most commonly around puberty (Barstow, 1999, p. 503). 'Cutters' are usually older women of the community. Available instruments such as razor blades are used, and in some instances a girl is cut without anaesthesia, sterilisers, pain-relief, or antibiotics (Barstow, 1999, p. 503). Qualified medical professionals might also carry out the procedure (Whitehorn, Ayonrinde, & Maingay, 2002, p. 163).

Due to individual differences among affected women and girls, and differences in the types of FGM, there are also variations in the effects of FGM (Earp, 2015, p. 92). Nevertheless, FGM has a number of possible immediate and long-term harmful consequences. Personal accounts recorded by victims reveal the pain, terror, and humiliation associated with the practice (Barstow, 1999, p. 503). Immediate complications can arise from haemorrhage and infection, post-operative shock, accidental puncture of neighbouring organs such as the urethra, bladder, and vaginal walls, and tetanus and septicaemia from the non-sterilisation of instruments used during the procedure. Deaths resulting from exsanguination following the procedure have also been reported (Barstow, 1999, p. 504). Long-term consequences include

chronic vaginal and uterine infections, keloid scars, painful menstruation, incontinence, nerve damage, infertility, childbirth complications, and an increased risk of HIV infection (Kellner, 1993, p. 119). There is also evidence that FGM can be harmful to sexual function. All variations of FGM damage nerves affecting sexual experiences, and for some women, vaginal intercourse may be difficult and painful (Kellner, 1993, p. 119).

Given the diversity of FGM procedures and the circumstances under which they are practiced, psychological morbidity associated with the practice can be expected to be highly variable (Parikh, Saruchera, & Liao, 2018). Girls and women living in societies where FGM is routine and their experiences therefore normalised might be minimally impacted. However, affected individuals have also reported feelings of betrayal, embarrassment, guilt, anger, shame, and inadequacy, manifesting in a range of psychological problems ranging from anxiety and depression, to chronic pain syndrome, post-traumatic stress disorder, and psychosexual dysfunction (Whitehorn et al., 2002, p. 166).

### **FGM: From Cultural Practice to Child Abuse**

Because FGM has been historically practiced by some communities primarily from Africa, but also Asia and the Middle East, it is described as a cultural practice (Vogt, Efferson, & Fehr, 2017, p. 283). Depending on the individual culture, many explanations have been offered to explain the practice, including as a means for securing better marriage by signalling fidelity and ensuring paternity (Whitehorn et

al., 2002, p. 163). FGM is thus understood as a parental practice powered by the belief that it is in the child's best interests (Costello, Quinn, Tatchell, Jordan, & Neophytou, 2015, p. 1264). Still, since their daughter's and their own status within the community partly depends on the performance of FGM, community members are under considerable pressure to consent (Christou & Fowles, 2015, p. 347).

FGM is understood as a practice that is meant to suppress and control women's sexual behaviour (Kellner, 1993, p. 120). International law recognises FGM as one of the most obvious and severe forms of violence against girls and women, and places an obligation on governments to take steps to prevent it (Christou & Fowles, 2015, p. 349). In the UK, it has been unlawful since 1985, and the law updated in 2003 to include procedures carried out abroad on UK nationals and residents. Calls for it to be treated as child abuse (e.g. Kellner, 1993; Larson, 1996) are also being increasingly heeded (Amasanti, Imcha, & Momoh, 2016). The 2015 Serious Crime Act introduced a mandatory reporting duty requiring health and social care professionals and teachers to report *known* cases of FGM (i.e. where a girl discloses she has undergone FGM) in under 18-year-olds.

### **The Possible Influence of Cultural Sensitivity and Socioeconomic Status (SES) on FGM Reporting**

The current figures available suggest that there remains a gap in the reporting of FGM. The UK's first paediatric FGM service was established in 2014 and a study describing its first year of activity found that 38 children were referred, 18 of which

(47%) were confirmed as having FGM. Three of those 18 cases were identified as illegal under UK law (the remaining 15 were performed before the child entered the UK, and were therefore not illegal under UK law) (Creighton, Dear, de Campos, Williams, & Hodes, 2016). The authors conclude that the number of referrals in the study was minimal in comparison with the numbers expected if daughters of adult women living in the UK with FGM undergo FGM, and that it is unclear whether this is due to a trend to the less invasive (but still criminal) type 4 FGM where medical complications and physical signs are few, that children are being taken out of the UK for FGM, or that FGM among children living in the UK is very uncommon (Creighton et al, p.4). The only previous similar report in the literature is a retrospective study from the same authors describing 48 children seen at a London safeguarding clinic over an 8-year period up to 2014 (Hodes, Armitage, Robinson, et al., 2016). Moreover, UK paediatricians reported 61 confirmed cases among under 16s to the British Paediatric Surveillance Unit from November 2015 to November 2017, most done prior to UK entry (O'Donnell et al, 2018).

Despite the assumed prevalence of FGM in the UK, the Metropolitan Police was involved with as few as 145 “incidents of concern” relating to FGM between 2008 and 2011 (Amasanti et al, 2016 p. 2). Moreover, a 2015 systematic review of health professionals’ FGM knowledge, attitudes and clinical practice found 10 studies confirming that not only do health professionals working in high income countries such as the UK, Australia, Italy, and Switzerland care for women and girls with FGM, some have also been approached to perform FGM in babies or young children

(Zurynski, Sureshkumar, Phu, & Elliott, 2015, p. 16). Zurynsk et al.'s review (2015, p. 16) also found that health professionals in the UK (and also Australia, Belgium, New Zealand, and Switzerland) believed that it was likely that some of their patients with FGM had the procedure done in these high income countries despite legislation criminalising FGM.

Reasons for the discrepancy between the assumed prevalence and reporting figures are unclear. One idea is that as society becomes increasingly culturally diverse, clinicians are required to be sensitive to culturally related issues (Terao, Borrego, & Urquiza, 2001, p. 158). They might thus face the challenge of determining the appropriate intervention where a client has committed what is considered a crime in the UK and an acceptable practice in another country (Terao et. al, 2001, p. 159). In the process, by trying to approach FGM in a culturally sensitive manner, there is a danger that they might avoid necessary interventions for fear of being considered racist (Whitehorn et al., 2002, p. 167). Indeed, the introduction of mandatory reporting aimed to break down professionals' concerns regarding cultural sensitivity (Mathers & Rymer, 2015, p. 283).

Zurynsk et al.'s review (2015, p. 16) found that some health professionals did not know about anti-FGM legislation or were unsure what these laws covered and what their obligations were under the laws. While their review was submitted before the UK FGM mandatory reporting duty came into force on 31st October 2015, it is possible that professionals' understanding of the duty suffers from similar

limitations. In an online survey of 1361 UK frontline practitioners and faith and community members, Oakley, Kinmond, Humphreys, & Dioumd (2017) assessed the extent of professional and community group awareness and understanding of child abuse linked to faith or belief (CALFB). They found that while 75% of professionals had heard of the term CALFB, only 33% were confident they would be able to identify this form of abuse, only 52% were confident they knew how to respond, and 74% had received no specific training on CALFB. Moreover, while witchcraft and spirit possession were the most common examples of CALFB given by professionals (8%), less than 2% offered FGM as an example of CALFB.

Research suggests that cultural factors (Terao et al., 2001, p. 160) and social factors such as SES (Lopez, 1989) are some of the many variables influencing a professional's decision to report child abuse. Turbett, & O'Toole (1983) found that although ethnicity had little or no relationship to teachers' and nurses' recognition and reporting of child abuse, for doctors, cases involving Black (vs. White) families were more likely to be labelled as abuse. Using US Department of Health and Human Services data, Hampton and Newberger (1985) examined the effects of a range of case characteristics on the reporting behaviour of hospital staff and found that cases where the child was Black or relatively less affluent were more likely to be reported. Zellman (1992, p. 70) built on these findings by using a vignette study to investigate variables shown to be important in the child abuse literature. They found that especially in cases of physical and sexual abuse, incidents involving Black and lower SES families were more likely to be considered abuse and reported.

Lopez and Hernandez (1986) surveyed the extent to which 118 clinicians in California considered cultural factors in the evaluation of culturally diverse patients. They found that clinicians were more likely to minimise the severity of a clinical issue considered cultural in nature. Finally, Mtezuka (1996) highlights the risk of minimising abuse when professionals assume that a practice is culturally normative. The studies described above are old, and social contexts and education changes may mean that their findings may not necessarily be representative of current norms. However, they may still be relevant to current trends, and more up-to-date research is required to investigate this.

### **Decision-Making Theories**

Judgments and decisions about child maltreatment can be difficult and complex. They may be based on ambiguous, non-factual, incomplete, and contradictory information (López, Fluke, Benbenishty, & Knorth, 2015, p. 2), and are typically time-pressured. The stakes are high. Decisions can have a large and lasting impact on the children and families involved. However, as we have seen, professionals' judgments and decisions may be subject to a number of biases.

Research on decision-making is a vast field. However, it is argued that the child protection field has struggled to take advantage of the knowledge gains and progress regarding decision-making research (Baumann, Kern, & Fluke, 1997, p. 4). Early theories such as Social Exchange Theory (Homans, 1958) and Attribution Theory

(Kelly, 1973) championed the notion of rational thought and rational decision-making, whereby humans calculated the costs and benefits of various options before making a decision, and weighed situational and personal forces before attributing a cause to someone's actions. However, later theories were based on research showing that reason had its limits. For example, Tversky and Kahneman (1974) provided a number of demonstrations of certain types of errors in decision-making, suggesting that under conditions of uncertainty, humans applied a number of heuristics (mental strategies that speed decision-making), that often led to error. Whether the use of heuristics is as error prone as had been previously thought is a matter of continued debate (Kahneman and Tversky, 1996).

While authors have highlighted the lack of theoretical developments that operationalise the decision-making processes in child protection (López et al., 2015, p. 6), there are a few noteworthy models that have been developed from the decision-making literature. In both Munro's (2005) systems approach model and the Decision-Making Ecology (DME) model (Baumann et al., 1997), human error is seen as the starting point for understanding professionals' decision-making. Munro's model incorporates individual factors such as skills and knowledge; resources and restrictions such as analytic versus intuitive judgment, with organisational contexts such as changes in thresholds. The DME model takes account of case factors such as ethnicity and SES, decision-maker or professional factors, organisational factors such as nature of caseload (e.g. proportion of ethnic minority families in a professional's caseload, time, and training) and external factors such as community resources. The

model suggests that understanding the source and extent of errors allows for policy and resources to be better targeted. For example, bias in reporting and placement decisions might be ameliorated by developing caseworkers' skills, particularly those involving cultural awareness (Baumann et al., 1997, p. 6).

Similarly, the model of Judgments and Decisions Processes in Context (JUDPIC) (Benbenishty & Arad-Davidson, 2012) posits that professionals' decisions are based on case characteristics of the child and family such as ethnicity, poverty, and risk; professionals' characteristics (e.g. personal experiences of abuse); and organisational features (e.g. placement policies and guidelines). Judgments and decisions are moderated by a large number of factors, such as policies on the threshold for child placements or about the benefits of maintaining the family unit; the availability of evidence, knowledge, and resources; and the wider contexts of the service system, including cultural contexts such as public attitudes toward the protective system or child welfare legislation. The authors argue that training focusing primarily on assessment of case characteristics is too narrow, and that it is important to make professionals aware of the other domains influencing their decisions, including the global, national, and cultural contexts in which they operate (Benbenishty et al., 2015 p. 64).

The two context-sensitive models of DME and JUDPIC have generated some empirical support for the theories they describe (López et al., 2015, p. 6). Whereas the applications of the DME model help illustrate the intersection between cultural

and organizational contexts, the use of the JUDPIC has illustrated the importance of the country-level context in professionals' decision-making (López et al., 2015, p. 7).

The DME model has been empirically tested in a study by Graham, Detlaff, Bauman and Fluke (2015). The study uses the model as a framework to investigate the impact of case, caseworker and organisational factors on professionals' decisions to place children in out-of-home care. Graham et al., (2015) surveyed 1,103 child protection caseworkers in Texas, cross-referencing against administrative records about each worker's demographic details and their caseload. They found that the DME model fit the data well, indicating a complex relationship between the variables. Case factors such as risk, the family's ethnicity, and their SES were found to be of most importance. While individual factors regarding the caseworkers themselves (such as the worker's own race or ethnicity), were not found to directly influence the placement decision.

The model of JUDPIC has also been empirically tested in a vignette study of 828 practitioners in Israel, the Netherlands, Northern Ireland, and Spain (Benbenishty et al., 2015). Participants in the study were presented with a vignette of a case of alleged child maltreatment and were asked to decide whether abuse was substantiated, assess risk, and recommend an intervention using structured instruments. Participants' child welfare attitudes were assessed. The case characteristic of mother's wish about the removal had no impact on judgments and decisions. However, practitioners' child welfare attitudes were significantly

associated with substantiation, risk assessments, and recommendations. Moreover, there were significant country differences on most measures. The authors conclude that the study's findings support most of the predictions of the JUDPIC model. In particular, the significant differences between practitioners from different countries highlights the importance of context in child protection decision making, and the need for professionals' training to alert them of these differences and their impact on practice decisions.

The scarcity of studies that might elucidate the application of both models is stark (López et al., 2015). This highlights that there remains a continued need for the child protection field to devote more effort to empirically understand the context of decisions that are made, the psychological process of decisionmaking, and the sources of errors that are made (Baumann, Dalgleish, Fluke, & Kern, 2011, p. 11).

### **The Present Study**

As already mentioned, the FGM mandatory reporting duty came into effect in 2015, and the reporting incidence has been surprisingly low. Thus far there has been no research to identify factors affecting reporting and non-reporting of FGM (Amasanti et al, 2016, p. 2), including the influence of cultural sensitivity and SES on clinicians' treatment of FGM. Moreover, aside from Oakley et al.'s 2017 UK study, the bulk of the existing literature is based on research conducted in the United States (US), with its specific history of race and cultural relations. The studies also do not untangle cultural factors from SES or race.

This study's primary research question is whether cultural sensitivity and SES influence UK professionals' decision to report FGM. Based on the child abuse literature, we hypothesise that: 1) professionals will be less likely to report FGM where a family is less integrated into UK culture, and 2) where the family is of higher SES. Research participants' characteristics that could influence reporting behaviours are assessed; these include: gender (see Finlayson and Koocher, 1991; Kenny, 2001), training background (e.g. social work vs. medicine; see Tilden et al., 1994; Turbett, & O'Toole, 1983), and amount of training (Kenny, 2001; King, Reece, Bendel, & Patel, 1998).

In order to answer the research question, an analogue vignette study was chosen because it allows for the controlled manipulation of the variables of interest (Hansen et al., 1997, p. 316). A sample of health and social care professionals and teachers were recruited to complete an online survey asking respondents to provide demographic and other pertinent descriptive information (gender, age, profession, level of training, race, parents' country of origin, and hours of FGM training received), followed by four hypothetical case vignettes presented in a list experiment.

A list experiment is used in an attempt to address the impact of socially desirable responding; i.e. the tendency for respondents in self-report studies to present themselves in a favourable manner (Mortel, 2008, p. 41). Participants are often

unwilling or unable to report truthfully on sensitive topics, due to ego-defensive or image management reasons (Fisher, 1993, p.303). The respondent may believe the information they report (self-deception), or may 'fake good' to conform to socially acceptable values, gain social approval, or to avoid criticism (King & Brunner 2000 p.81). This results in data that are systematically biased towards what respondents perceive to be "correct" or socially acceptable (Fisher, 1993, p.303). The phenomenon is known as social desirability bias and has been found to exist in almost all types of self-report measures (Fisher, 1993). It has been found to affect the measurement of personality factors (e.g. Mick, 1996), attitudes (e.g. Fisher, 1993), and self-reported behaviors (e.g., Mensch & Kandel, 1988), and is most likely to occur in responses to socially sensitive questions (King & Brunner, 2000).

An instrument is valid if it accurately measures what it aims to measure (Beanland, Schneider, LoBiondo-Wood, & Haber, 1999). Social desirability response bias therefore affects the validity of a questionnaire (Huang, Liao, & Chang, 1998). Social desirability bias can confound relationships among the variables of interest in a study by suppressing or obscuring relationships among variables or producing artificial relationships between variables (King & Brunner 2000 p.81).

Health related research often covers socially sensitive topics, therefore researchers must "identify situations in which data may be systematically biased toward respondents' perceptions of what is socially acceptable, to determine the extent to which this represents contamination of the data, and to implement the most

appropriate methods of control” (King and Brunner 2000 p.80). In the present study, social pressure (social desirability effects) is expected to play an important role due to the sensitive nature of the topic and the reporting duty being mandatory. The list experiment (or item/unmatched count technique) in the current study was therefore designed to elicit responses to the vignettes in an attempt to measure professionals’ true views. Used mainly in political science to examine voters’ attitudes and attitudes about race, De Cao & Lutz (2015) have pioneered the use of a list experiment for researching attitudes on FGM. The idea being that if a sensitive question is asked indirectly, the respondent is more likely to answer truthfully (De Cao & Lutz, 2015, p. 3).

In sum, the study aimed to determine whether cultural sensitivity and SES influence UK professionals’ decision to report FGM, and to explore the relationships between demographic variables and reporting behaviour. The study was approved by the NHS Health Research Authority (IRAS project ID 247542, Appendix B) and Royal Holloway University’s ethics committee in July 2018. A non-substantial amendment (Appendix C) was granted in October 2018 (following piloting).

## **Method**

### **Design**

A 2 x 2 factorial design was utilised in this study, with two independent variables: cultural integration and SES. Both had two levels each: high and low. Two binary

variables were required to model four vignettes: 'Vignette High Cultural Integration UK' with values 0 (No) or 1 (Yes), and 'Vignette High SES' also with values 0 (No) or 1 (Yes). These indicator (or 'dummy') variables therefore allowed for four combinations of values for the four vignettes.

Participants were randomly assigned to either the control group or the experimental group (between-subjects). Each participant irrespective of group was presented with all four vignettes in a randomly assigned order. Vignettes were presented with a list experiment and also without.

### **Participants**

The data were gathered via convenience sampling. The study utilised an online survey delivered through Qualtrics. Individual services within Camden and Islington NHS Foundation Trust (e.g. iCope Psychological Therapies Services, Early Intervention Services, Sexual Problems Clinic) were contacted by the principal investigator (FJ) both in person and via email to request they forward their team an online link to the study. Professionals were also recruited via the researchers' networks, including on social media.

There were 226 attempts at completing the online survey of professionals' perceptions of hypothetical case scenarios. Twenty-one logons did not result in survey completion (for reasons unknown). Therefore, a total of 205 health and social care professionals and teachers completed the online survey.

Participants who completed the online survey were primarily female (79.51%) and White (77.07%). Participants also reported themselves as Black (8.78%), Asian (6.34%), Mixed (4.88%), and Other (2.93%). As race is too broad a category to infer participants' cultural characteristics (Kesner et al., 2016, p. 324), respondents were also asked to name both their parents' country of birth in order to determine whether they might have links to a FGM practicing country. Of the 205 participants who completed the survey, 11 (5.37%) did not provide this information. The exact reasons for this are unknown. However, Ireland was mistakenly omitted from the list of countries provided to participants so it is possible that some chose to discontinue the survey as a result. Where participants (n=4) contacted the principal investigator (FJ) about this they were given an apology and asked to report themselves as coming from the UK or another European country. For those participants who completed the survey, most had mothers (93.30%) and fathers (90.77%) who were born in non-FGM practicing countries, primarily the UK (mothers 68.04%; fathers 66.15%) but also Germany, Italy, India, and South Africa (2.05% of fathers for each). A minority of participants had a father (9.23%) or mother (6.70%) born in a FGM practicing country, namely Nigeria, Kenya, The Gambia, Iraq, Ghana, and Somalia.

Almost half of respondents were psychologists (47.32%), namely clinical (41.46%), counselling (5.37%), and health (0.49%). For comparison, in 2016 the British Psychological Society's (BPS) UK membership (which is likely to be less ethnically diverse than that of London) was 76.73% female and racially 91.56% White, 6.75%

Other, and 1.69% Black (BPS, 2016). The sample also included doctors (19.02%), social workers and social care workers (10.73% and 1.95% respectively), teachers (7.80%), nurses (4.39%), occupational therapists (3.90%), midwives (2.44%), psychiatrists (0.98%), biomedical scientists (0.98%), and pharmacists (0.49%).

Participants were offered a choice of categories regarding their ages. Reported ages ranged from 16-20 (although participants had to be at least 18 to be eligible) to 66-70. The majority were aged 26-30 (31.71%), 31-35 (24.39%), and 36-40 (14.63%). 7.80% of the remaining sample were aged 41-45, 6.83% were aged 46-50, 6.34% were aged 21-25, 3.41% were aged 51-55, and 3.41% were aged 56-60. The categories 16-20, 61-65, and 66-70 had 0.48% of participants each.

Almost half of professionals had been qualified for 5 or more years (42.44%), while 25.37% were newly qualified (<5yrs) and 32.20% were trainees. The majority of participants (77.07%) had received either no FGM training at all (37.07%) or 1-2 hours only (40%). 13.17% had received half a day of FGM training, followed by 5.85% who received one day, 1.46% who received two days, 0.98% who received 5-6 hours, and one (0.49%) who had received over a week of FGM training.

Due to limitations of the software used for this study it was not possible to stratify participants by gender and race. However, results show that for gender there were 84 female participants in the experimental group and 68 in the control group. For male participants there was an equal distribution with 18 in the experimental group

and 18 in the control group. 17 participants did not complete enough of the survey to be randomised into group condition (which occurred after collection of demographic data and before the presentation of vignettes). Of those 17, 10 were female and 7 male. For race, 88 non-Black participants were in the experimental group, and 84 in the control group. Black participants were the most unequally distributed group, however they constituted only 8.78% of the sample, so this is unlikely to have impacted results. 13 of the Black participants were in the experimental group, and 3 in the control group. Of the 17 who were not randomised, 15 were non-Black and 2 were Black.

## **Materials**

### **Survey section 1: Demographics.**

The first section of the online survey (Appendix D) required participants to complete a demographic questionnaire that asked about their gender, age, race, and parents' country of birth. They were also asked to indicate their profession, how long they had been qualified or if they were a trainee, and whether they had received any FGM training.

**Survey section 2: The Indirect Questioning Method - Vignettes presented with a list experiment.**

### ***The vignettes.***

Four vignettes (Appendix E) were used for this study. These vignettes manipulated two variables: level of integration into UK culture and level of SES. Table 3 describes these variables in further detail. Table 4 outlines the possible vignette combinations.

**Table 3: Independent variables manipulated within the vignettes.**

IV	Level	Indicators
Cultural integration	High	Second generation, English speaking, few ties to Somali community
	Low	First generation, non-English speaking, strong ties to Somali community
SES	High	Wealthy neighbourhood, university educated, banking profession
	Low	Poor neighbourhood, non-university educated, security guard

**Table 4: The possible vignette combinations.**

Vignette	Integration	SES
1	Low	Low
2	High	High
3	High	Low
4	Low	High

Potentially confounding variables were kept the same across all vignettes. The child's age was set at 12 for two reasons: 1) studies have consistently found that professionals are less likely to report cases involving older children (Terao et al., 2001, p. 161); and 2) FGM is generally performed on girls between ages 4 and 12 (Feldman-Jacobs & Clifton, 2010, p. 1). The family was described as coming from Somalia because it has a 98% FGM incidence rate (Cook, 2016, p. 91), and because the UK has the largest and longest-established Somali community in Europe, most of whom live in London (Hammond, 2013, p. 1005). Since the 2015 mandatory

reporting duty requires mandated professionals to report *known* cases of FGM, all vignettes involved the child telling the respondent that she was 'cut during a special celebration'.

### ***The list experiment.***

The list experiment is an indirect questioning technique, which provides respondents some amount of privacy protection via aggregation (Glynn, 2010, p. 2). The idea behind a list experiment is that if a sensitive question is asked indirectly, the respondent is more likely to reveal a truthful answer (De Cao & Lutz, 2015, p. 3). Used mainly in political science to examine voters' attitudes and attitudes about race, the list experiment has also been shown to reduce social desirability bias when researching attitudes on FGM (De Cao & Lutz, 2015). Using the list experiment means that respondents are asked how many of a list of questions apply to them. As long as the entire list does not apply to them, they can be assured that their answer to the sensitive question will remain unknown (Glynn, 2010, p. 2).

Participants were randomly assigned to either the control group or the experimental group (between-subjects). Each participant irrespective of group was presented with all four vignettes in a randomly assigned order. The control group however did not answer any FGM-sensitive items.

The control group respondents received a list of four non-sensitive, yes/no items and were asked how many of the listed items they agree on, but were told they do not

have to state which items (see below). The treatment group respondents instead received the same list of four non-sensitive, yes/no items, plus a FGM-sensitive yes/no item ('are you concerned enough to report this family directly to the police, rather than continuing to monitor them within your service?'). As with the control group, the treatment group respondents were also asked the number of items they agree on, but without saying which (see below).

In the second section of the online survey, each vignette was presented in turn in a randomly assigned order with a list of four non-sensitive yes/no items for the control group, and with a list of four non-sensitive yes/no items **plus** the FGM-sensitive yes/no item for the treatment group. Four separate lists of four non-sensitive yes/no items were developed and piloted. Each vignette was always presented with the same list. Responses to the lists (specifically, the number of yes responses) constituted the dependent variable. The difference in the total number of items between control and treatment group identified the proportion of respondents that agree with the sensitive item (De Cao & Lutz, 2015, p. 3).

### **Survey section 3: The Direct Questioning Method - Vignettes presented with the FGM-sensitive question.**

In the third section, respondents were again presented with the same four case vignettes they were presented with in the second section (in the same randomly assigned order) and asked to directly answer the FGM-sensitive yes/no item for each. The number of yes responses constituted the dependent variable.

Vignettes, whether presented with a list experiment or not, were followed by a statement acknowledging that the question/s might be difficult to answer and that in real life participants would likely attempt to gather additional information before making a reporting decision, but that we would appreciate their providing their most likely answers based on the information provided.

### **Piloting**

Extensive piloting of the vignettes, the FGM-sensitive question, and the list experiment was carried out. Vignettes should provide realistic scenarios that include sufficient detail so that it is reasonable to ask for a response (Hughes, 1998). The four vignettes and the FGM-sensitive question were administered to seven professionals (two clinical psychologists, two doctors, two pharmacists, and one teacher) for the purposes of content-validity. Pilot participants were asked to give feedback on the vignettes and FGM-sensitive question. Based on their feedback, some parts of the vignettes were changed for readability and clarity. For example, the vignettes were changed to state that the child had been 'cut during a special celebration' in order to alert professionals of FGM rather than other types of physical abuse. Clearer indicators of cultural integration and SES were also applied. The FGM-sensitive question underwent several transformations in order to maximise its sensitivity. The original question 'would you report this to the police?' was deemed too unambiguous to be reflective of real-life practice, hence the final choice of 'are

you concerned enough to report this family directly to the police, rather than continuing to monitor them within your service?'.

A research expert was consulted to properly design the list experiment. Variations of the lists were piloted with several laymen and professionals including clinical psychologists, nurses, and teachers. It was advised that each of the four lists should include negatively correlated control items (where a subject who endorses item 1 is unlikely to endorse item 2) and at least one easy-to-endorse and one hard-to-endorse item (Blair & Imai, 2010; Glynn, 2010). Empirical examples (e.g. Kuklinski, Cobb, & Gilens, 1997) often use a 4-items list (Coffman et al., 2013; Kuklinski et al., 1997; McKenzie and Siegel, 2013) and the sensitive item is often last (De Cao & Lutz, 2015, p. 11). Aims were to minimise ceiling and floor effects and response variance, thus ensuring our ability to detect any increase in the number of yes responses when assigned to the sensitive list.

## **Procedure**

The survey tool Qualtrics was utilised. The first two pages of the online survey consisted of an information sheet (Appendix F) and consent form (Appendix G). Participants were required to be over 18 and a health and social care professional or teacher working in the UK. Participation was voluntary and participants were allowed to withdraw from the study at any time without penalty. Consent was obtained by the participant pressing 'Yes, I agree' on the web page. Participants took

the survey online on their own computers (which provided privacy from the researcher) and their responses were anonymous.

The first section of the survey collected demographic information such as age and gender. In the second section automated procedures were utilised (a coding expert was consulted so that a Java script could be embedded into Qualtrics) to randomise participants to either the experimental or control groups, after which they were presented with the vignettes in the list experiment. In the third section the vignettes were presented again but with the FGM-sensitive question only. The last page of the survey contained a debrief sheet (Appendix H) which included information on FGM and FGM training. Participants' responses were automatically logged on to Qualtrics. See Appendix I for a diagram of the study's procedure. The survey was active from November 2018 to April 2019.

### **Data Analysis**

Responses to the list experiment that followed the vignettes (by estimating the overall proportion of respondents that agreed on the FGM-sensitive items; De Cao & Lutz, 2015, p. 10) were analysed using multivariate regression. We further examined responses where the FGM-sensitive item was asked directly, and investigated the relationship between the FGM-sensitive items and the respondents' characteristics. Missing data was dealt with by using both the complete cases method (CC) and the inverse probability weighting method (IPW). Data analysis is further elaborated as follows:

### **Power analysis.**

Power analysis was carried out for the combined relationship between the independent variables and the dependent variables. The analysis was based on a regression resulting in a total of 8 parameters, namely:

- Four vignettes totalling four parameters (i.e. the intercept for the baseline vignette, plus three additional parameters for the three remaining vignettes).
- The control list vs sensitive list, totalling one parameter (i.e. the dummy variable indicating assignment to the sensitive list; the control group doubles the required sample size).
- And three first order interactions, totalling three parameters (i.e. to capture the additional effects of assignment to the sensitive list for the three remaining vignettes).

Power was analysed with small (0.02), medium (0.15), and large (0.35) effect sizes, a significance level of 0.05, and power at 0.8 (Cohen, 1992). Results indicated a sample of 725, 103, and 49 for small, medium, and large effect sizes respectively. It was feasible to aim for a medium effect size and our sample size of 226 was sufficient for this.

### **The Complete Cases Method (CC).**

This is the simplest method for the analysis of incomplete data regression models and involves the analysis of the set of observations with no missing values. This assumes that data is “missing completely at random” so that ‘missingness’ is not related to any factor, known or unknown, in the study (Horton & Kleinman, 2007, p. 2). We can assume this because the order of the vignettes in the study was randomised.

### **Inverse Probability Weighting Method (IPW).**

For this approach a model for the probability of missingness is fit and the inverse of these probabilities are used as weights for the complete cases (Horton & Kleinman, 2007, p. 6). The number of observations is therefore the same as in the corresponding CC regressions. It is anticipated that the calculated weights correct for any possible bias in missingness. Due to randomisation we did not expect any bias related to the vignettes but carried out these analyses in order to confirm the robustness of results.

## **Results**

This study aimed to determine whether cultural sensitivity and SES influence UK professionals’ decision to report FGM.

There were two hypotheses of interest:

1) Professionals would be less likely to report FGM where a family is less integrated into UK culture.

2) Professionals would be less likely to report FGM where the family is of higher SES.

Table 5 displays the average responses to the four vignettes (i.e. the total number of items the participants said yes to) when vignettes were presented within the list experiment and without. Figures are provided for the control and the treatment groups separately.

Demographic variables and their relationship to reporting behaviour were controlled for and examined for exploratory purposes (namely professionals' gender, profession, level of training, hours of FGM training received, race, age, and parents' country of origin).

In order to simplify analyses and retain power, these variables were collapsed into binary categories with values 0 (No) or 1 (Yes) assigned for each. For example, if a participant responded that they were female, they would be assigned the value 1 for Yes. A male participant would be assigned the value 0 for No. Age was thought to correlate with training level so it was left out.

The variable categories (also listed in Tables 6 and 7) are as follows:

1. Intercept: Did participants on average answer yes to questions in the list experiment, yes or no? Did participants on average answer yes to the FGM-sensitive question, yes or no?
2. Gender: Female, yes or no?

3. Profession: Clinical Psychologist, yes or no?
4. Level of Training: Qualified, yes or no?
5. Level of Training: Trainee, yes or no?
6. Hours of FGM training received: Some FGM training received, yes or no?
7. Race: Black, yes or no?
8. Race: Non-White, yes or no?
9. Parents' country of origin: One or more parents from a FGM practicing country, yes or no?
10. Order vignette presentation: Did respondents on average answer yes to more questions as they moved from one vignette to the next, yes or no?
11. High UK cultural integration: Did respondents on average provide more yes responses to the vignettes which described high cultural integration, yes or no?
12. High SES: Did respondents on average provide more yes responses to the vignettes which described high SES, yes or no?
13. FGM sensitive list: Did participants in the experimental condition (i.e. those shown the FGM sensitive list) provide on average more yes responses, yes or no?
14. High UK cultural integration x high SES: Did respondents on average provide more yes responses to the vignettes which described both high cultural integration and high SES, yes or no?
15. High UK cultural integration x FGM sensitive list: Did respondents in the experimental condition (i.e. those shown the FGM sensitive list), on

average provide more yes responses to the vignettes which described high cultural integration, yes or no?

16. High SES x FGM sensitive list: Did respondents in the experimental condition (i.e. those shown the FGM sensitive list), on average provide more yes responses to the vignettes which described high SES, yes or no?

17. High UK cultural integration x high SES x FGM sensitive list: Did respondents in the experimental condition (i.e. those shown the FGM sensitive list), on average provide more yes responses to the vignettes which described both high cultural integration and high SES, yes or no?

## **The List Experiment Analyses**

### **Model 1a.**

For the first model the list experiment unweighted complete cases were analysed (i.e. the number of observations with values for all of the variables in the list experiment). Participants' responses to the list experiment constituted the dependent variable. There were 4 observations per participant so a maximum possible total of 904 complete cases (226 participants x 4 observations). The list experiment models have 725 complete cases.

### **Model 2a.**

Then the same regression was carried out but without the variable relating to whether either of the respondents' parents was born in a FGM practicing country. This is because 5.37% (n = 11) of participants did not provide responses for this variable and we wanted to understand whether their exclusion would impact results.

### **Model 3a.**

The final regression model that was carried in this set was the list experiment unweighted complete cases but without the control variables. This was in order to analyse the list experiment alone.

### **Models 1b, 2b, and 3b.**

As with model 1a, 2a, and 3a, but using IPW instead of CC.

## **The FGM-Sensitive Question Analyses**

### **Model 4a.**

This model analysed the FGM-sensitive unweighted complete cases using a linear probability model (heteroskedastic errors were corrected for by calculating the cluster robust standard errors). Yes/no responses to the FGM-sensitive question about reporting constituted the dependent variable. The FGM-sensitive models have 697 complete cases.

### **Model 5a.**

Again the same regression used for model 4a was carried out, but without the variable relating to whether either of the respondents' parents was born in a FGM practicing country.

### **Model 6a.**

The final regression analysed all of the FGM-sensitive question unweighted complete cases minus the control variables.

### **Model 4b, 5b, and 6b.**

As with 4a, 5a, and 6a, but with IPW rather than CC.

**Table 5: List Experiment and FGM-Sensitive Question Means.**

	List Experiment				FGM-Sensitive Question			
	Control group		Treatment group		Control group		Treatment group	
Vignette	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
Low Cultural Integration x Low SES	2.04	0.29	2.85	0.63	0.74	0.44	0.79	0.41
High Cultural Integration x Low SES	2.08	0.35	2.90	0.72	0.75	0.44	0.75	0.44
Low Cultural Integration x High SES	1.94	0.44	2.65	0.73	0.74	0.44	0.79	0.41
High Cultural Integration x High SES	2.04	0.29	2.84	0.62	0.70	0.46	0.74	0.44

### Results for the List Experiment

See Table 6 (models 1a and 2a) for results for the list experiment. The main findings are as follows:

1. We can see from the intercept value of just above 2 that our list experiment was created successfully. Participants mostly answered yes to 2 of the questions in the list experiment ( $p < 0.001$ ). Deviations from this intercept value allow us to make inferences about our variables of interest.
2. As expected, there was a highly significant effect for randomisation into the sensitive list (experimental condition) so that participants in the experimental condition provided on average 0.79 more yes responses, which we attribute

to their responses to the FGM-sensitive question ( $p < 0.001$ ). In other words, participants shown the sensitive list increased the number of yes responses by 79%.

3. However, this was not conditional on vignette type. There were no significant main effects for vignette type. Asking the FGM-sensitive question in an indirect manner via the list experiment did not significantly impact professionals' responses to any of the vignettes.
4. There were no significant main effects for the control variables, except that Black respondents on average answered yes to 0.30 (30%) more questions in the list experiment ( $p < 0.05$ ) (model 1a). However, this is not a robust result since it disappears in the subsequent model (2a). Moreover, Black respondents constituted only 8.78% of the sample, and we know that 81.25% of them were allocated to the experimental group and therefore had more opportunities to respond yes to questions than those in the control group.
5. When we remove the control variables and analyse only the list experiment (model 3a) we see that the coefficients are similar to the larger models, thus confirming that the experiment was properly designed.

6. Moreover, carrying out analyses with IPW instead of CC (models 1b, 2b, and 3b) did not impact results, further supporting the reliability of our results (see Appendix J).

**Table 6: Results for the List Experiment**

Variable	Model 1a			Model 2a		
	Estimate	Std. Error*	p-value	Estimate	Std. Error*	p-value
Intercept	2.03	0.10	<. 0.001***	2.03	0.10	<. 0.001***
Female	-0.01	0.06	0.89	-0.01	0.06	0.91
Clinical Psychologist	0.03	0.09	0.74	0.03	0.09	0.76
Training: Qualified	0.01	0.09	0.88	0.014	0.09	0.87
Training: Trainee	0.02	0.08	0.85	0.024	0.08	0.75
Training: Some FGM	0.04	0.07	0.55	0.04	0.07	0.57
Black	0.30	0.15	0.04**	0.26	0.14	0.06
Non-White	0.034	0.09	0.72	0.01	0.09	0.88
One/More Parents From FGM Practicing Country	-0.10	0.11	0.36	-	-	-
Order Vignette Presentation	-0.02	0.01	0.19	-0.02	0.01	0.19
High UK Cultural Integration	0.048	0.05	0.33	0.05	0.05	0.33
High SES	-0.09	0.06	0.13	-0.09	0.06	0.13
FGM Sensitive List	0.79	0.07	<. 0.001***	0.80	0.07	<. 0.001***
High UK Cultural Integration x High SES	0.04	0.09	0.61	0.04	0.09	0.61
High UK Cultural Integration x FGM Sensitive List	-0.01	0.08	0.93	-0.01	0.08	0.93
High SES x FGM Sensitive List	-0.10	0.10	0.31	-0.10	0.10	0.31
High UK Cultural Integration x High SES x FGM Sensitive List	0.11	0.14	0.44	0.11	0.14	0.44

\*\*p < 0.05. \*\*\*p < 0.01. \*Standard errors are cluster robust, clustered by participant.

## Results for the FGM-Sensitive Question

See Table 7 (models 4a and 5a). These models consider how participants responded when they were directly asked whether they were concerned enough to report the family to the police. The main findings are as follows:

1. The intercept value of 0.78 shows that on average participants answered yes to the FGM-sensitive question 78% of the time ( $p < 0.001$ ).
2. There was also a very small significant effect for the order in which the vignettes were presented to participants. Respondents on average answered yes to 0.01 (1%) less questions as they moved from one vignette to the next ( $p < 0.05$ ). This was controlled for in our analyses.
3. Interestingly, there was also a significant interaction between high cultural integration and high SES. On average participants provided 0.05 less yes responses to the vignette which combined high cultural integration and high SES. In other words, they were 5% less likely to say they would report the family to the police when the family was described as being both highly integrated into UK culture and affluent ( $p < 0.05$ ). Although not a large effect, this effect is robust since it holds in all of the relevant models (4a, 5a, 6a).

4. When we remove the control variables and analyse only the responses to the FGM-sensitive question (model 6a) we see that the coefficients are again similar to the larger models.
5. Moreover, the analyses are unaffected by the use of IPW instead of CC (models 4b, 5b, and 6b) (see Appendix J).
6. No other effects were significant, including experimental condition.

**Table 7: Results for the FGM-Sensitive Question**

Variable	Model 4a			Model 5a		
	Estimate	Std. Error*	p-value	Estimate	Std. Error*	p-value
Intercept	0.78	0.11	<. 0.001***	0.79	0.11	<. 0.001***
Female	-0.06	0.08	0.43	-0.07	0.08	0.41
Clinical Psychologist	-0.03	0.08	0.71	-0.03	0.08	0.75
Training: Qualified	0.06	0.08	0.44	0.06	0.08	0.45
Training: Trainee	0.04	0.09	0.61	0.034	0.09	0.71
Training: Some FGM	0.01	0.07	0.83	0.01	0.07	0.80
Black	0.03	0.12	0.77	0.09	0.13	0.50
Non-White	-0.10	0.11	0.38	-0.07	0.11	0.53
One/More Parents From FGM Practicing Country	0.13	0.09	0.12	-	-	-
Order Vignette Presentation	-0.01	0.01	0.03 **	-0.01	0.01	0.03 **
High UK Cultural Integration	0.01	0.02	0.51	0.01	0.02	0.51
High SES	0.00	0.01	0.96	0.00	0.02	0.96
FGM Sensitive List	0.06	0.07	0.39	0.06	0.07	0.40
High UK Cultural Integration x High SES	-0.051	0.024	0.03 **	-0.05	0.024	0.03 **
High UK Cultural Integration x FGM Sensitive List	-0.057	0.034	0.09	-0.06	0.034	0.09
High SES x FGM Sensitive List	0.00	0.02	0.95	0.00	0.02	0.94
High UK Cultural Integration x High SES x FGM Sensitive List	0.04	0.04	0.25	0.04	0.04	0.25

\*\*p <. 0.05. \*\*\*p <. 0.01. \*Standard errors are cluster robust, clustered by participant.

## Discussion

This study aimed to answer the question of whether cultural sensitivity and SES influence UK professionals' decision to report FGM. It was hypothesised that: 1) professionals would be less likely to report FGM where a family is less integrated into UK culture, and 2) where the family is of higher SES. Participants' characteristics that could influence reporting behaviours were also assessed (exploratory variables were professionals' gender, age, profession, level of training, race, parents' country of origin, and hours of FGM training received).

Direct questioning in our study suggested that professionals were concerned enough to report the family to the police 78% of the time. Unlike in the list experiment, vignette type within direct questioning did have some effect. When asked directly, professionals were 5% less likely to say they would report to the police when the family was described as being both highly integrated into UK culture and of high SES. In other words, professionals were less likely to report a family, only when that family was portrayed as being both affluent and highly acculturated into UK culture. Vignette type did not otherwise significantly impact results. Nor did the exploratory variables.

Results for the list experiment indicated that it was satisfactorily designed. Moreover, participants who were shown the sensitive list said that they would report the family to the police 79% of the time. However, this was irrespective of vignette type. Asking the FGM-sensitive question in an indirect manner via the list

experiment did not significantly impact professionals' responses to any of the vignettes. While Black participants on average answered yes to 30% more questions in the list experiment, suggesting that they might be more likely to report, this result was not robust and is likely due to 81% of Black participants being allocated to the experimental group. This is because they had considerably more opportunities to respond to the FGM-sensitive list, which contained five items, as opposed to the control list, which contained four items. Assignment to group condition was random, so this occurred by chance, as a result of the relatively small number of Black participants in the sample.

Judgments and decisions about child maltreatment are dilemmatic. Both the DME (Baumann et al., 1997) and the JUDPIC model (Benbenishty & Arad-Davidson, 2012) suggest that when making such decisions, professionals take into account case factors, professional factors, organisational factors, and external factors such as cultural norms. The present study contributes to the literature by attempting to understand professionals' application of the FGM reporting duty across all of these levels. It was considered that cultural sensitivity (an interaction between all the levels of factors) and SES (case factors) might influence UK professionals' decision to report FGM. We controlled for professional factors shown to be influential of reporting behaviour including gender, links to a FGM practicing country, and amount of FGM training received, and also organisational factors (i.e. their training background). The use of a list experiment to reduce any social desirability bias

targeted external factors such as the mandatory nature of the legislation and the current cultural climate.

Global commitment to eradicate FGM has been accelerating (e.g. the 2012 adoption of UN resolution Intensifying global efforts for the elimination of female genital mutilation by all 194 members of the General Assembly, including the 29 practising countries; national laws banning FGM) (Muthumbi, Svanemyr, Scolaro, Temmerman, & Say, 2015, p. 33). Nevertheless, it remains an acceptable practice in some communities, perpetuated by sporadic enforcement of FGM laws, and delays in raising awareness and thus shifting the underlying socio-cultural norms behind the practice (Muthumb et al., p. 38). This potentially leaves practitioners in a bind when deciding the best course of action for a child at risk of FGM. Strict enforcement might have unwanted consequences for the children that the laws are meant to protect. Girls might be traumatised by the criminalisation of their parents, and they and other children in the family may be left in a more vulnerable situation if their caregivers are imprisoned (Boyden, Pankhurst, & Tafere, 2012, p. 519). Moreover, professionals might not be fully aware of the anti-FGM legislation and their obligations under it (Zurynsk et al., 2015, p. 16).

The FGM mandatory reporting duty came into effect in October 2015, requiring all UK health and social care professionals and teachers to report known cases of FGM in under 18-year-olds. Despite FGM being made illegal in the UK in 1985 and the law updated in 2003 to include procedures carried out abroad on UK nationals and

residents, it was deemed necessary to introduce a FGM reporting duty, distinct from other child abuse safeguarding regulations. Authors have argued that due to the disparity between assumed prevalence and actual referrals, its purpose was to break down professionals' concerns regarding cultural sensitivity (Mathers & Rymer, 2015, p. 283). In the absence of other UK reporting statistics, our results suggest that for the most part it may have succeeded in this aim. The majority of professionals, whether asked indirectly within a list experiment or directly, say that they would report a known case of FGM just under 80% of the time. Moreover, direct questioning in our study suggested that rather than being less likely to report families with strong ties to their potentially FGM-practising communities, professionals showed less concern when the family was described as being both well integrated into UK culture and of high SES.

The United Kingdom and France are the European countries with the largest immigrant communities from FGM-practising countries (Guiné & Fuentes, 2007, p. 481). However, they have taken divergent paths in their handling of the issue. France is the only European country to have actively used the courts by prosecuting families as well as practitioners, with over 40 prosecutions since 1979 resulting in the conviction of over 100 parents and cutters (Rahman & Toubia, 2000). In contrast, Britain's first prosecution did not occur until 2012 and did not result in a conviction (Jefferson, 2015, p. 418). The first successful prosecution did not occur until February 2019, when a 37-year-old mother from east London became the first person convicted of cutting her 3-year-old daughter. This occurred during data collection for

this study with widespread reporting in the media and it is possible that as a result professionals were particularly primed to say that they would report a known case of FGM.

Media reporting of high-profile cases such as the Baby P case has been shown to increase referrals for child protection assessments (Ray, Conn, Liebeschuetz, Costoli, & Tan, 2013, p. 21). The impact of media reporting may have been particularly relevant because our results show that the majority of participants sampled had received very little FGM training. 37% had received no FGM training at all, while 40% had received only 1-2 hours. This is line with Oakley et al.'s (2017) findings in which 74% of professionals had received no specific training on child belief-related child abuse, including FGM.

In our study acculturation and SES were manipulated in combination so results appear to be due to the interaction of the two variables. Vignettes that manipulated high acculturation with low SES and low acculturation with high SES did not produce significantly different reporting decisions. The incidence and prevalence literature suggests that individuals from lower income backgrounds are at greater risk for all forms of maltreatment (Hansen et al., p. 328). Moreover, socioeconomic inequalities are especially correlated to deaths from child abuse (Gilbert et al., 2009, p. 72). Accordingly, there is evidence that lower family SES is predictive of professionals' increased likelihood to report abuse (e.g. Hampton & Newberger, 1985; Hansen et

al., 1997; Zellman, 1992) and remove children from the family home (e.g. Lindsey, 1991, p. 279).

Our results do not fully substantiate these findings since low SES in the absence of high acculturation did not produce significantly higher reporting. Professionals in the sample were slightly more reluctant to report a family only when that family was both highly acculturated into UK culture and of high SES. While the reasons for this are unclear, the literature suggests that practitioners' stereotypes of an "abuser" may determine which parents of an injured child are considered as possible abusers. Moreover, the label "child abuser" may be less likely to be applied where the decision-maker and suspected abuser share similar characteristics such as SES (Hampton & Newberger, 1985, p. 57) (and level of acculturation). It may also be that professionals felt that a child from an affluent family that was well integrated into UK culture may have more to lose from the criminalisation of her otherwise well-meaning parents.

Unexpectedly, it was answers to the direct questions that showed some variability due to vignette type. The list experiment has gained in popularity in recent years and there is some evidence that its use increases estimates of socially undesirable behaviours in comparison to direct questioning (e.g. LaBrie & Earleywine, 2000; Tsuchiya, Hirai, & Ono, 2007). Moreover, De Cao & Lutz (2015) have pioneered its use in researching attitudes on FGM. However, in a meta-analysis of studies, Tourangeau and Yan (2007, p. 873) found significant variation across studies. They

found that studies using undergraduate samples were likely to yield positive results (i.e. increased reporting under the list experiment), in contrast to the one general population survey included in the study which received negative results.

Emerging research in the behavioural sciences suggests that university student samples are particularly unusual compared to the general population (see Henrich, Heine, & Norenzayan, 2010). It is possible that there is something specific about undergraduate samples (such as their relative youth and lack of self-assurance) that makes them more susceptible to the social desirability bias, and therefore more likely to make use of a list experiment. List experiments have also been traditionally used in political science to examine Americans' attitudes about sensitive topics such as race and political stance. The US has a unique set of social and cultural dynamics characterised by often difficult race relations. Perhaps in such a climate study participants might welcome the chance to mask their true answers. Whereas participants in the present study might have been suspicious of the list experiment, preferring instead to lower their guard when asked their views directly.

### **Strengths, Limitations, and Directions for Future Research**

There were a number of limitations to the study, and the generalisability of the findings is limited by the study's participant characteristics. The sample was mostly composed of White females, almost half of whom were clinical psychologists. Whilst this demographic is quite representative of the BPS membership, due to variability amongst professions, services, and cities, it is unknown if professionals across the

country would have responded in a different manner. Additionally, as is always the case with online surveys, sample selection effects are possible (Sieracki, Fuller, Leon, Jhe Bai, & Bryant, 2015, p. 9). Additional research, with larger samples that might come closer to being nationally representative, would be useful to cross validate the present findings. Stratification of the sample by gender, race, and other pertinent variables would also be advantageous.

While results show that the list experiment was satisfactorily designed, it was direct questioning that generated variability in participants' responses to the vignettes. This suggests that similar future research sampling UK professionals might benefit from focusing on direct questioning techniques that are less likely to raise professionals' suspicions about the researchers' intentions. However, as we can only guess at why direct questioning yielded more variability in responses than the list experiment, qualitative methods might complement the results of this study by allowing for a detailed exploration of professionals' thought-processes (Braun & Clark, 2013). Moreover, despite the vignettes being pre-screened and extensively piloted, because they were created exclusively for the present study, reliability and validity are unknown.

The analogue vignette study has the advantage of flexibility; allowing for variables to be manipulated in a controlled manner. Much of the previous research in child maltreatment reporting has successfully used similar vignette formats and self-report items (e.g., Finlayson and Koocher, 1991; Zellman, 1990). However, this

method also compromises external validity. In real-life situations practitioners' decisions are based on a much greater and more nuanced amount of information, received through multiple sensory systems. Verbal descriptions in a vignette cannot fully portray the extent of the information received (visual, aural, tactile, or olfactory), and therefore nor can they impact on professionals' decision-making to the same effect (Mandel et al., 1995, p. 919). Working with a child (and their family) and developing a clinical or therapeutic relationship with them is also bound to affect professionals' judgments in ways that a verbal vignette study cannot capture. Future research could capitalise on advances in technology, particularly virtual reality computer-based simulations, which have started to be used in clinical psychology research and practice, and which can provide real-life experiences that are emotionally engaging (Riva, 2008). Although beyond the scope of this study, in due course, service, regional, and national data gathered following the mandatory reporting duty might also be analysed for an understanding of the direct impact of the duty (e.g. Ards & Harrell, 1993).

The mandatory reporting duty refers to known cases of FGM, meaning cases where a child has disclosed to the professional that she has undergone FGM. The vignettes in this study therefore depicted a child telling the professional that she had been 'cut in a special ceremony'. But how likely is this to occur in the real world? Studies suggest that relying on children's statements about abuse can be problematic since many remain silent about abuse and may even deny that abuse has occurred (London, Bruck, Ceci, & Shuman, 2005, p. 195). Parents are often the first to notice signs of

something being amiss so a key resource is missed if they are in fact the perpetrators of the abuse (Fontes & Plummer, 2010, p. 492). Moreover, cultural values such as shame or modesty around sexual organs, or family respect and honour might prevent affected girls from disclosing (Fontes & Plummer, 2010, p. 496). A study assessing the reliability of self-reported FGM in women (comparing the extent of cutting verified by clinical examination with the corresponding WHO classification) found that even amongst adults, there is considerable under-reporting of the extent of FGM (Elmusharaf, Elhadi, & Almroth, 2006). This suggests that for the reporting duty to be as effective as possible, it is imperative that researchers try to understand the factors impacting children's disclosure of FGM. Moreover, in their study describing the first year of activity of the UK's first paediatric FGM service, Creighton et al. (2016, p. 5) warn that the unexpectedly high proportion of girls with type 4 FGM found in their study, may indicate that families may be changing to practices that are less detectable. Further research on FGM trends within FGM-practising communities is therefore vital.

Making it clear in the vignettes that the child had undergone FGM, removed ambiguity, thus making decision-making considerably easier. Tversky and Kahneman's (1974) research has demonstrated that under conditions of uncertainty, humans use a number of heuristics that can lead to error. For example, using the 'representativeness' heuristic might mean that we estimate the likelihood of an event by comparing it to an existing example that already exists in our minds, thus potentially leading to the use of stereotypes which might introduce bias. It is

possible that if the vignettes had presented more of an ambiguous scenario, and looked to understand professionals' initial concerns (rather than their reporting behaviour), a clearer effect concerning the influence of professionals' cultural sensitivity and families' SES on professionals' decision-making may have been found.

Finally, twenty percent of professionals sampled said that they would not report to the police despite being presented with a clear disclosure of FGM. Future research, possibly qualitative (as this would allow for practitioners to outline their reasoning in their own words), is required to understand these decisions.

### **Conclusion**

FGM continues to be a highly emotive and fiercely debated topic amongst researchers and healthcare professionals not just in the UK, but across the West, and in countries where it continues to be regularly practiced (e.g. Earp, 2016; Johnsdotter, 2019). It sits right at the interface between tradition and progress; minority and dominant culture; cultural bias and cultural sensitivity. Professionals are required to respond to fast-changing policy and legislative developments with regards to the practice. While the mandatory reporting duty came into effect in 2015 amidst such controversy, thus far there has been no research to identify factors affecting its implementation by professionals. It is only relatively recently that FGM has become understood as child abuse and the implications of such a shift are not yet fully understood. This study contributes to the literature by attempting to understand professionals' application of the FGM reporting duty, specifically

whether cultural sensitivity and family SES might influence professionals' reporting decisions.

Whether asked indirectly or directly, the majority of professionals sampled said that they would report a known case of FGM in just under 80% of instances. Moreover, direct questioning suggested that professionals were slightly less likely to report to the police when the family was described as being both well acculturated and of high SES. In the absence of other UK FGM reporting statistics, our results tentatively suggest that for the most part the mandatory reporting duty appears to have been successful in breaking down professionals' concerns about needing to be cultural sensitive towards practising families. It is also possible that perceived affluence coupled with high acculturation might dissuade professionals from reporting for fear of causing more harm to the child. The clinical implications of this must be considered carefully so that economically and socially disadvantaged families are not further penalised by the systems meant to serve them. Results indicated that the majority of professionals sampled had received either none or very little FGM training. High-profile media reporting should not be the main source of information for professionals delegated with the duty of protecting children from the negative consequences of FGM.

Mandatory reporting of known cases of FGM may be effective in sending a message to professionals (and practising families) that FGM is detrimental to girls' wellbeing and is therefore unacceptable, irrespective of cultural background. The reporting

duty however, cannot be truly impactful unless we also understand the conditions for such disclosures. Professionals need to be provided with appropriate training to develop skills in cultural awareness so that they are aware of potential biases and the extent of influences on their decision-making in this area. Facilitation of more accurate identification and reporting of FGM in children will in turn facilitate prevention of maltreatment and increase positive outcomes for families and professionals.

## **Integration, Impact, and Dissemination Summary**

## Integration

The female genital mutilation (FGM) mandatory reporting duty came into effect in October 2015 via the Serious Crime Act 2015. It requires all United Kingdom (UK) health and social care professionals and teachers to report *known* cases of FGM in under 18-year-olds. Despite previous legislation outlawing FGM, and an increasing acknowledgement that it is a form of child abuse (due to its potentially serious physical and psychological consequences), it was deemed necessary to introduce mandated reporting of FGM distinct from other child abuse safeguarding regulations. Due to FGM being a traditional and cultural practice of a number of communities in Africa, Asia, and the Middle East, it is argued that this was in order to circumvent professionals' concerns about being culturally sensitive towards FGM-practicing communities in the UK (Mathers & Rymer, 2015, p. 283). Research suggests that there are a many factors influencing professionals' decision-making regarding child abuse reporting, including cultural factors and social factors such as socioeconomic status (SES). The aim of the doctoral thesis was therefore to explore the impact of the reporting duty, by investigating whether professionals' cultural sensitivity and a family's SES might influence UK professionals' decision to report FGM.

When reading around the topic, it became apparent from the child maltreatment literature that while FGM might be cultural practice that is only relatively recently being understood as child abuse, it is one of several practices that is normative to some groups, but also potentially harmful to children within those groups. Corporal punishment (illegal in some countries but not others), ridding-evil (e.g. Catholic

exorcisms), witch-craft, and refusal of medical procedures (e.g. Jehovah's Witnesses' refusal of blood transfusions) are other such examples. This led to a broader question about the impact of cultural differences on understandings of child maltreatment and how definitions of child abuse are demarcated and then interpreted and implemented by the practitioners tasked with child protection. It was hoped that an exploration of this in the systematic review would provide a comprehensive overview of the role of cultural factors in professionals' decisions about child maltreatment. The empirical study could then narrow down and provide a detailed understanding of the role of cultural sensitivity (and SES) on professionals' reporting of a currently salient cultural practice, namely FGM. The systematic review therefore provided an overview and clear rationale for the empirical article, and assisted its development in a number of ways, which are outlined here.

The review provided an outline of the current classification of child maltreatment and discussed problems with this. It particularly considered the impact of social and cultural differences which make cross-cultural comparisons problematic. Potentially, these differences may also introduce cultural bias into professionals' decision-making in ways that are not yet fully understood. By focusing on cultural factors, it also aimed to move the literature on from its previously narrow focus on race to distinguish population groups, as racial categories do not capture the full diversity of cultural preferences present in all racial groups. The review thus provided an-up-to-date summary of what the literature says about the role that cultural factors play in Western-based professionals' decisions about child maltreatment. Reading for the

review involved studying published reviews on related topics such as Zurynski, Sureshkumar, Phu, and Elliott's (2015) systematic review of health professionals' knowledge, attitudes and clinical practice around FGM, and this provided a useful source of information for the empirical paper. This meant that review findings were relevant to the empirical study and could inform thinking and dialogue throughout the article. Moreover, having mostly had previous research experience in qualitative methods, I felt much less familiar with quantitative methodology. Reading and interpreting results of the included studies in the review therefore helped to build my confidence in understanding and writing-up the results of the empirical study, the first quantitative study I have worked on in which I have been the principal investigator.

Review findings formed part of the background literature for the empirical paper. While the review did not find any studies looking specifically at FGM (providing further rationale for the necessity of the empirical study), reading for the review, and the reviews' findings concerning evidence of cultural bias regarding child maltreatment, helped to provide context for the empirical study. For example, Oakley, Kinmond, Humphreys, and Dioumd's 2017 study, which was included in the review, found that while the majority of professionals said that they had heard of the term child abuse linked to faith or belief (CALFB), less than 2% offered FGM as an example of CALFB, and 74% said that they had received no specific training on CALFB. Similarly, the review findings were able to inform the discussion section of the empirical study, providing contextual background for the study's results. As an

example, it was discussed that the majority of participants sampled in the empirical study reported having received either no FGM training at all, or very little, in line with Oakley et al.'s 2017 findings.

Review findings also gave a strong rationale for the methodological design of the empirical study. The majority of studies in the review employed vignettes, in the form of hypothetical written scenarios, to infer whether cultural biases existed in practitioners' treatment of suspected child abuse cases. Most of the studies employed identical vignettes in which the race or ethnicity of the child or family were manipulated in order to measure the impact of case characteristics on professionals' attitudes and decisions. Although not a perfect methodology, the analogue vignette study provides experimental flexibility by allowing for variables to be manipulated in a controlled manner, hence its popularity amongst studies included in the review. This provided a clear rationale for the use of this methodology in the empirical study. The empirical study therefore made use of vignettes to manipulate variables of interest, namely the family's level of cultural integration and their level of SES, while controlling for the possible influence of the child's age, race, and country of origin.

The review highlighted that researchers have grouped variables influencing professionals' decision to report child abuse into professional, perpetrator, and case characteristics (Terao, Borrego, & Urquiza, 2001, pp. 158). Cultural factors might be relevant to all of these categories, however, studies in the review did not use theory to inform how cultural factors might influence professionals' decision-making. The review also highlighted the existence of cultural bias concerning child maltreatment

decisions among professionals of different disciplines, regarding both case and professional variables. However, the extent and nature was unclear. Consequently, review findings emphasised that there was a continuing need to develop practitioners' cultural competence in order to mitigate potentially damaging cultural biases. It was therefore felt necessary to link the empirical study to relevant theories. Background reading for the empirical paper suggested that more comprehensive decision-making models such as the Decision-Making Ecology (DME) model (Baumann, Kern, & Fluke, 1997) and the model of Judgments and Decisions Processes in Context (JUDPIC) (Benbenishty & Arad-Davidson, 2012) might be able to adequately explain the impact of cultural sensitivity and/or cultural bias on professionals' judgements.

While discussions with supervisors were crucial to the design of the empirical study, quality appraisal of the included studies within the review also provided a good understanding of some of the pitfalls of low quality studies and the characteristics of good quality studies. This allowed not only for the further development of rationale around the choice of study methodology, but also helped enrich discussion around the empirical study's limitations. For example, whereas many of the review studies used non-randomised samples, and some failed to randomise the order of vignettes, the empirical study did randomise participants to either a control group or experimental group, and the vignette order was also randomised in order to eliminate any systematic relationship between order and vignette. Some of the review studies also did not discuss piloting and validating of the vignettes and

measures used, which alerted me to the importance of doing so for the empirical study. Similarly, social desirability effects were found for half of the studies in the review, and it was considered important that due to the mandatory nature of the FGM reporting duty, that this should be addressed for the empirical study. As a result, a list experiment by which participants could 'mask' their answers was designed (Glynn, 2010, p. 2). Many of the review studies had non-representative samples due to having low proportions of non-White professionals and male participants, and one study also suffered due to recruiting from a single workshop. This alerted me to the importance of trying to recruit as diverse and representative a sample as possible.

However, attempting the above, was not without its challenges. Randomising the empirical study's participants to either the experimental or control groups, then randomising the order in which they were presented with vignettes within the list experiment, and then ensuring that they again received the same order of vignettes when they were asked the FGM-sensitive question directly, involved a considerable amount of work. It was beyond the scope of the Qualtrics software and I had to be resourceful in finding a coding expert that was able and willing to help me develop a Java script that could be embedded into Qualtrics for these purposes. My supervisor also consulted a research expert who could help us properly design the list experiment. There are several key rules to properly designing a list experiment, and it was required that four lists (to accompany each of the four vignettes) be created that included negatively correlated control items (where a subject who

endorses item 1 is unlikely to endorse item 2; e.g. 'is the torture of prisoners acceptable?' vs. 'should the death penalty be banned?') and at least one easy-to-endorse (e.g. 'do you like spending time with your friends?') and one hard-to-endorse item (e.g. 'do you regularly sleep-walk?'). While I enjoyed the creative aspect of designing the lists, this required extensive piloting under time-limited conditions.

Moreover, the design and piloting of the four vignettes was also very labour and time intensive. I was keen to sharpen thinking for the study and ensure that it was as relevant as possible to developments in FGM. Consequently, at the start of the project, I arranged telephone meetings with my supervisors and the prominent anti-FGM activist and psychotherapist Leyla Hussein (founder of the Dahlia project, a specialist service for victims of FGM, and co-founder of Daughters of Eve, a charity dedicated to ending gender-based violence including FGM), whom I had previously collaborated with on anti-FGM campaigns. The intention was for Leyla to act as an external supervisor throughout the empirical study, and although she initially agreed, she was unable to continue for personal reasons. However, my remaining supervisors and I met to design the vignettes and I then piloted them extensively for the purposes of content-validity. Although labour intensive, review findings (and results of the early versions of the vignettes) made it clear that piloting was vital to properly designing the empirical study. Ultimately, it was another creative and collaborative process that I learnt a considerable amount from.

Piloting of the vignettes and the list experiment also brought me in to close contact with professionals and service users, thus helping me to maximise the relevance of the empirical study. The Department of Health (Department of Health, 1999, 2000) has stated the importance, and benefits, of involving service users in the research process. Minogue, Boness, Brown, and Girdlestone (2005, p. 104) have identified five levels of service user involvement; consultation, collaboration or partnership, user-commissioned, user-controlled or user-led, and user dissemination. Using this framework, I think that the current study sits between the 'consultation' (i.e. service-users are asked for input but have limited influence) and 'collaboration' (i.e. service-users can make suggestions and influence outcomes) levels. While it was not possible to involve service users at all stages of the empirical research (due to limited time and resources), members of the public's views were sought during the design stage. I used contacts made from FGM campaigning to invite women who have publicly shared their experiences of having undergone FGM to share comments on the study methodology, particularly the vignettes. Sharing views with potential service users who are directly impacted by FGM, also allowed for the research to benefit from some of the advantages of service user involvement, that have been identified by the literature (e.g. Hewlett, p. 679). For example, as researcher, I benefited from a greater understanding of the research issue, and a sense that my efforts were worthwhile, which has been incredibly motivating; while feedback from the women has been that asking their views felt empowering, and that it was stimulating to reflect on recent research developments.

Finally, the review findings emphasised that it would be important to try and recruit a representative sample. The review highlighted the need to try and recruit from as many sites as possible. Ethics for the study was sought so that recruitment could be carried out at many different services within Camden and Islington NHS Foundation Trust, a large and diverse trust, and through the researchers' networks, including social media. Most of the review studies were limited by non-representative samples made up mostly of White and female participants, and unfortunately, despite hoping to recruit more of a diverse sample, the empirical study was similarly restricted. While it was discussed with my supervisor that I could try to specifically target professionals from minority backgrounds, it was decided against this, as doing so might have introduced bias. Moreover, while it was originally proposed that randomisation of the study participants into group condition would be stratified by gender and race in order to minimise bias, this proved too ambitious (in the time available) even with the consultation of a coding expert to help set up the study.

Overall, in spite of these challenges (and also because of them), I am pleased with the integration of the multiple components of the thesis. It was my aim that the thesis should have a clear narrative. Due to my background and my personal experiences, I am passionate about the issue of FGM. Prior to training, I worked part-time as a FGM Prevention Trainer for the charity Family Action, and during training I continued to facilitate FGM training workshops where possible, including to my cohort. This meant that I had many discussions with various professional groups (including psychologists, psychotherapists, and teachers) about their understanding

of FGM, and the challenges they face in their work around the issue, including their views on the reporting duty. I felt that it was important to provide a proper overview for the empirical study. In the end, this resulted in a review that is cognisant of the fact that FGM is only one of many practices in which the wider issues around cultural competency in professionals' decisions about child maltreatment play an important role. The framework of the thesis was influenced by having access to the remarkable expertise of my supervisors, and by my personal experiences of working with other FGM-campaigners, talking to FGM-impacted women, and the rich discussions I have had with professionals when providing them with FGM training.

### **Impact**

International law recognises FGM as one of the most obvious and severe forms of violence against girls and women, and places an obligation on governments to take steps to prevent it (Christou & Fowles, 2015, p. 349). However, the manner in which to do this is unclear, and as a result, FGM continues to be a highly emotive and fiercely debated topic amongst researchers and healthcare professionals across the world. Just the terminology around the practice is highly contentious (i.e. the use and connotations of the term “female genital mutilation” versus “female genital cutting,” “female circumcision,” “female genital alteration”, and so on; Earp, 2016). Moreover, the ways in which governments and mandated practitioners decide to tackle the problem of FGM is also under constant deliberation and debate (e.g. Johnsdotter, 2019). Ongoing migration is leading to increasingly ethnically diverse populations, and clinicians are increasingly required to be sensitive to culturally related issues

(Terao et al., 2001, pp. 158). By attempting to understand factors that might affect professionals' implementation of the 2015 FGM mandatory reporting duty; and more broadly, how cultural sensitivity and/or cultural bias might influence professionals' decisions about child maltreatment cases in general; the current review and empirical research impact on clinical, policy, and research developments in a number of important ways that are outlined below.

Potentially, the research findings could have a far-reaching impact, as they relate to the practice of all professionals working with children and families. In the UK, this is all health and social care professionals or teachers, including psychologists, doctors, dentists, nurses, midwives, social workers, and so on. Viewing the findings of the systematic review could lead health and care professionals and teachers, and others responsible for their training (e.g. universities, service managers, professional bodies, clinical supervisors), to develop cultural competency aspects of the training. This would mean that professionals are trained and supported to recognise, acknowledge, and where appropriate, mitigate cultural biases. An understanding of the background and findings of the empirical study, could help professionals to be aware of the debate surrounding FGM, to feel more confident of their mandated duties concerning FGM, and to be aware that their cultural sensitivity coupled with their perceptions about a family's SES (and its impact), might introduce some bias into their decision-making about reporting FGM. This is important so that already socially and economically underprivileged children and families are not further disadvantaged.

Cultural competence training, including FGM training, is being increasingly reported in the literature and is gaining the attention of health and social care administrators and educators, and policy makers (Beach et al., 2005; Price et al., 2005). Beach et al.'s systematic review of cultural competency educational interventions (2005, p. 6) evaluated several different curricular or training methods and content areas (e.g. experiential learning involving cultural immersion, or involving discussion with members of another culture), and found that there is excellent or good evidence that cultural competence training impacts intermediate outcomes such as the knowledge, attitudes, and skills of health professionals, and good evidence that cultural competence training impacts patient satisfaction. Their review also found that there were no two studies evaluating the exact same educational experience, or comparing different types of training methods or content. However, almost all studies reported a positive effect, suggesting that employing any intervention may be effective, including both longer and shorter duration interventions, and experiential and non-experiential methods.

Moreover, there was little uniformity across studies in measurement of outcomes. Even where standardised measures of cultural self-efficacy were used, the studies measured too wide a range of attitudes to allow for comparison (p. 7). Therefore, not only would it be beneficial for training incorporating the findings of the systematic review and the empirical study to be properly evaluated for an understanding of its impact (via pre/post quantitative measures, alongside qualitative feedback). Training provision would also benefit from the development of standardised measures of

cultural competence, which would allow for comparisons between studies in the future. Furthermore, since Beach et al. (2005, p. 7) found no studies that have evaluated the direct impact of cultural competency training on patient health outcomes, future assessments of the link between training and relevant patient outcomes of interest will be advantageous.

The findings of both the systematic review and the empirical study may be of interest to service users, particularly those impacted by FGM, but also any who are from a minority background. Both papers could help them to make sense of how professionals might relate to their culture and cultural practices, and potentially this could in turn impact their own decision-making around these practices. For example, exposure to the findings of the empirical study might deter parents from FGM-practicing communities who are ambivalent about the practice from cutting their daughters, since the study suggests that if it were to be disclosed, most likely they would be reported to the police. Exposure to the findings of the systematic review could help service users (and parents of school children) to make sense of any cultural bias they may have experienced by professionals, and empower them to better understand the framework within which decisions about them are made.

Several organisations (e.g. United Nations Children's Fund (UNICEF), the National Society for the Prevention of Cruelty to Children (NSPCC), FORWARD (Foundation for Women's Health Research and Development)) and community agencies and charities (e.g. Daughters of Eve, the Dahlia Project, Hawa Trust, and Family Action) offer

information, support, and guidance about FGM. The current findings could be used and incorporated into their campaigns to help professionals and families from FGM-practising communities understand and adhere to FGM legislation, potentially preventing FGM. Key findings of the empirical paper (i.e. that just under 80% of UK professionals said that they would report a disclosed case of FGM to police) could be reported on their websites. Alongside media reporting of FGM trials and convictions, this could help these organisations to share understanding of the implications of the reporting duty and promote abandonment of FGM as a practice.

There is a chance that the reporting duty, and consequently the findings of the empirical study, may drive FGM further underground, or lead to an increase in forms of FGM which are less detectable (such as type 4 FGM). It is therefore hoped that the findings of the review and the empirical study will be taken up by child protection researchers. It will be important to cross-reference the findings of the empirical paper which are in relation to hypothetical case scenarios presented to professionals, with real-world service, regional, and national data gathered following the mandatory reporting duty. It will be important to investigate the reporting duty's actual impact on clinical outcomes. The empirical study also highlighted that it is important that research also address the factors impacting children's disclosure of FGM, since the reporting duty only applies to known cases of FGM. The findings of the empirical study also contribute to the wider literature about UK professionals' reporting by suggesting that the use of direct questioning methods might be more advantageous than using a list experiment (a form of indirect questioning meant to bypass social

desirability bias) to elicit professionals' views.

Findings of both the papers, but particularly the systematic review also highlight that while broad definitions of child maltreatment allow courts enough flexibility to respond to real-life situations, this stance also negates social and cultural differences, making any cross-cultural comparisons especially problematic. Policy-makers might therefore benefit from being exposed to the review findings. Particularly in the current climate of ongoing migration, and continued debate around issues of immigration and the cultural integration of minority groups. Therefore, current study findings could be outlined and summarised to provide helpful suggestions of how guidelines can be written to facilitate professionals' understanding of their duties with regards to practices that are normative to some groups, but also potentially harmful to children.

Finally, working on the thesis has also had a personal impact. Immersing myself in the literature around both child maltreatment and FGM has been very eye-opening. I have personal experience of the consequences of FGM, including feelings of ambivalence about the ways in which otherwise well-meaning FGM-practicing parents (such as my own) should be viewed and treated by professionals and the law. It has been interesting to see some of my confused thoughts and feelings play out on the pages of other research studies. Comforting even, to see that it really is not simple at all. Are parents to blame? If not, who is? What is the best way forward? Is it too easy for Western culture to look and judge the other? Or would we prefer

that it turn a blind eye to the potential suffering of 'others'? What is best for the child? What would have been best for me? These are just some of the questions I have asked myself throughout work on this thesis. As a researcher it has been important to put my personal feelings and beliefs aside and try to be as impartial as possible in conducting both studies. However, my personal experiences have motivated me to want to do justice to the issues that are discussed in this thesis. The experience has taught me about the challenges of working in child protection, and how important it is to keep talking and asking questions, particularly about potentially sensitive or taboo subjects such as FGM. This will also be beneficial to my clinical practice as I have a better understanding of the manner in which cultural sensitivity and bias might impact my own work.

### **Dissemination**

For the research to reach a diverse audience and have as much impact as possible, the dissemination of the research will be organised through several channels, including traditional routes such as publication in relevant journals, more modern platforms such as social media, and my own networks gained through work campaigning about FGM and delivering FGM training to professionals. The main aim is to make the research widely available to encourage further discussion and understanding about cultural factors affecting professionals' decisions about child abuse, and about the impact of professionals' cultural sensitivity and family SES on professionals' understanding of the FGM mandatory reporting duty.

Firstly, dissemination will be via the provision of a plain English summary of the research to the study participants who emailed to indicate that they would like to receive a summary of findings. This process will involve discussion with the research supervisors to agree on the version of the summary before sharing with participants. Participants were not remunerated in any way for taking part in the study, so it is important to ensure that they are at least provided with a summary of the study's findings.

Secondly, the findings of the empirical study have been disseminated locally via a presentation to staff and students at Royal Holloway University. Responding to questions following the presentation has contributed to the write-up of the thesis, particularly the discussion section of the empirical paper, as it alerted me to some of the questions that readers of the study might have. Moreover, the empirical study will also be disseminated during a Continuing Professional Development session at one of the recruitment sites.

In order to reach an audience beyond the immediate participants and recruitment sites, dissemination to a wider audience will be sought. It was agreed early on with one of the supervisors of the thesis that the empirical study will be submitted to peer-reviewed journals. To this end, prior to data collection, the empirical study was registered with Open Science Framework (an online project that facilitates open collaboration in science research by allowing for research projects to be publicly pre-registered). Carrying out the systematic review has alerted me to the type of

publications that might be best suited to the study. These include Child Abuse & Neglect (this journal will be prioritised since the many of the review studies were published here), Journal of Child and Family Studies, Journal of Family Violence, Child Maltreatment, and Journal of Criminal Law. An application will be made to present the findings at relevant mental health conferences such as Early Intervention for Children and Young Adults and the CAMHS National Summit. If the application is accepted, this would provide an excellent means of maximising the findings of the research to a targeted nationwide audience. Although not initially planned, it has been agreed that the systematic review will also be submitted for publication.

Fourthly, in order to facilitate dissemination to professionals I will incorporate the findings of both the systematic review and the empirical study into the FGM training that I deliver to various professional bodies. In the past, this has included first year clinical psychology doctoral trainees, multi-disciplinary team members on my various doctoral placements, teachers at various schools in London, psychologists at Chelsea and Westminster Hospital, and the Association of Psychoanalytical Psychotherapists in the NHS. Due to my previous experience, I am regularly approached to provide training workshops, and I also hope to contact various organisations that might potentially find the training useful. This could be achieved through contact with the recruitment sites, and through continued discussions with professionals I come into contact with. The impact of this will be maximised if the training is evidence-based. Feedback will be sought after each workshop to ensure that impact is maximised, and also evidenced.

Finally, social media is a useful tool by which the research can be promoted. A short summary of the research will be posted on LinkedIn, Facebook, and Twitter in order to disseminate findings. This would enable communication to a wide audience, networking with stakeholders of the research, which could potentially lead to further opportunities to disseminate the research.

## References

- Alvarez, K.M., Kenny, M.C., Donohue, B., & Carpin, K.M. (2004). Why are professionals failing to initiate mandated reports of child maltreatment, and are there any empirically based training programs to assist professionals in the reporting process? *Aggression and Violent Behavior, 9*, 563-578.
- Amasanti, M. L., Imcha, M., & Momoh, C. (2016). Compassionate and proactive interventions by health workers in the United Kingdom: a better approach to prevent and respond to female genital mutilation? *PLoS Medicine, 13*(3).
- Ards, S., & Harrell, A. (1993). Reporting of child maltreatment: A secondary analysis of the National Incidence surveys. *Child Abuse & Neglect, 17*(3), 337-344.
- Ashton, V. (1999). Worker judgements of seriousness about and reporting of suspected child maltreatment. *Child Abuse & Neglect, 23*(6), 539-548.
- Ashton, V. (2004). The effect of personal characteristics on reporting child maltreatment. *Child Abuse & Neglect 28*(9), 985-997.
- Barstow, D.G. (1999). Female genital mutilation: the penultimate gender abuse. *Child Abuse and Neglect, 23*, 501-510.
- Barth, J., Bermetz, L., Heim, E., Trelle, S., & Tonia, T. The current prevalence of child sexual abuse worldwide: a systematic review and meta-analysis. *Int J Public Health. 2013, 58*, 469-83.
- Baumann, D. J., Dagleish, L., Fluke, J. D., & Kern, H. (2011). The decision-making ecology. Washington, DC: American Humane Association.
- Baumann, D., Kern, H., & Fluke, J. (1997). Foundations of the decision making ecology and overview. In Kern, H., Baumann, D. J., & Fluke, J. (Eds.). *Worker Improvements to the Decision and Outcome Model (WISDOM): The child welfare decision enhancement project*. The Children's Bureau, Washington, D.C.
- Beach, M. C., Price, E. G., Gary, T. L., Robinson, K. A., Gozu, A., Palacio, A., ... Cooper, L. A. (2005). Cultural competence: A systematic review of health care provider educational interventions. *Medical Care, 43*, 356-373.
- Benbenishty, R., & Arad-Davison, B. (September, 2012). A controlled study of placement and reunification decision in Israel. Presented in EUSARF, Glasgow.
- Beanland, C., Schneider, Z., LoBiondo-Wood, G., & Haber, J. (1999). *Nursing research: methods, critical appraisal and utilisation*. Harcourt: Sydney, Australia.

Benbenishty, R., Davidson-Arad, B., Lopez, M., Devaney, J., Spratt, T., Koopmans, C. ... Hayes, D. (2015). Decision making in child protection: An international comparative study on maltreatment substantiation, risk assessment and intervention recommendations, and the role of professionals' child welfare attitudes. *Child Abuse and Neglect*, 49, 63-75.

Betancourt, J.R., Green, A.R., Carrillo, J.E., & Ananeh-Firempong. (2003). Defining cultural competence: a practical framework for addressing racial/ethnic disparities in health and health care. *Public Health Rep.*, 118(4), 293-302.

Besharov, D.J. (1981). Towards better research on child abuse and neglect: Making definitional issues an explicit methodological concern. *Child Abuse & Neglect*, 5(4), 383-391.

Blair, G., & K. Imai. (2012). Statistical analysis of list experiments. *Political Analysis*, 20, 47-77.

Bottoms, B. L., Shaver, P. R., Goodman, G. S., & Qin, J. (1995). In the name of God: A profile of religion-related child abuse. *Journal of Social Issues*, 51(2), 85–111.

Boyden, J., Pankhurst, A., & Tafere, Y. (2012). Child protection and harmful traditional practices: female early marriage and genital modification in Ethiopia. *Development in Practice*, 22(4), 510-522.

Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. London: Sage.

British Psychological Society. (2016, retrieved 22 April 2019). Achieving representation in psychology. *The Psychologist*.  
[<https://thepsychologist.bps.org.uk/volume-29/april/representativeness-psychologists-bps-response>]

Britner, P.A., & Mossler, D.G. (2002). Professionals' decision-making about out-of-home placements following instances of child abuse. *Child Abuse & Neglect*, 26(4), 317-32.

Cardiff University. (2019, accessed 29 March 2019). Health Evidence Bulletin Wales: critical appraisal checklists.  
[<https://www.cardiff.ac.uk/specialist-unit-for-review-evidence/resources/critical-appraisal-checklists>].

CASP, NHS. Critical Appraisal Skills Programme (CASP): appraisal tools. Public Health Resource Unit, NHS 2003.

Castilho L.S., Ferreira, E.F., & Perini, E. (2009). Perceptions of adolescents and young people regarding endemic dental fluorosis: psychosocial suffering. *Health and Social Care in the Community*, 17(6), 557-563.

- Christou, T., & Fowles, S. (2015). Failure to protect girls from female genital mutilation. *The Journal of Criminal Law*, 79(5), 344-357.
- Cook, K. (2016). Female genital mutilation in the UK population: a serious crime. *The Journal of Criminal Law* 2016, 80(2), 88–96.
- Costello, S., Quinn, M., Tatchell, A., Jordan, L., & Neophytou, K. (2015). In the best interests of the child: preventing female genital cutting (FGC). *British Journal of Social Work*, 45, 1259-1276.
- Christou, T., & Fowles, S. (2015). Failure to protect girls from female genital mutilation. *The Journal of Criminal Law*, 79(5), 344-357.
- De Cao, E., & Lutz, C. (2015). Measuring attitudes regarding female genital mutilation through a list experiment. *Technical report*.
- De Houwer, J., & Moors, A. (2010). Implicit measures: Similarities and differences. In: Gawronski, B., & Payne, B.K., eds. *Handbook of implicit social cognition: measurement, theory, and applications*. New York: The Guilford Press; 2010.
- Department of Health (1999), *National Service Framework for Mental Health*, Department of Health, London.
- Department of Health (2000), *NHS Plan*, Department of Health, London.
- Dustin, M. (2010). Female genital mutilation/cutting in the UK: challenging the inconsistencies. *European Journal of Women's Studies*, 17(1), 7-23.
- Earp, B. D. (2015). Female genital mutilation and male circumcision: toward an autonomy-based ethical framework. *Medicolegal and Bioethics*, 5, 89-104.
- Earp, B. D. (2016). Between moral relativism and moral hypocrisy: reframing the debate on "FGM." *Kennedy Institute of Ethics Journal*, 26(2), 105-44.
- Egger, M, Schneider, M, & Smith, G, D. (1998). Meta-analysis spurious precision? Meta-analysis of observational studies. *British Medical Journal*, 316(7125), 140-144.
- Egu, C. L., & Weiss, D. J. (2003). The role of race and severity of abuse in teachers' recognition or reporting of child abuse. *Journal of Child and Family Studies*, 12(4), 465–474.
- Elmusharaf, S., Elhadi, N., & Almroth, L. (2006). Reliability of self-reported form of female genital mutilation and WHO classification: Cross-sectional study. *British Medical Journal*, 333, 124-129.

- English, D.J. (1997). LONGSCAN Investigators. Modified Maltreatment Classification System (MMCS). Available at: <http://www.iprc.unc.edu/longscan/>.
- Feldman-Jacobs, C., & Clifton, D. (2010). *Female genital mutilation/cutting: data and trends*. Population Reference Bureau, USAID.
- Finlayson, L. M., & Koocher, G. P. (1991). Professional judgment and child abuse reporting in sexual abuse cases. *Professional Psychology: Research and Practice*, 22(6), 464-472.
- Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, 20(2), 303-315.
- Fontes, L., & Plummer, C. (2010). Cultural issues in disclosures of child sexual abuse. *Journal of Child Sexual Abuse*, 19, 491-518.
- Gibeau, A.M. (1998). Female genital mutilation: when a cultural practice generates clinical and ethical dilemmas. *Journal of obstetric, gynecologic, & neonatal nursing*, 27(1), 85-91.
- Gilbert, R., Widom, C. S., Browne, K., Fergusson, D., Webb, E., & Janson, S. (2009). Burden and consequences of child maltreatment in high-income countries. *Lancet*, 373, 68-81.
- Glachan, M. (1991) Child abuse: a social and cultural phenomenon. *Early Child Development and Care*, 74(1), 95-102.
- Glynn, A. N. (2010). What can we learn with statistical truth serum? Design and analysis of the list experiment (Technical report). Cambridge, MA: Harvard University.
- Gough, D. (1996). Defining the problem. *Child Abuse & Neglect*, 20(11), 993-1002.
- Graham, J. C., Detlaff, A. J., Bauman, D. J., & Fluke, J. (2015). The decision making ecology of placing a child into foster care: A structural equation model. *Child Abuse & Neglect*, 49, 12-23.
- Guiné, A., & Fuentes, F. J. M. (2007). Engendering redistribution, recognition, and representation: the case of female genital mutilation (FGM) in the United Kingdom and France. *Polit. Soc.* 35(3), 477-519.
- Hall, W., Chapman, M., Lee, K., Merino, Y., Thomas, T., Payne, B., Eng, E., Day, S., Coyne-Beasley, T. (2015). Implicit racial/ethnic bias among health care professionals and its influence on health care outcomes: a systematic review. *American Journal of Public Health*, 105(12), E60-76.

- Hammond, L. (2013). Somali transnational activism and integration in the UK: mutually supporting strategies. *Journal of Ethnic and Migration Studies*, 39(6), 1001-1017.
- Hampton, R. L., & Newberger, E. H. (1985). Child abuse incidence and reporting by hospitals: Significance of severity, class, and race. *American Journal of Public Health*, 75, 56-60.
- Hansen, D. J., Bumby, K. M., Lundquist, L. M., Chandler, R. M., Le, P. T., & Futa, K. T. (1997). The influence of case and professional variables on the identification and reporting of child maltreatment: a study of licensed psychologists and certified Masters social workers. *Journal of Family Violence*, 12(3), 313-332.
- Hewlett, S., De Wit, M., Richards, P., et al. (2006). Patients and professionals as research partners: challenges, practicalities and benefits. *Arthritis Rheum*, 55, 676-80.
- HM Government. (2016, retrieved 02 November 2017). Multi-agency statutory guidance on female genital mutilation. [[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/512906/Multi\\_Agency\\_Statutory\\_Guidance\\_on\\_FGM\\_-\\_FINAL.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/512906/Multi_Agency_Statutory_Guidance_on_FGM_-_FINAL.pdf)]
- Horton, N. J., & Kleinman, K. P. (2007). Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. *Am Statistician*, 61(1), 79-90.
- Hoyano L., & Keenan, C. (2007). *Child Abuse: Law and policy across boundaries*. Oxford, UK: Oxford University Press.
- Huang, C., Liao, H., & Chang, S. (1998). Social desirability and the clinical self-report inventory: methodological reconsideration. *Journal of Clinical Psychology*, 54(4), 517-528.
- Hughes, R. (1998). Considering the vignette technique and its application to a study of drug injecting and HIV risk and safer behaviour. *Sociology of health & illness*, 20(3), 381-400.
- Ibanez, E.S., Borrego, J., Pemberton, J.R., & Terao, S. (2006). Cultural factors in decision-making about child physical abuse: Identifying reporter characteristics influencing reporting tendencies. *Child Abuse & Neglect*, 30, 1365-1379.
- Jackson, H., & Nuttall, R. (1994). Effects of gender, age, and history of abuse on social workers' judgments of sexual abuse allegations. *Social Work Research*, 18(2), 105-114.

- Jefferson, M., (2015). FGM/cutting: contextualising recent legal developments. *The Journal of Criminal Law*, 79(6), 411–421.
- Johnsdotter, S. (2019). Meaning well while doing harm: compulsory genital examinations in Swedish African girls. *Sexual and Reproductive Health Matters*, 27(2), 1-13.
- Kellner, N. (1993). Under the knife: female genital mutilation as child abuse. *Journal of Juvenile Law*, 14, 118-132.
- Kemoli, A.M., & Mavindu, M. (2014). Child abuse: a classic case report with literature review. *Contemporary Clinical Dentistry*, 5(2), 256-259.
- Kenny, M. C. (2001). Child abuse reporting: Teachers' perceived deterrents. *Child Abuse & Neglect*, 25(1), 81–92.
- Kesner, J., Kwon, K.A., & Lim, C. (2016). The role of race in preservice teachers' perceptions of and attitudes towards corporal punishment & child maltreatment. *Teaching and Teacher Education*, 59, 318–326.
- Kesner, J., & Stenhouse, V. (2018). Investigating the potential effect of race and culture on preservice teachers' perceptions of corporal punishment and its subsequent effect on mandated reporting. *Australian Journal of Teacher Education*, 43(11), 70-83.
- King, G., Reece, R., Bendel, R., & Patel, V. (1998). The effects of sociodemographic variable, training and attitudes on the lifetime reporting practices of mandated reporters. *Child Maltreatment*, 3(3), 276-283.
- King, M., & Bruner, G. (2000). Social desirability bias: a neglected aspect of validity testing. *Psychology and Marketing*, 17(2), 79-103.
- Kirmayer, L. J. (2012). Rethinking cultural competence. *Transcultural Psychiatry*, 49(2), 149-164.
- Kuklinski, J.H., Cobb, M.D., & Gilens, M. (1997). Racial attitudes and the "New South". *The Journal of Politics* 59, 323-349.
- LaBrie, J, W., & Earleywine, M. (2000). Sexual risk behaviors and alcohol: higher base rates revealed using the unmatched-count technique. *Journal of Sex Research*, 37, 321-26.
- Larson, L. (1996). Female genital mutilation in the United States: child abuse or constitutional freedom? *Women's Rights Law Reporter*, 17(2), 237-257.
- Levi, B.H., & Crowell, K. (2011). Child abuse experts disagree about the threshold for mandated reporting. *Clinical Pediatric*, 50(4) 321-329.

- Liberati, A. (1995). "Meta-analysis: statistical alchemy for the 21st century": discussion. A plea for a more balanced view of meta-analysis and systematic overviews of the effect of health care interventions. *J Clin Epidemiol.*, *48*(1), 81-86.
- Lindsey, D. (1991). Factors affecting the foster care placement decision. *American Journal of Orthopsychiatry*, *61*, 272-281.
- London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2005). Disclosure of child sexual abuse: What does the research tell us about the ways that children tell? *Psychology, Public Policy, & the Law*, *11*, 194-226.
- London safeguarding children board. (2009, accessed 02 November 2017). London female genital mutilation resource pack. [[https://www.reducingtherisk.org.uk/cms/sites/reducingtherisk/files/folders/resources/fgm/resource\\_pack.pdf/](https://www.reducingtherisk.org.uk/cms/sites/reducingtherisk/files/folders/resources/fgm/resource_pack.pdf/)].
- Lopez, S.R. (1989). Patient variable biases in clinical judgment: conceptual overview and methodological considerations. *Psychological Bulletin*, *106*(2), 184-203.
- López, M., Fluke, J. D., Benbenishty, R., & Knorth, E. J. (2015). Commentary on decision-making and judgments in child maltreatment prevention and response: An overview. *Child Abuse & Neglect*, *49*, 1-11.
- Lopez, S., & Hernandez, P. (1986). How culture is considered in the evaluation of mental health patients. *Journal of Nervous and Mental Disease*, *174*, 598-606.
- Macfarlane, A., & Dorkenoo, E. (2015). *Prevalence of female genital mutilation in England and Wales: National and local estimates*. London: City University London and Equality Now.
- Maiter, S. (2004). Considering context and culture in child protection services to ethnically diverse families: an example from research with parents from the Indian subcontinent (South-Asians). *Journal of Social Work Research and Evaluation*, *5*(1), 63-80.
- Mandel, D.R., Lehman, D.R., & Yuille, J.C. (1995). Reasoning about the removal of a child from home: A comparison of police officers and social workers. *Journal of Applied Social Psychology*, *25*(10), 906-921.
- Mathers, N., & Rymer, J. (2015). Mandatory reporting of female genital mutilation by healthcare professionals. *The British Journal of General Practice*, *65*(635), 282-283.
- May-Chahal, C., & Cawson, P. (2005). Measuring child maltreatment in the United Kingdom: a study of the prevalence of child abuse and neglect. *Child Abuse & Neglect*, *29*, 969-984.

- McAuley, L., Pham, B., Tugwell, P., & Moher, D. (2000). Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet*, *356*, 1228-1231.
- Mick, D. G. (1996). Are studies of dark side variables confounded by socially desirable responding? The case of materialism. *Journal of Consumer Research*, *23*, 106-119.
- Minogue, V., Boness, J., Brown, A., & Girdlestone, J. (2005). The impact of service user involvement in research. *Int J Health Care Qual Assur*, *18*(2), 103-12.
- Mensch, B. S., & Kandel, D. B. (1988). Underreporting of substance use in a national longitudinal youth cohort. *Public Opinion Quarterly*, *52*, 100-124.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. The PRISMA Group. (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement*. *British Medical Journal*, *339*, b2535.
- Moody, G., Cannings-John, R., Hood, K., Kemp, A., & Robling, M. (2018). Establishing the international prevalence of self-reported child maltreatment: a systematic review by maltreatment type and gender. *BMC Public Health*, *18*(1), 1-15.
- Mortel, T. F. (2008). Faking it: social desirability response bias in self-report research. *Australian Journal of Advanced Nursing*, *25*(4), 40-48.
- Mtezuka, M. (1996). Issues of race and culture in child abuse. In B. Fawcett, B. Featherstone, J. Hearn, & C. Toft (Eds.), *Violence and gender relations: Theories and interventions* (p. 171-177). London, England: Sage.
- Mulrow, C., Langhorne, P., & Grimshaw, J. (1997). Integrating heterogeneous pieces of evidence in systematic reviews. *Ann Intern Med*, *127*(11), 989-995.
- Munro, E. (2005) Improving practice: Child protection as a systems problem. *Children and Youth Services Review*, *27*, 375-391.
- Muthumbi, J., Svanemyr, J., Scolaro, E., Temmerman, M., & Say, L. (2015). Female genital mutilation: a literature review of the current status of legislation and policies in 27 African countries and Yemen. *African Journal of Reproductive Health*, *19*, 32-40.
- Nieto, S., & Bode, P. (2011). *Affirming diversity: The sociopolitical context of multicultural education (6th ed.)*. Boston: Pearson Allyn and Bacon.
- Oakley, L., Kinmond, K., Humphreys, J., & Dioumd, M. (2017). Practitioner and communities' awareness of CALFB: Child abuse linked to faith or belief. *Child Abuse & Neglect*, *72*, 276-282.

- Paradies, Y., Truong, M., & Priest, N. (2013). A systematic review of the extent and measurement of healthcare provider racism. *Journal of General Internal Medicine*, 29(2), 364-387.
- Pećnik, N., & Brunnberg, E. (2005). Professionals' characteristics, victim's gender, and case assessments as predictors of professional judgments in child protection. *Review of Psychology*, 12(2), 133–146.
- Pereda, N., Guilera, G., Forns, M., & Gómez-Benito, J. (2009). The prevalence of child sexual abuse in community and student samples: a meta-analysis. *Clin Psychol Rev*, 29, 328-338.
- Pfeffer N. (1998). Theories of race, ethnicity and culture. *BMJ (Clinical research ed.)*, 317(7169), 1381–1384.
- Pinderhughes, E. E., Dodge, K. A., Bates, J. E., Pettit, G. S., & Zelli, A. (2000). Discipline responses: influences of parents' socioeconomic status, ethnicity, beliefs about parenting, stress, and cognitive-emotional processes. *Journal of Family Psychology*, 14, 380-400.
- Portwood, S. G. (1999). Coming to terms with a consensual definition of child maltreatment. *Child Maltreatment*, 4(1), 56-68.
- Price, E. G., Beach, M. C., Gary, T. L., Robinson, K. A., Gozu, A., Palacio, A., ... Cooper, L. A. (2005). A systematic review of the methodological rigor of studies evaluating cultural competence training of health professionals. *Academic Medicine*, 80, 578-586.
- Public Health Wales Observatory (PHWO), Gwasanaeth Tystiolaeth Evidence Service. (2014, accessed 29 March 2019). Questions to assist with the critical appraisal of a cross-sectional study (Type IV evidence). [[http://www2.nphs.wales.nhs.uk:8080/PubHObservatoryProjDocs.nsf/\(\\$All\)/E7B0C80995DC1BA380257DB80037C699/\\$File/Cross%20sectional%20study%20checklist.docx?OpenElement](http://www2.nphs.wales.nhs.uk:8080/PubHObservatoryProjDocs.nsf/($All)/E7B0C80995DC1BA380257DB80037C699/$File/Cross%20sectional%20study%20checklist.docx?OpenElement)].
- Rahman, A., & Toubia, N. (2000). *Female genital mutilation: A practical guide to worldwide laws & policies*. London, England: Zed Books.
- Ranganath, K.A., Smith, C.T., & Nosek, B, A. (2008). Distinguishing automatic and controlled components of attitudes from direct and indirect measurement methods. *J Exp. Soc. Psychol.*, 44(2), 386-396.
- Ray, S., Conn, K., Liebeschuetz, S., Costoli, V., & Tan, M. (2013). The impact of media reporting of high-profile cases on child protection medical assessments. *Child Abuse Review*, 22(1), 20-24.

- Reyes-García, V., Paneque-Gálvez, J., Luzb A., Guezeb, M., Maciád, M., Orta-Martínez, M., & Pinof, J. (2014). Cultural change and traditional ecological knowledge. An empirical analysis from the Tsimane' in the Bolivian Amazon. *Hum Organ*, 73(2), 162-173.
- Richerson, P., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago: University of Chicago Press.
- Riva, G. (2009). Virtual reality: an experiential tool for clinical psychology. *British Journal of Guidance & Counselling*, 37 (3), 337-345.
- Rubin, G. B. (1992). Multicultural considerations in the application of child protection laws. *Journal of Social Distress and the Homeless*, 1(3/4), 249-271.
- Saravanan, S., Kalyani, V., Vijayarani, M.P., Jayakodi, P., Felix, J., Arunmozhi, P., et al. (2008). Caries prevalence and treatment needs of rural school children in Chidambaram Taluk, Tamil Nadu, South India. *Indian Journal of Dental Research*, 19 (3), 186-90.
- Sieracki, J.H., Fuller, A.K., Leon, S.C., Jhe Bai, G., & Bryant, F.B. (2015). The role of race, socioeconomic status, and System of Care services in placement decision-making. *Children and Youth Services Review*, 50, 3-11.
- Tang, C.S.K. (1998). The rate of physical child abuse in Chinese families: a community survey in Hong Kong. *Child Abuse & Neglect*, 22(5), 381-91.
- Terao, S.Y., Borrego, Jr. J., & Urquiza, A.J. (2001). A reporting and response model for culture and child maltreatment. *Child Maltreatment*, 6(2), 158-168.
- Tilden, V.P., Schmidt, T.A., Limandri, B.J., Chiodo, G.T., Garland, M.J., & Loveless, P.A. (1994). Factors that influence clinicians' assessment and management of family violence. *American Journal of Public Health*, 84, 628-633.
- Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133(5), 859-883.
- Tsuchiya, T., Hirai, Y., & Ono, S. (2007). A Study of the Properties of the Item Count Technique. *Public Opinion Quarterly* 71, 253-272.
- Turbett, J. P., & O'Toole, R. (1983). Teachers' recognition and reporting of child abuse. *Journal of School Health*, 53(10), 605-609.
- United Nations Children's Fund (2013, accessed 11 September 2017). Female genital mutilation/cutting: a statistical overview and exploration of the dynamics of change. [[https://www.unicef.org/publications/index\\_69875.html](https://www.unicef.org/publications/index_69875.html)].

Varenne, B., Petersen, P.E., & Ouattara S. (2004). Oral health status of children and adults in urban and rural areas of Burkina Faso, Africa. *International Dental Journal*, 54, 83-89.

Vogt, S., Efferson, C., & Fehr, E. (2017). The risk of female genital cutting in Europe: comparing immigrant attitudes toward uncut girls with attitudes in a practicing country. *SSM – Population Health* 3, 283-293.

Weightman, A., Barker, J., & Lancaster, J. (2000). Health Evidence Bulletins Wales Project Methodology 3. Cardiff: UWCM.

Whaley, A. L., & Davis, K. E. (2007). Cultural competence and evidence based practice in mental health services: A contemporary perspective. *American Psychologist*, 62, 563-574.

Whitehorn, J., Ayonrinde, O. & Maingay, S. (2002). Female genital mutilation: cultural and psychological implications. *Sexual and Relationship Therapy*, 17 (2), 161–170.

Wissow, L. S. (2001). Ethnicity, income, and parenting contexts of physical punishment in a national sample of families with young children. *Child Maltreatment*, 6, 118-129.

World Health Organization. (1999, accessed 2 January 2019). Report of the consultation on child abuse prevention. Geneva: WHO. [<http://apps.who.int/iris/handle/10665/65900>].

World Health Organization (2008, accessed 31 October 2017). Eliminating female genital mutilation: an interagency statement. OHCHR, UNAIDS, UNDP, UNECA, UNESCO, UNFPA, UNHCR, UNICEF, UNIFEM, WHO. [[http://www.un.org/womenwatch/daw/csw/csw52/statements\\_missions/Interagency\\_Statement\\_on\\_Eliminating\\_FGM.pdf](http://www.un.org/womenwatch/daw/csw/csw52/statements_missions/Interagency_Statement_on_Eliminating_FGM.pdf)].

World Health Organization. (2014, retrieved 9 January 2019). Global status reports on violence prevention 2014. [[https://www.who.int/violence\\_injury\\_prevention/violence/status\\_report/2014/en/](https://www.who.int/violence_injury_prevention/violence/status_report/2014/en/)]

Wright, R. W., Brand, R. A., Dunn, W., & Spindler, K. P. (2007). How to write a systematic review. *Clin Orthop Relat Res.*, 455, 23-29.

Zeanah, C. H., & Humphreys, K. L., (2018). Child abuse and neglect. *Journal of the American Academy of Child & Adolescent Psychiatry*, 57(9), 637-644.

Zellman, G. L. (1990a). Child abuse reporting and failure to report among mandated reporters: prevalence, incidence, and reasons. *Journal of Interpersonal Violence*, 5(1), 3-22.

Zellman, G. L. (1990b). Linking schools and social services: the case of child abuse reporting. *Educational Evaluation and Policy Analysis*, 12(1), 41–55.

Zellman, G. (1992). The impact of case characteristics on child abuse reporting decisions. *Child abuse and neglect: the international journal*, 16(1), 57-74.

Zellman, G. L., & Faller, K. C. (1996). Reporting of child maltreatment. In J. Briere, L. Berliner J. A. Bulkley, C. Jenny, & T. Reid (Eds.) *The APSAC handbook on child maltreatment* (pp. 359-381). Thousand Oaks, CA: Sage Publications, Inc.

Zurynski, Y. A., Sureshkumar, P., Phu, A., & Elliott, E. (2015). Female genital mutilation and cutting: A systematic literature review of health professionals' knowledge, attitudes and clinical practice. *BMC International Health and Human Rights*, 15(32), 1-18.

## **Appendices**

## Appendix A

### Public Health Wales Observatory Critical Appraisal Tool for Cross-Sectional Studies (2014).

#### Questions to assist with the critical appraisal of a cross-sectional study (Type IV evidence)<sup>1</sup>

#### Paper citation:

#### A. What is this paper about?

	Yes	Can't tell	No
<b>1. Is the study relevant to the needs of the Project?</b>			
<b>2. Does the paper address a clearly focused issue? In terms of:</b> <ul style="list-style-type: none"> <li>• Aims of the investigation?</li> <li>• Setting (location and dates)?</li> <li>• The population studied?</li> <li>• The variables measured?</li> </ul>			
<b>3. Is the choice of study method appropriate to the study question?</b>			

#### Is it worth continuing? (Delete as appropriate)

#### YES/NO/Discuss

Only complete the next section if the answer to the question above was 'Yes'

#### B. Can I trust this paper?

	Yes	Can't tell	No
<b>4. Is the population studied appropriate?</b>			
• Was the sample representative of			

<sup>1</sup> Sources used: Critical Appraisal Skills Programme (CASP, Anglia and Oxford RHA) questions and Polgar A, Thomas SA. Chapter 22. Critical evaluation of published research in Introduction to research in the health sciences. 3rd edition. Melbourne: Churchill Livingstone, 1995; Undertaking systematic reviews of research on effectiveness. University of York: NHS Centre for Reviews & Dissemination, 2001; Weightman AL, Barker, JM, Lancaster J. Health Evidence Bulletins Wales Project Methodology 3. Cardiff: UWCM, 2000.

its target population?			
<ul style="list-style-type: none"> <li>• How was the sample selected? Random, stratified?</li> </ul>			
<ul style="list-style-type: none"> <li>• If appropriate, was a power calculation made?</li> </ul>			
<b>5. Have confounding and bias been considered?</b>			
<ul style="list-style-type: none"> <li>• Did the study achieve a good response rate?</li> </ul>			
<ul style="list-style-type: none"> <li>• Were rigorous processes used to develop the questions? (E.g. were the questions piloted/validated?)</li> </ul>			
<b>6. Is the study design and/or execution flawed to the extent that the results are unreliable?</b>			

**Is it worth continuing? (Delete as appropriate)**

**YES/NO/Discuss**

Only complete the next two sections if the answer to the question above was 'Yes'

**C. What did they find?**

	<b>Yes</b>	<b>Can't tell</b>	<b>No</b>
<b>7. Are tables/graphs adequately labelled and understandable?</b>			
<b>8. Are you confident with the authors' choice and use of statistical methods, if employed?</b> <ul style="list-style-type: none"> <li>• If sub-group/interactions analyses have been undertaken is there an explanation of how/why sub-groups have been formed?</li> <li>• Is there an explanation of how potential confounding factors have been controlled for?</li> <li>• Is there an explanation of how missing data have been handled?</li> </ul>			

<ul style="list-style-type: none"> <li>• Are both unadjusted and adjusted (i.e. for confounding) results given if appropriate?</li> <li>• Is the precision of estimates (95% CI) given?</li> <li>• Do you believe the results?</li> </ul>			
---	--	--	--

**D. Are the results relevant locally?**

	<b>Yes</b>	<b>Can't tell</b>	<b>No</b>
<p><b>9. Can the results be applied to the local situation?</b> Consider differences between the local and study populations (e.g. cultural, geographical, ethical) which could affect the relevance of the study.</p>			
<p><b>10. Were all important variables considered?</b></p>			
<p><b>11. Accept for further use as Type IV evidence?</b></p>			

**If the answer to question 11 above was 'Yes' then record this study as 'Included' and proceed to data extraction:**

**Comments:**

**Appendix B**  
**NHS Health Research Authority Ethics Approval (IRAS project ID 247542)**



Ymchwil Iechyd  
a Gofal Cymru  
Health and Care  
Research Wales



Miss Fatoumata Jatta  
Trainee Clinical Psychologist  
Camden and Islington NHS Foundation Trust  
Royal Holloway University of London  
Egham Hill  
Egham  
Surrey  
TW20 0EX  
[fatoumata.jatta.2016@live.rhul.ac.uk](mailto:fatoumata.jatta.2016@live.rhul.ac.uk)

Email: [hra.approval@nhs.net](mailto:hra.approval@nhs.net)

27 July 2018

Dear Miss Jatta

**HRA and Health and Care  
Research Wales (HCRW)  
Approval Letter**

**Study title:** Influence of culture and socioeconomic status on the mandatory reporting of female genital mutilation (FGM) by health and social care professionals and teachers.

**IRAS project ID:** 247542

**Sponsor:** Royal Holloway, University of London

I am pleased to confirm that [HRA and Health and Care Research Wales \(HCRW\) Approval](#) has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

**How should I continue to work with participating NHS organisations in England and Wales?**  
You should now provide a copy of this letter to all participating NHS organisations in England and Wales, as well as any documentation that has been updated as a result of the assessment.

Participating NHS organisations in England and Wales **will not** be required to formally confirm capacity and capability before you may commence research activity at site. As such, you may commence the research at each organisation 35 days following sponsor provision to the site of the local information pack, so long as:

- You have contacted participating NHS organisations (see below for details)
- The NHS organisation has not provided a reason as to why they cannot participate
- The NHS organisation has not requested additional time to confirm.

You may start the research prior to the above deadline if the site positively confirms that the research may proceed.

Page 1 of 7

If not already done so, you should now provide the [local information pack](#) for your study to your participating NHS organisations. A current list of R&D contacts is accessible at the [NHS RD Forum website](#) and these contacts MUST be used for this purpose. After entering your IRAS ID you will be able to access a password protected document (password: **Summer14**). The password is updated on a monthly basis so please obtain the relevant contact information as soon as possible; please do not hesitate to contact me should you encounter any issues.

Commencing research activities at any NHS organisation before providing them with the full local information pack and allowing them the agreed duration to opt-out, or to request additional time (unless you have received from their R&D department notification that you may commence), is a breach of the terms of HRA and HCRW Approval. Further information is provided in the “*summary of assessment*” section towards the end of this document.

It is important that you involve both the research management function (e.g. R&D office) supporting each organisation and the local research team (where there is one) in setting up your study. Contact details of the research management function for each organisation can be accessed [here](#).

#### **How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?**

HRA and HCRW Approval does not apply to NHS/HSC organisations within the devolved administrations of Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report (including this letter) has been sent to the coordinating centre of each participating nation. You should work with the relevant national coordinating functions to ensure any nation specific checks are complete, and with each site so that they are able to give management permission for the study to begin.

Please see [IRAS Help](#) for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

#### **How should I work with participating non-NHS organisations?**

HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to [obtain local agreement](#) in accordance with their procedures.

#### **What are my notification responsibilities during the study?**

The attached document “*After HRA Approval – guidance for sponsors and investigators*” gives detailed guidance on reporting expectations for studies with HRA and HCRW Approval, including:

- Registration of Research
- Notifying amendments
- Notifying the end of the study

The [HRA website](#) also provides guidance on these topics and is updated in the light of changes in reporting expectations or procedures.

#### **I am a participating NHS organisation in England or Wales. What should I do once I receive this letter?**

IRAS project ID	247542
-----------------	--------

You should work with the applicant and sponsor to complete any outstanding arrangements so you are able to confirm capacity and capability in line with the information provided in this letter.

The sponsor contact for this application is as follows:

Name: Annette Lock  
Tel: 01784 414 388  
Email: [annette.lock@rhul.ac.uk](mailto:annette.lock@rhul.ac.uk)

**Who should I contact for further information?**

Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is **247542**. Please quote this on all correspondence.

Yours sincerely

**Gemma Oakes**  
**Assessor**

Email: [hra.approval@nhs.net](mailto:hra.approval@nhs.net)

Copy to: *Ms Annette Lock, Royal Holloway, University of London [Sponsor Contact]*  
[annette.lock@rhul.ac.uk](mailto:annette.lock@rhul.ac.uk)  
*Ms Mabel Salli, Noclor NHS [Lead NHS R&D Contact]*  
[mabel.salli@nhs.net](mailto:mabel.salli@nhs.net)

### List of Documents

The final document set assessed and approved by HRA and HCRW Approval is listed below.

<i>Document</i>	<i>Version</i>	<i>Date</i>
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only) [Evidence of Sponsor indemnity]		26 July 2017
IRAS Application Form [IRAS_Form_13062018]		13 June 2018
IRAS Application Form XML file [IRAS_Form_13062018]		13 June 2018
IRAS Checklist XML [Checklist_13062018]		13 June 2018
IRAS Checklist XML [Checklist_27072018]		27 July 2018
Non-validated questionnaire [Online Questionnaire]	2	27 July 2018
Non-validated questionnaire [Non-validated questionnaire]	1	31 May 2018
Other [Summary CV for supervisor ]	1	11 May 2018
Other [Summary CV for supervisor ]		
Other [Proposal approval]	1	22 March 2018
Other [Reference list]	3	31 May 2018
Participant consent form [Participant Consent Form]	4	27 July 2018
Participant information sheet (PIS) [Participant Information Sheet]	5	27 July 2018
Research protocol or project proposal [Research Proposal/Protocol]	4	27 July 2018
Summary CV for Chief Investigator (CI) [Summary CV for Chief Investigator (CI)]		24 May 2018
Summary CV for student [Summary CV for student]		24 May 2018
Summary CV for supervisor (student research) [Summary CV for supervisor ]		22 May 2018
Summary, synopsis or diagram (flowchart) of protocol in non-technical language [Diagram or Procedure]	2	27 July 2018

## Appendix C

### NHS Health Research Authority (IRAS project ID 247542) Non-Substantial Amendment Approval

**From:** TSITSIPA, Eirini (CENTRAL AND NORTH WEST LONDON NHS FOUNDATION TRUST)  
<eirini.tsitsipa1@nhs.net>  
**Sent:** 08 November 2018 12:06  
**To:** Jatta, Fatoumata (2016)  
**Cc:** Lock, Annette; NOCLOR, Contact (CENTRAL AND NORTH WEST LONDON NHS FOUNDATION TRUST)  
**Subject:** 247542 - Approval of NSA1-C&I

Dear **Fatoumata Jatta**,

<b>Study title:</b>	Influence of culture and socioeconomic status on the mandatory reporting of female genital mutilation (FGM) by health and social care professionals and teachers.		
<b>IRAS ID:</b>	247542	<b>REC number:</b>	REC Review Exempt
<b>Amendment No /Sponsor Ref:</b>	NSA1		
<b>Date amendment submitted to HRA</b>	16 October 2018		
<b>Amendment type</b>	Non Substantial		
<b>Brief Summary</b>	<ul style="list-style-type: none"><li>• Online questionnaire and diagram of procedure (case vignettes and questions amended)</li><li>• Research/project proposal (case vignettes and questions amended, and stratification removed)</li></ul>		

Following review of the amendment for the above study, Camden & Islington NHS Foundation Trust confirm continued capacity and capability.

The amendment can therefore be implemented at this site under the existing HRA Approval in accordance with sponsor instructions.

Kind regards,  
Eirini

**Eirini Tsitsipa**  
Research Facilitator  
Direct: 020 76852897 Team: 020  
76855949

## Appendix D

### Section 1 of the Online Survey: Demographics (Gender, Age, Profession, Training Level, FGM Training, Race, and Parents' Country of Origin)

Q1. What is your gender?

- Male
- Female

Q2. What is your age?

- 16-20
- 21-25
- 26-30
- 31-35
- 36-40
- 41-45
- 46-50
- 51-55
- 56-60
- 61-65
- 66-70
- 71-75
- 76-80

Q3. What is your profession?

- Art/Drama/Music Therapist
- Biomedical Scientist
- Chiropractor
- Chiropodist/Podiatrist
- Clinical Scientist
- Dentist
- Dietitian
- Doctor/Medical Practitioner/Physician
- Hearing Aid Dispenser
- Midwife
- Nurse
- Occupational Therapist
- Operating Department Practitioner
- Optician
- Orthoptist
- Osteopath
- Paramedic
- Pharmacist
- Physiotherapist
- Prosthetist/Orthotist
- Psychiatrist
- Psychologist - Clinical
- Psychologist - Counselling

- Psychologist - Educational
- Psychologist - Forensic
- Psychologist - Health
- Psychologist - Occupational
- Psychologist - Sport and Exercise
- Radiographer
- Social Care Worker
- Social Worker
- Speech and Language Therapist
- Teacher

Q4. What is your professional training level?

- Qualified (5+ years)
- Newly qualified (under 5 years)
- Trainee

Q5. How many hours of FGM training have you received to date?

- None
- 1-2
- 3-4 (half a day)
- 5-6
- 7-8 (one day)
- 9-10
- 11-12 (one and a half days)
- 13-14
- 15-16 (two days)
- 17-18
- 19-20 (two and a half days)
- 21-22
- 23-24 (three days)
- 25-26
- 27-28 (three and a half days)
- 29-30
- 31-32 (four days)
- 33-34
- 35-36 (four and a half days)
- 37-38
- 39-40 (five days/one week)
- 40+ (over a week)

Q6. What is your race?

- White
- Black/ African/Caribbean/Black British
- Asian/Asian British
- Mixed
- Other

Q7. What country was your father born in?

Q8. What country was your mother born in?

- Afghanistan

- Albania
- Algeria
- Andorra
- Angola
- Anguilla
- Antigua & Barbuda
- Argentina
- Armenia
- Australia
- Austria
- Azerbaijan
- Bahamas
- Bahrain
- Bangladesh
- Barbados
- Belarus
- Belgium
- Belize
- Benin
- Bermuda
- Bhutan
- Bolivia
- Bosnia & Herzegovina
- Botswana
- Brazil
- Brunei Darussalam
- Bulgaria
- Burkina Faso
- Burma / Myanmar
- Burundi
- Cambodia
- Cameroon
- Canada
- Cape Verde
- Cayman Islands
- Central African Republic
- Chad
- Chile
- China
- Colombia
- Comoros
- Congo
- Costa Rica
- Croatia
- Cuba
- Cyprus

- Czech Republic
- Democratic Republic of the Congo
- Denmark
- Djibouti
- Dominica
- Dominican Republic
- Ecuador
- Egypt
- El Salvador
- Equatorial Guinea
- Estonia
- Ethiopia
- Fiji
- Finland
- France
- French Guiana
- Gabon
- Gambia
- Georgia
- Germany
- Ghana
- Greece
- Grenada
- Guadeloupe
- Guatemala
- Guinea
- Guinea-Bissau
- Guyana
- Haiti
- Honduras
- Hungary
- Iceland
- India
- Indonesia
- Iran
- Iraq
- Israel and the Occupied Territories
- Italy
- Ivory Coast (Cote d'Ivoire)
- Jamaica
- Japan
- Jordan
- Kazakhstan
- Kenya
- Kosovo
- Kuwait

- Kyrgyz Republic (Kyrgyzstan)
- Laos
- Latvia
- Lebanon
- Lesotho
- Liberia
- Libya
- Liechtenstein
- Lithuania
- Luxembourg
- Republic of Macedonia
- Madagascar
- Malawi
- Malaysia
- Maldives
- Mali
- Malta
- Martinique
- Mauritania
- Mauritius
- Mayotte
- Mexico
- Moldova, Republic of
- Monaco
- Mongolia
- Montenegro
- Montserrat
- Morocco
- Mozambique
- Myanmar (formerly Burma)
- Namibia
- Nepal
- Netherlands
- New Zealand
- Nicaragua
- Niger
- Nigeria
- Korea, Democratic Republic of (North Korea)
- Norway
- Oman
- Pacific Islands
- Pakistan
- Panama
- Papua New Guinea
- Paraguay
- Peru

- Phillipines
- Poland
- Portugal
- Puerto Rico
- Qatar
- Reunion
- Romania
- Russia
- Rwanda
- Saint Kitts and Nevis
- Saint Lucia
- Saint Vincent's & Grenadines
- Samoa
- Sao Tome and Principe
- Saudi Arabia
- Senegal
- Serbia
- Seychelles
- Sierra Leone
- Singapore
- Slovakia
- Slovenia
- Solomon Islands
- Somalia
- South Africa
- South Korea
- South Sudan
- Spain
- Sri Lanka
- Sudan
- Suriname
- Swaziland
- Sweden
- Switzerland
- Syria
- Tajikistan
- Tanzania
- Thailand
- Timor Leste
- Togo
- Trinidad & Tobago
- Tunisia
- Turkey
- Turkmenistan
- Turks & Caicos Islands
- Uganda

- Ukraine
- United Arab Emirates
- United Kingdom
- United States of America (USA)
- Uruguay
- Uzbekistan
- Venezuela
- Vietnam
- Virgin Islands (UK)
- Virgin Islands (US)
- Yemen
- Zambia
- Zimbabwe

*NB. Ireland was mistakenly omitted from the list.*

## Appendix E

### **The List Experiment**

The list experiment is an indirect questioning technique, which provides respondents some amount of privacy protection via aggregation (Glynn, 2010, p. 2). The idea behind a list experiment is that if a sensitive question is asked indirectly, the respondent is more likely to reveal a truthful answer (De Cao & Lutz, 2015, p. 3). Used mainly in political science to examine voters' attitudes and attitudes about race, the list experiment has also been shown to reduce social desirability bias when researching attitudes on FGM (De Cao & Lutz, 2015).

Using the list experiment means that respondents are asked how many of a list of questions apply to them. As long as the entire list does not apply to them, they can be assured that we will not know their answer to the sensitive question (Glynn, 2010, p. 2).

Participants were randomly assigned to either the control group or the experimental group (between-subjects). Each participant irrespective of group was presented with all four vignettes in a randomly assigned order. The control group however did not answer any FGM-sensitive items.

The control group respondents received a list of four non-sensitive, yes/no items and were asked how many of the listed items they agree on, but were told they do not have to state which items (see below). The treatment group respondents instead received the same list of four non-sensitive, yes/no items, plus a FGM-sensitive yes/no item ('are you concerned enough to report this family directly to the police, rather than continuing to monitor them within your service?'). As with the control group, the treatment group respondents were also asked the number of items they agree on, but without saying which (see below).

Each vignette was presented in turn in a randomly assigned order with a list of four non-sensitive yes/no items for the control group, and with a list of four non-sensitive yes/no items **plus** the FGM-sensitive yes/no item for the treatment group. Four separate lists of four non-sensitive yes/no items were developed and piloted. Each vignette was always presented with the same list. Responses to the lists (specifically, the number of yes responses) constituted the dependent variable. The difference in the total number of items between control and treatment group identified the proportion of respondents that agree with the sensitive item (De Cao & Lutz, 2015, p. 3).

### **Survey Section 2: Vignettes presented with a list experiment Experimental Group**

#### ***Vignette 1 (Low Cultural Integration x Low SES).***

Please carefully read the vignette below AND the list of 4\* questions below it.

Halima and Ahmed were both born in Somalia. They have strong ties to their Somali community and try to visit friends and family in Somalia when they can. They recently arrived in London, and live in Whitechapel in east London with their 12 year old daughter Zahra.

Having both left school early, Ahmed works long hours as a security guard at a supermarket while Halima looks after the home.

Most of their friends are Somali, and they also speak Somali at home.

During the course of your work with Zahra (their 12 year old daughter), she tells you that she was cut during a special celebration. However when asked she has not wanted to give any further details.

1. Should children be encouraged to play outside?
2. Should the law protect against animal cruelty?
3. Should water be more expensive than soft drinks?
4. Should animals today be killed for their fur?
5. Are you concerned enough to report this family directly to the police, rather than continuing to monitor them within your service?

We appreciate that some of these questions may be difficult to answer. In real life you may want to gather further information before making a decision. Nonetheless, please attempt to give us your most likely answer based on the information provided.

Please respond with HOW MANY of the questions you have ANSWERED YES to. We do not want to know which questions, just HOW MANY.

- Answered yes to 1 question
- Answered yes to 2 questions
- Answered yes to 3 questions
- Answered yes to 4 questions
- Answered yes to 5 questions

***Vignette 2 (High Cultural Integration x High SES).***

Please carefully read the vignette below AND the list of 4\* questions below it.

Fatima and Mohamed both have a Somali background, but were born in London and have few ties to their Somali community. They have never been to Somalia.

They live in Notting Hill in west London with their 12 year old daughter Hawa.

Having both gone to university in London, Mohamed works long hours as a banker for an investment bank while Fatima looks after the home.

Most of their friends are English, and they also speak English at home.

During the course of your work with Hawa (their 12 year old daughter), she tells you that she was cut during a special celebration. However when asked she has not wanted to give any further details.

1. Should people be encouraged to wear seat-belts?
2. Should women be banned from serving in the army?
3. Should corporal (physical) punishment be allowed in schools?
4. Should women be allowed the right to an abortion?

5. Are you concerned enough to report this family directly to the police, rather than continuing to monitor them within your service?

We appreciate that some of these questions may be difficult to answer. In real life you may want to gather further information before making a decision. Nonetheless, please attempt to give us your most likely answer based on the information provided.

Please respond with HOW MANY of the questions you have ANSWERED YES to. We do not want to know which questions, just HOW MANY.

- Answered yes to 1 question
- Answered yes to 2 questions
- Answered yes to 3 questions
- Answered yes to 4 questions
- Answered yes to 5 questions

***Vignette 3 (High Cultural Integration x Low SES).***

Please carefully read the vignette below AND the list of 4\* questions below it.

Maryam and Ismail both have a Somali background, but were born in London and have few ties to their Somali community. They have never been to Somalia.

They live in Whitechapel in east London with their 12 year old daughter Yasmin.

Having both left school early, Ismail works long hours as a security guard at a supermarket while Maryam looks after the home.

Most of their friends are English, and they also speak English at home.

During the course of your work with Yasmin (their 12 year old daughter), she tells you that she was cut during a special celebration. However when asked she has not wanted to give any further details.

1. Do you like spending time with your friends?
2. Should there be a right to free health care?
3. Do you regularly sleep-walk?
4. Should the government make it more difficult for people in need to receive benefits?
5. Are you concerned enough to report this family directly to the police, rather than continuing to monitor them within your service?

We appreciate that some of these questions may be difficult to answer. In real life you may want to gather further information before making a decision. Nonetheless, please attempt to give us your most likely answer based on the information provided.

Please respond with HOW MANY of the questions you have ANSWERED YES to. We do not want to know which questions, just HOW MANY.

- Answered yes to 1 question
- Answered yes to 2 questions
- Answered yes to 3 questions
- Answered yes to 4 questions
- Answered yes to 5 questions

***Vignette 4 (Low Cultural Integration x High SES).***

Please carefully read the vignette below AND the list of 4\* questions below it.

Khadija and Hasan were both born in Somalia. They have strong ties to their Somali community and try to visit friends and family in Somalia when they can. They recently moved to London after Hasan was transferred here for work, and live in Notting Hill in west London with their 12 year old daughter Amina. Having both gone to university in Somalia, Hasan works long hours as a banker who caters to wealthy Somalis while Khadija looks after the home. Most of their friends are Somali, and they also speak Somali at home. During the course of your work with Amina (their 12 year old daughter), she tells you that she was cut during a special celebration. However when asked she has not wanted to give any further details.

1. Should employees have regular days off work such as weekends?
2. Is the torture of prisoners acceptable?
3. Do you like receiving calls from telemarketers?
4. Should the death penalty be banned?
5. Are you concerned enough to report this family directly to the police, rather than continuing to monitor them within your service?

We appreciate that some of these questions may be difficult to answer. In real life you may want to gather further information before making a decision. Nonetheless, please attempt to give us your most likely answer based on the information provided.

Please respond with HOW MANY of the questions you have ANSWERED YES to. We do not want to know which questions, just HOW MANY.

- Answered yes to 1 question
- Answered yes to 2 questions
- Answered yes to 3 questions
- Answered yes to 4 questions
- Answered yes to 5 questions

\*This was an error as respondents are presented with 5 questions below the vignettes and not 4.

## **Control Group**

### ***Vignette 1 (Low Cultural Integration x Low SES).***

Please carefully read the vignette below AND the list of 4 questions below it.

Halima and Ahmed were both born in Somalia. They have strong ties to their Somali community and try to visit friends and family in Somalia when they can.

They recently arrived in London, and live in Whitechapel in east London with their 12 year old daughter Zahra.

Having both left school early, Ahmed works long hours as a security guard at a supermarket while Halima looks after the home.

Most of their friends are Somali, and they also speak Somali at home.

During the course of your work with Zahra (their 12 year old daughter), she tells you that she was cut during a special celebration. However when asked she has not wanted to give any further details.

1. Should children be encouraged to play outside?
2. Should the law protect against animal cruelty?
3. Should water be more expensive than soft drinks?
4. Should animals today be killed for their fur?

We appreciate that some of these questions may be difficult to answer. In real life you may want to gather further information before making a decision. Nonetheless, please attempt to give us your most likely answer based on the information provided.

Please respond with HOW MANY of the questions you have ANSWERED YES to. We do not want to know which questions, just HOW MANY.

- Answered yes to 1 question
- Answered yes to 2 questions
- Answered yes to 3 questions
- Answered yes to 4 questions

### ***Vignette 2 (High Cultural integration x High SES).***

Please carefully read the vignette below AND the list of 4 questions below it.

Fatima and Mohamed both have a Somali background, but were born in London and have few ties to their Somali community. They have never been to Somalia.

They live in Notting Hill in west London with their 12 year old daughter Hawa.

Having both gone to university in London, Mohamed works long hours as a banker for an investment bank while Fatima looks after the home.

Most of their friends are English, and they also speak English at home.

During the course of your work with Hawa (their 12 year old daughter), she tells you that she was cut during a special celebration. However when asked she has not wanted to give any further details.

1. Should people be encouraged to wear seat-belts?
2. Should women be banned from serving in the army?
3. Should corporal (physical) punishment be allowed in schools?
4. Should women be allowed the right to an abortion?

We appreciate that some of these questions may be difficult to answer. In real life you may want to gather further information before making a decision. Nonetheless, please attempt to give us your most likely answer based on the information provided.

Please respond with HOW MANY of the questions you have ANSWERED YES to. We do not want to know which questions, just HOW MANY.

- Answered yes to 1 question
- Answered yes to 2 questions
- Answered yes to 3 questions
- Answered yes to 4 questions

***Vignette 3 (High Cultural Integration x Low SES).***

Please carefully read the vignette below AND the list of 4 questions below it.

Maryam and Ismail both have a Somali background, but were born in London and have few ties to their Somali community. They have never been to Somalia. They live in Whitechapel in east London with their 12 year old daughter Yasmin. Having both left school early, Ismail works long hours as a security guard at a supermarket while Maryam looks after the home. Most of their friends are English, and they also speak English at home. During the course of your work with Yasmin (their 12 year old daughter), she tells you that she was cut during a special celebration. However when asked she has not wanted to give any further details.

1. Do you like spending time with your friends?
2. Should there be a right to free health care?
3. Do you regularly sleep-walk?
4. Should the government make it more difficult for people in need to receive benefits?

We appreciate that some of these questions may be difficult to answer. In real life you may want to gather further information before making a decision. Nonetheless, please attempt to give us your most likely answer based on the information provided.

Please respond with HOW MANY of the questions you have ANSWERED YES to. We do not want to know which questions, just HOW MANY.

- Answered yes to 1 question
- Answered yes to 2 questions
- Answered yes to 3 questions
- Answered yes to 4 questions

***Vignette 4 (Low Cultural Integration x High SES).***

Please carefully read the vignette below AND the list of 4 questions below it.

Khadija and Hasan were both born in Somalia. They have strong ties to their Somali community and try to visit friends and family in Somalia when they can. They recently moved to London after Hasan was transferred here for work, and live in Notting Hill in west London with their 12 year old daughter Amina. Having both gone to university in Somalia, Hasan works long hours as a banker who caters to wealthy Somalis while Khadija looks after the home. Most of their friends are Somali, and they also speak Somali at home. During the course of your work with Amina (their 12 year old daughter), she tells you that she was cut during a special celebration. However when asked she has not wanted to give any further details.

1. Should employees have regular days off work such as weekends?
2. Is the torture of prisoners acceptable?
3. Do you like receiving calls from telemarketers?
4. Should the death penalty be banned?

We appreciate that some of these questions may be difficult to answer. In real life you may want to gather further information before making a decision. Nonetheless, please attempt to give us your most likely answer based on the information provided.

Please respond with HOW MANY of the questions you have ANSWERED YES to. We do not want to know which questions, just HOW MANY.

- Answered yes to 1 question
- Answered yes to 2 questions
- Answered yes to 3 questions
- Answered yes to 4 questions

**Survey section 3: Vignettes presented with the FGM-sensitive question (for both the Experimental AND Control Groups)**

***Vignette 1 (Low Cultural Integration x Low SES)***

Please carefully read the vignette below AND the question below it.

Halima and Ahmed were both born in Somalia. They have strong ties to their Somali community and try to visit friends and family in Somalia when they can.

They recently arrived in London, and live in Whitechapel in east London with their 12 year old daughter Zahra.

Having both left school early, Ahmed works long hours as a security guard at a supermarket while Halima looks after the home.

Most of their friends are Somali, and they also speak Somali at home.

During the course of your work with Zahra (their 12 year old daughter), she tells you that she was cut during a special celebration. However when asked she has not wanted to give any further details.

1. Are you concerned enough to report this family directly to the police, rather than continuing to monitor them within your service?

We appreciate that this question may be difficult to answer. In real life you may want to gather further information before making a decision. Nonetheless, please attempt to give us your most likely answer based on the information provided.

- Yes
- No

***Vignette 2 (High Cultural Integration x High SES)***

Please carefully read the vignette below AND the question below it.

Fatima and Mohamed both have a Somali background, but were born in London and have few ties to their Somali community. They have never been to Somalia.

They live in Notting Hill in west London with their 12 year old daughter Hawa.

Having both gone to university in London, Mohamed works long hours as a banker for an investment bank while Fatima looks after the home.

Most of their friends are English, and they also speak English at home.

During the course of your work with Hawa (their 12 year old daughter), she tells you that she was cut during a special celebration. However when asked she has not wanted to give any further details.

1. Are you concerned enough to report this family directly to the police, rather than continuing to monitor them within your service?

We appreciate that this question may be difficult to answer. In real life you may want to gather further information before making a decision. Nonetheless, please attempt to give us your most likely answer based on the information provided.

- Yes

- No

### ***Vignette 3 (High Cultural Integration x Low SES)***

Please carefully read the vignette below AND the question below it.

Maryam and Ismail both have a Somali background, but were born in London and have few ties to their Somali community. They have never been to Somalia.

They live in Whitechapel in east London with their 12 year old daughter Yasmin.

Having both left school early, Ismail works long hours as a security guard at a supermarket while Maryam looks after the home.

Most of their friends are English, and they also speak English at home.

During the course of your work with Yasmin (their 12 year old daughter), she tells you that she was cut during a special celebration. However when asked she has not wanted to give any further details.

1. Are you concerned enough to report this family directly to the police, rather than continuing to monitor them within your service?

We appreciate that this question may be difficult to answer. In real life you may want to gather further information before making a decision. Nonetheless, please attempt to give us your most likely answer based on the information provided.

- Yes
- No

### ***Vignette 4 (Low Cultural Integration x High SES)***

Please carefully read the vignette below AND the question below it.

Khadija and Hasan were both born in Somalia. They have strong ties to their Somali community and try to visit friends and family in Somalia when they can.

They recently moved to London after Hasan was transferred here for work, and live in Notting Hill in west London with their 12 year old daughter Amina.

Having both gone to university in Somalia, Hasan works long hours as a banker who caters to wealthy Somalis while Khadija looks after the home.

Most of their friends are Somali, and they also speak Somali at home.

During the course of your work with Amina (their 12 year old daughter), she tells you that she was cut during a special celebration. However when asked she has not wanted to give any further details.

1. Are you concerned enough to report this family directly to the police, rather than continuing to monitor them within your service?

We appreciate that this question may be difficult to answer. In real life you may want to gather further information before making a decision. Nonetheless, please attempt to give us your most likely answer based on the information provided.

- Yes
- No

## Appendix F

### Information Sheet



**IRAS Project ID 247542; Version no. 5; Dated 27.7.18**  
**PARTICIPANT INFORMATION SHEET**  
**PROJECT TITLE: FEMALE GENITAL MUTILATION (FGM) AND THE**  
**MANDATORY REPORTING DUTY**

Chief Investigator: Fatoumata Jatta  
Sponsor Organisation: Royal Holloway, University of London  
Primary Academic Supervisor: Dr Charles Efferson  
Secondary Academic Supervisor: Dr Jane Vosper

*We are researchers at Royal Holloway, University of London, and we would like to invite you to take part in a study exploring professionals' decision-making. This research study is being conducted as part of an educational qualification (Doctoral Programme in Clinical Psychology, Royal Holloway, University of London).*

*This information sheet is to help you decide if you would like to take part. Before you decide, it is important that you understand why the research is being done and what it would involve for you. Please read this information carefully, and feel free to discuss this with others if you wish. If there is anything that is not clear, or you would like to know more, please contact Fatoumata Jatta via email at [fatoumata.jatta.2016@live.rhul.ac.uk](mailto:fatoumata.jatta.2016@live.rhul.ac.uk) or telephone on 01784 414 012 (this is the university answer machine, you will be asked to say whom the message is for and to leave your name and contact details so that Fatoumata can call you back).*

#### **What is the purpose of the study?**

The purpose of this study is to explore the factors that influence a professionals' decision whether or not to report cases of female genital cutting.

#### **Why have I been invited?**

You have been invited to take part because you are a health and social care professional or teacher working in the UK. We aim to involve around 150 health and social care professionals and teachers aged between 18 and 65.

#### **Do I have to take part?**

No, you do not have to take part in this study. You can withdraw from the study at any time, without giving a reason. If you decide that you do not wish to take part this will not affect your employment or legal rights in any way.

**What will the study involve?**

If you decide to take part in the research, you will be invited to complete an online survey. The first section of the survey will ask you to provide some background information about yourself. In the second section, you will be presented with four case vignettes in a randomly assigned order. After each vignette, you will be presented with a number of questions that may or may not relate to the vignette. You will then simply be asked the total number of questions with which you agree. In the third section, you will again be presented with the same four case vignettes, followed by a question, and asked if you agree with each of the four questions.

**What will happen if I decide not to carry on with the study?**

That is fine and it will not affect you in any way. If you wish to leave the study you will need to contact the research team via email: [fatoumata.jatta.2016@live.rhul.ac.uk](mailto:fatoumata.jatta.2016@live.rhul.ac.uk). This will mean that the data that you have given us will not be used in the study.

**What are the potential disadvantages of taking part?**

The survey will take approximately fifteen minutes of your time to complete. There are no known risks to taking part in this study.

**What are the potential benefits of taking part?**

There is no personal benefit to participating in this study. By taking part in the study you will improve knowledge of the impact of the 2015 mandatory reporting duty on professionals' practice. You will help to increase our understanding of factors that influence professional decision-making, and ultimately, help professionals like yourself to fulfil their reporting obligations more consistently.

**Will the data provided by myself be kept confidential?**

All privacy laws and procedures will be followed during all elements of this study. Information collected from you during the study will be kept confidential and safe. Only members from the research team will have access to your data.

You will only ever be identified by an identification number rather than by name. None of the information that you provide us will be attached to your name, and the results from the study will not be linked to any identifiable information. Being an online study, data will be stored encrypted and password protected on the secure survey platform Qualtrics. And if necessary, an encryption key that adheres to NHS confidentiality standards will be used. When the study has finished, data which has only been collected for the purpose of this research will be stored in a locked cupboard and destroyed after 5 years.

**Who has reviewed the study?**

The research is being led by Fatoumata Jatta under the supervision of Dr Charles Efferson (Lecturer in Psychology), Dr Jane Vosper (Lecturer in Clinical Psychology / Principle Clinical Psychologist), and Dr Lih-Mei Liao (Consultant Clinical Psychologist). The study is being funded by Royal Holloway, University of London, as part of the

doctorate programme in Clinical Psychology. This study has been reviewed and approved by the NHS Health Research Authority and the Royal Holloway, University of London Research Ethics Committee.

**What will happen to the results of this study, and how will I be informed?**

The research will be submitted in partial fulfilment of a doctorate degree in Clinical Psychology. For participants who opt-in to be informed of the results of the study, overall findings will be fed back via email. We aim to publish results in a peer-reviewed journal. Results may also be used for training and information-sharing purposes in relevant services and at mental health conferences. The published data will be anonymised and no participants will be identified.

**What if there is a problem?**

If you have any concerns about any aspects of the research, you can contact the research supervisor Dr Charles Efferson via email at [charles.efferson@rhul.ac.uk](mailto:charles.efferson@rhul.ac.uk). If you have any further problems or complaints about the study then please contact Dr Gary Brown, Senior

Lecturer in Clinical Psychology at Royal Holloway, University of London by email at [gary.brown@rhul.ac.uk](mailto:gary.brown@rhul.ac.uk).

**What happens next?**

If you decide to take part in the study you will need to read the consent statement and provide your consent to participate. You can do this by clicking the link to the study below. Please take your time to think about whether you would like to do this and please ask any questions that you have.

**How do I find out more?**

If you would like to know more about the study, please contact Fatoumata Jatta, Trainee Clinical Psychologist, via email at [fatoumata.jatta.2016@live.rhul.ac.uk](mailto:fatoumata.jatta.2016@live.rhul.ac.uk).

**GDPR TRANSPARENCY STATEMENT**

Royal Holloway, University of London is the sponsor for this study based in the United Kingdom. We will be using information from you in order to undertake this study and will act as the data controller for this study. This means that we are responsible for looking after your information and using it properly. Royal Holloway, University of London will keep identifiable information about you for 5 years after the study has finished.

Your rights to access, change or move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. If you withdraw from the study, we will keep the information about you that we have already obtained. To safeguard your rights, we will use the minimum personally-identifiable information possible.

You can find out more about how we use your information at:

<https://www.royalholloway.ac.uk/about-us/more/governance-and-strategy/data-protection/>.

As a university we use personally-identifiable information to conduct research to improve health, care and services. As a publicly-funded organisation, we have to ensure that it is in the public interest when we use personally-identifiable information from people who have agreed to take part in research. This means that when you agree to take part in a research study, we will use your data in the ways needed to conduct and analyse the research study. Your rights to access, change or move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. If you withdraw from the study, we will keep the information about you that we have already obtained. To safeguard your rights, we will use the minimum personally-identifiable information possible.

Health and care research should serve the public interest, which means that we have to demonstrate that our research serves the interests of society as a whole. We do this by following the [UK Policy Framework for Health and Social Care Research](#).

If you wish to raise a complaint on how we have handled your personal data, you can contact our Data Protection Officer who will investigate the matter. If you are not satisfied with our response or believe we are processing your personal data in a way that is not lawful you can complain to the Information Commissioner's Office (ICO).

You can contact our Data Protection Officer at [dataprotection@rhul.ac.uk](mailto:dataprotection@rhul.ac.uk).

Royal Holloway, University of London will collect information from you for this research study in accordance with our instructions.

Royal Holloway, University of London will use your name and contact details to contact you about the research study, and make sure that relevant information about the study is recorded for your care, and to oversee the quality of the study. Individuals from Royal Holloway, University of London and regulatory organisations may look at your research records to check the accuracy of the research study. Camden and Islington NHS Foundation Trust will pass these details to Royal Holloway, University of London along with the information collected from you. The only people in Royal Holloway who will have access to information that identifies you will be people who need to contact you to invite you to participate in the study or to audit the data collection process. The people who analyse the information will not identify you, nor your name or contact details.

Royal Holloway, University of London will keep identifiable information about you from this study for 5 years after the study has finished.

**Thank you for taking the time to read this information sheet.**

## Appendix G

### Consent Form



IRAS Project ID: 247542/ Version 4; Dated 27.7.18

#### **PARTICIPANT CONSENT FORM**

#### **PROJECT TITLE: FEMALE GENITAL MUTILATION (FGM) AND THE MANDATORY REPORTING DUTY**

Chief Investigator: Fatoumata Jatta

Sponsor Organisation: Royal Holloway, University of London

Primary Academic Supervisor: Dr Charles Efferson

Secondary Academic Supervisor: Dr Jane Vosper

Participant no. \_\_\_\_\_

*This is a consent form. Before you decide, it is important that you understand why the research is being done and what it would involve for you. Please read the information sheet for this project carefully, and feel free to discuss this with others if you wish. If there is anything that is not clear, or you would like to know more, please contact Fatoumata Jatta via email at [fatoumata.jatta.2016@live.rhul.ac.uk](mailto:fatoumata.jatta.2016@live.rhul.ac.uk) or telephone on 01784 414 012 (this is the university answer machine, you will be asked to say whom the message is for and to leave your name and contact details so that Fatoumata can call you back).*

#### Statement by participant

1. I confirm that I have read and understood the information sheet for the above study. I have been informed of the purpose, risks, and benefits of taking part. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.
2. I understand what my involvement will entail and any questions have been answered to my satisfaction.
3. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my legal rights being affected.
4. I understand that all information obtained will be confidential.
5. I understand that the information collected about me may be used in an anonymous form to support research. It will not be possible for me to be identified by it.
6. Contact information has been provided should I wish to seek further information from the investigator at any time for purposes of clarification.
7. I agree to take part in this study.

**DO YOU AGREE TO GIVE CONSENT TO PARTICIPATE IN THIS STUDY AND CONTINUE TO THE ONLINE QUESTIONNAIRE?**

## Appendix H

### Debrief Sheet



IRAS Project ID 247542/ Version no. 1; Dated 20.7.18

#### PARTICIPANT DEBRIEF SHEET

#### **PROJECT TITLE: FEMALE GENITAL MUTILATION (FGM) AND THE MANADATORY REPORTING DUTY**

Chief Investigator: Fatoumata Jatta

Sponsor Organisation: Royal Holloway, University of London

Primary Academic Supervisor: Dr Charles Efferson

Secondary Academic Supervisor: Dr Jane Vosper

Field Supervisor: Dr Lih-Mei Liao

*Many thanks for completing this survey. Your participation in this research is very much appreciated.*

#### **What is the purpose of the study?**

This study is being written up and submitted in partial fulfilment of a doctorate degree in Clinical Psychology. The purpose of this study is to explore the factors that influence a professionals' decision whether or not to report cases of female genital mutilation (FGM), which has been unlawful in the United Kingdom since 1985. In 2015 a compulsory reporting duty was introduced requiring health and social care professionals and teachers to report *known* cases of FGM (i.e. where a girl discloses she has undergone FGM) in under 18-year-olds. Despite this, data suggest that FGM continues to be underreported.

Research suggests that there are a number of factors influencing a professional's decision to report child abuses such as FGM. Studies have found that professionals were more likely to consider issues of a cultural nature, and those involving a family of lower socioeconomic status (SES) as less serious. This study's primary research question is therefore whether cultural sensitivity and SES influence professionals' decision to report FGM. We hypothesised that clinicians would be less likely to report FGM where a family is less integrated into UK culture, and where the family is of higher SES.

#### **Why was the study designed in this way?**

Due to the compulsory reporting duty, social pressure was expected to play an important role in responses. Therefore, in order to measure professionals' true views, an indirect questioning technique known as a list experiment was used to allow participants to hide their individual answers following four vignettes in which case characteristics (level of integration into UK culture, SES) were systematically varied. Moreover, since other factors shown to be influential of professionals' reporting behaviour include their gender, training background (e.g. social work vs. medicine), and amount of training, these data were also collected to ascertain whether they play a similar role in the reporting of FGM.

### **How can I find out about the results of the study?**

If you are interested in hearing about the results and conclusions of the study, please inform the principal researcher via email (Fatoumata.Jatta.2016@live.rhul.ac.uk) who will send you a summary once the research is complete.

### **Where can I access further information on FGM?**

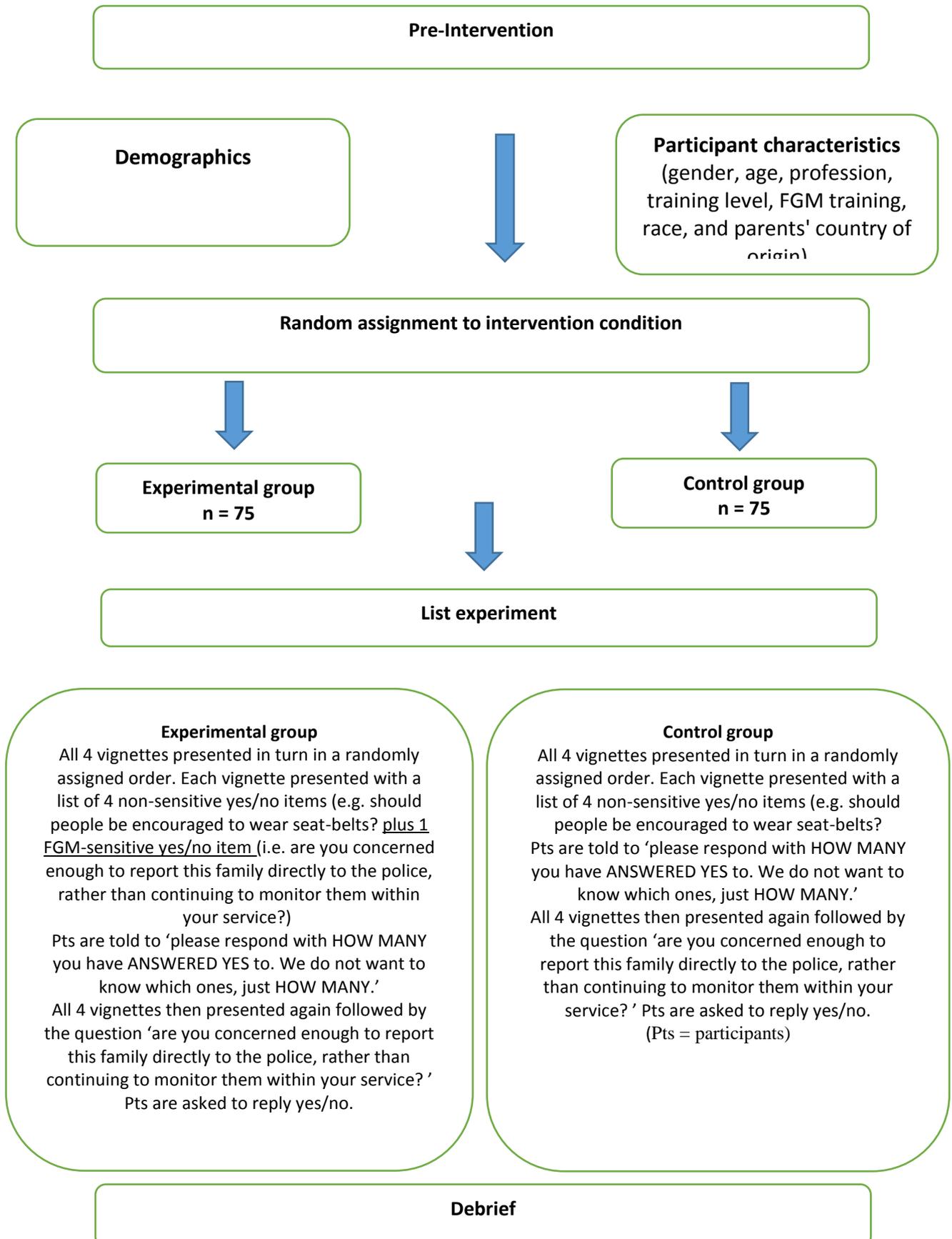
For further information on FGM and FGM training please visit:

- ❖ FGM specialist clinics (NHS England provide a full list), GPs, the police, and social services
- ❖ The Home Office free online learning package for professionals: <https://www.fgmelearning.co.uk/>. See also NSPCC (FGM Helpline), UNICEF, and FORWARD
- ❖ For government resources explaining the mandatory reporting duty please see <https://www.gov.uk/government/publications/fgm-mandatory-reporting-in-healthcare>
- ❖ Community agencies and charities e.g. online platform Magool (<https://www.themaven.net/magool/>), Daughters of Eve, Dahlia Project, Forward, Hawa Trust, and Family Action
- ❖ The Face of Defiance project for personal accounts of FGM

**Thank you for taking the time to read this information sheet.**

## Appendix I

### Diagram of procedure



## Appendix J

### Full Copy of Results of Analyses for the Empirical Study

results29April2019 (1)

R version 3.5.1 (2018-07-02) -- "Feather Spray"  
 Copyright (C) 2018 The R Foundation for Statistical Computing  
 Platform: x86\_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.  
 You are welcome to redistribute it under certain conditions.  
 Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.  
 Type 'contributors()' for more information and  
 'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or  
 'help.start()' for an HTML browser interface to help.  
 Type 'q()' to quit R.

```
> setwd('Dropbox/ip/fgcProfessionalsUK/data')
Error in setwd("Dropbox/ip/fgcProfessionalsUK/data") :
cannot change working directory
> setwd('.;./Dropbox/ip/fgcProfessionalsUK/data')
> source('importConfigureRunRegs.txt')
Loading required package: zoo
```

Attaching package: 'zoo'

The following objects are masked from 'package:base':

as.Date, as.Date.numeric

[1] "coefTest\_reg1\_list\_compCases:"

MODEL 1A

z test of coefficients:

	value	Pr(> z )	Estimate	Std. Error	z
(Intercept)	20.3185	< 2e-16 ***	2.0323139	0.1000227	
female	-0.1402	0.88850	-0.0077323	0.0551502	
clinPsychologist	0.3356	0.73714	0.0292914	0.0872685	
as.factor(professionalTrainingLevel)qualified	0.1480	0.88238	0.0128374	0.0867685	
as.factor(professionalTrainingLevel)trainee	0.1926	0.84728	0.0151438	0.0786298	
someTrainingFGC	0.5947	0.55204	0.0404043	0.0679405	
black	2.0374	0.04161 *	0.3006186	0.1475518	
notwhite	0.3623	0.71715	0.0343315	0.0947690	
oneOrMoreParentsFromPracticingCountry	-0.9159	0.35973	-0.1031290	0.1126001	
orderForPresentation	-1.3027	0.19268	-0.0192414	0.0147705	
vigHighIntegrationUK	0.9747	0.32971	0.0478540	0.0490964	
vigHighSES	-1.5005	0.13348	-0.0938783	0.0625640	
sensitiveList	10.9750	< 2e-16 ***	0.7940169	0.0723476	
I(vigHighIntegrationUK * vigHighSES)	0.5031	0.61487	0.0433395	0.0861392	
I(vigHighIntegrationUK * sensitiveList)	-0.0862	0.93131	-0.0067895	0.0787647	

**NB** STANDARD ERRORS ARE CLUSTER ROBUST. I.E. CLUSTERED BY PARTICIPANT.

```

                                results29April2019 (1)
I(vigHighSES * sensitiveList)      -0.1042058  0.1039893
-1.0021  0.31630
I(vigHighIntegrationUK * vigHighSES * sensitiveList)  0.1057215  0.1363756
0.7752  0.43821

```

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

```

[1] "Num complete cases: 725"
[1] "coefTest_reg2_list_compCases:"  MODEL 2A

```

z test of coefficients:

	value	Pr(> z )	Estimate	Std. Error	z
(Intercept)	20.3899	< 2e-16 ***	2.0270249	0.0994130	
female	-0.1141	0.90914	-0.0063009	0.0552089	
clinPsychologist	0.3062	0.75949	0.0267112	0.0872483	
as.factor(professionalTrainingLevel)qualified	0.1579	0.87450	0.0137066	0.0867825	
as.factor(professionalTrainingLevel)trainee	0.3132	0.75412	0.0242885	0.0775459	
someTrainingFGC	0.5750	0.56529	0.0391738	0.0681281	
black	1.8493	0.06442 .	0.2560564	0.1384648	
notwhite	0.1482	0.88215	0.0130456	0.0879981	
orderForPresentation	-1.3051	0.19184	-0.0192298	0.0147338	
vigHighIntegrationUK	0.9754	0.32938	0.0478532	0.0490618	
vigHighSES	-1.5086	0.13140	-0.0942076	0.0624468	
sensitiveList	11.0314	< 2e-16 ***	0.7953504	0.0720990	
I(vigHighIntegrationUK * vigHighSES)	0.5077	0.61166	0.0436712	0.0860166	
I(vigHighIntegrationUK * sensitiveList)	-0.0870	0.93069	-0.0068462	0.0787133	
I(vigHighSES * sensitiveList)	-1.0053	0.31478	-0.1044350	0.1038895	
I(vigHighIntegrationUK * vigHighSES * sensitiveList)	0.7734	0.43928	0.1053149	0.1361711	

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

```

[1] "Num complete cases: 725"
[1] "coefTest_reg3_list_compCases:"  MODEL 3A

```

z test of coefficients:

	value	Pr(> z )	Estimate	Std. Error	z
(Intercept)	40.0122	<2e-16 ***	2.0846583	0.0521005	
orderForPresentation	-1.3442	0.1789	-0.0198177	0.0147430	
vigHighIntegrationUK	0.9813	0.3264	0.0478943	0.0488068	
vigHighSES	-1.4674	0.1423	-0.0916279	0.0624414	
sensitiveList	11.3889	<2e-16 ***	0.8159560	0.0716451	
I(vigHighIntegrationUK * vigHighSES)	0.4778	0.6328	0.0409684	0.0857523	

```

                                results29April2019 (1)
I(vigHighIntegrationUK * sensitiveList) -0.0068352 0.0783648
-0.0872 0.9305
I(vigHighSES * sensitiveList) -0.1064067 0.1038534
-1.0246 0.3056
I(vigHighIntegrationUK * vigHighSES * sensitiveList) 0.1096498 0.1357534
0.8077 0.4193

```

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

[1] "Num complete cases: 725"

[1] "coefTest\_reg1\_basicQ\_compCases:" MODEL 4A

z test of coefficients:

	Estimate	Std. Error	z
value Pr(> z )			
(Intercept)	0.78117316	0.10715838	
7.2899 3.102e-13 ***			
female	-0.06331141	0.07965536	
-0.7948 0.42672			
clinPsychologist	-0.03033746	0.08182778	
-0.3707 0.71083			
as.factor(professionalTrainingLevel)qualified	0.06390933	0.08355266	
0.7649 0.44433			
as.factor(professionalTrainingLevel)trainee	0.04802592	0.09430789	
0.5092 0.61058			
someTrainingFGC	0.01464640	0.06808851	
0.2151 0.82968			
black	0.03497304	0.12163622	
0.2875 0.77371			
notwhite	-0.09808935	0.11283483	
-0.8693 0.38467			
oneOrMoreParentsFromPracticingCountry	0.13362571	0.08637290	
1.5471 0.12184			
orderForPresentation	-0.01243144	0.00585707	
-2.1225 0.03380 *			
vigHighIntegrationUK	0.01369573	0.02077770	
0.6592 0.50980			
vigHighSES	0.00089866	0.01738691	
0.0517 0.95878			
sensitiveList	0.05703271	0.06581118	
0.8666 0.38616			
I(vigHighIntegrationUK * vigHighSES)	-0.05178741	0.02448403	
-2.1152 0.03442 *			
I(vigHighIntegrationUK * sensitiveList)	-0.05696873	0.03407534	
-1.6718 0.09455 .			
I(vigHighSES * sensitiveList)	0.00156031	0.02303339	
0.0677 0.94599			
I(vigHighIntegrationUK * vigHighSES * sensitiveList)	0.04417052	0.03871740	
1.1408 0.25393			

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

[1] "Num complete cases: 697"

[1] "coefTest\_reg2\_basicQ\_compCases:" MODEL 5A

z test of coefficients:

	Estimate	Std. Error	z
value Pr(> z )			
(Intercept)	0.7882435	0.1071010	
7.3598 1.842e-13 ***			
female	-0.0659958	0.0792981	
-0.8323 0.40527			
clinPsychologist	-0.0260264	0.0818827	
-0.3178 0.75060			
as.factor(professionalTrainingLevel)qualified	0.0628194	0.0839333	
0.7484 0.45419			

```

                                results29April2019 (1)
as.factor(professionalTrainingLevel)trainee    0.0345078  0.0935212
0.3690  0.71214
someTrainingFGC                                0.0175019  0.0680804
0.2571  0.79712
black                                           0.0872544  0.1278793
0.6823  0.49504
notwhite                                       -0.0667090  0.1075168
-0.6205  0.53496
orderForPresentation                          -0.0124224  0.0058566
-2.1211  0.03392 *
vigHighIntegrationUK                          0.0136945  0.0207626
0.6596  0.50952
vigHighSES                                     0.0008980  0.0173740
0.0517  0.95878
sensitiveList                                 0.0555330  0.0662997
0.8376  0.40225
I(vigHighIntegrationUK * vigHighSES)          -0.0517848  0.0244663
-2.1166  0.03430 *
I(vigHighIntegrationUK * sensitiveList)       -0.0569680  0.0340504
-1.6730  0.09432 .
I(vigHighSES * sensitiveList)                0.0015592  0.0230162
0.0677  0.94599
I(vigHighIntegrationUK * vigHighSES * sensitiveList) 0.0441165  0.0387067
1.1398  0.25439

```

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

[1] "Num complete cases: 697"

[1] "coefTest\_reg3\_basic0\_compCases:" MODEL 6A

z test of coefficients:

	Estimate	Std. Error	z
value Pr(> z )			
(Intercept)	0.76580182	0.05113917	
14.9749 < 2e-16 ***			
orderForPresentation	-0.01249537	0.00581594	
-2.1485 0.03168 *			
vigHighIntegrationUK	0.01370421	0.02065642	
0.6634 0.50705			
vigHighSES	0.00090328	0.01728644	
0.0523 0.95833			
sensitiveList	0.05479169	0.06497125	
0.8433 0.39905			
I(vigHighIntegrationUK * vigHighSES)	-0.05180589	0.02434400	
-2.1281 0.03333 *			
I(vigHighIntegrationUK * sensitiveList)	-0.05697369	0.03387663	
-1.6818 0.09261 .			
I(vigHighSES * sensitiveList)	0.00156833	0.02289973	
0.0685 0.94540			
I(vigHighIntegrationUK * vigHighSES * sensitiveList)	0.04455222	0.03844471	
1.1589 0.24651			

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

[1] "Num complete cases: 697"

[1] "IMPORTANT NOTE: The ipw analysis analyzes the complete cases, as above, it just weights the observations according to rarity. Thus, the number of observations here will be exactly the same as the corresponding regressions above."

[1] "coefTest\_reg1\_list\_ipw:" MODEL 1B

z test of coefficients:

	Estimate	Std. Error	z
value Pr(> z )			
(Intercept)	2.0324875	0.1000309	
20.3186 < 2e-16 ***			

```

                                results29April2019 (1)
female                           -0.0078162  0.0551409
-0.1417  0.88728
clinPsychologist                 0.0293125  0.0872753
0.3359  0.73697
as.factor(professionalTrainingLevel)qualified  0.0127992  0.0867715
0.1475  0.88273
as.factor(professionalTrainingLevel)trainee  0.0150607  0.0786390
0.1915  0.84812
someTrainingFGC                 0.0404147  0.0679451
0.5948  0.55197
black                            0.3003272  0.1474859
2.0363  0.04172 *
notwhite                        0.0344474  0.0947744
0.3635  0.71626
oneOrMoreParentsFromPracticingCountry -0.1033664  0.1125608
-0.9183  0.35845
orderForPresentation            -0.0192685  0.0147804
-1.3036  0.19235
vigHighIntegrationUK            0.0478559  0.0490963
0.9747  0.32969
vigHighSES                      -0.0938753  0.0625659
-1.5004  0.13351
sensitiveList                   0.7940345  0.0723552
10.9741 < 2e-16 ***
I(vigHighIntegrationUK * vigHighSES)  0.0433307  0.0861405
0.5030  0.61495
I(vigHighIntegrationUK * sensitiveList) -0.0067903  0.0787639
-0.0862  0.93130
I(vigHighSES * sensitiveList)     -0.1041987  0.1039896
-1.0020  0.31634
I(vigHighIntegrationUK * vigHighSES * sensitiveList) 0.1057336  0.1363765
0.7753  0.43816

```

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

[1] "coefTest\_reg2\_list\_ipw:" **MODEL 2B**

z test of coefficients:

	Estimate	Std. Error	z
value Pr(> z )			
(Intercept)	2.0271955	0.0994211	
20.3900 < 2e-16 ***			
female	-0.0063817	0.0551995	
-0.1156 0.90796			
clinPsychologist	0.0267267	0.0872554	
0.3063 0.75937			
as.factor(professionalTrainingLevel)qualified	0.0136698	0.0867857	
0.1575 0.87484			
as.factor(professionalTrainingLevel)trainee	0.0242260	0.0775571	
0.3124 0.75476			
someTrainingFGC	0.0391814	0.0681339	
0.5751 0.56525			
black	0.2556622	0.1384531	
1.8466 0.06481 .			
notwhite	0.0131110	0.0879938	
0.1490 0.88155			
orderForPresentation	-0.0192603	0.0147448	
-1.3062 0.19147			
vigHighIntegrationUK	0.0478554	0.0490615	
0.9754 0.32935			
vigHighSES	-0.0942054	0.0624485	
-1.5085 0.13142			
sensitiveList	0.7953707	0.0721054	
11.0307 < 2e-16 ***			
I(vigHighIntegrationUK * vigHighSES)	0.0436625	0.0860175	
0.5076 0.61173			
I(vigHighIntegrationUK * sensitiveList)	-0.0068473	0.0787124	

results29April2019 (1)

-0.0870 0.93068  
 I(vigHighSES \* sensitiveList) -0.1044277 0.1038895  
 -1.0052 0.31481  
 I(vigHighIntegrationUK \* vigHighSES \* sensitiveList) 0.1053270 0.1361721  
 0.7735 0.43924

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

[1] "coefTest\_reg3\_list\_ipw:" MODEL 3B

z test of coefficients:

value Pr(> z )	Estimate	Std. Error	z
(Intercept)	2.0847178	0.0521032	
40.0113 <2e-16 ***			
orderForPresentation	-0.0198414	0.0147519	
-1.3450 0.1786			
vigHighIntegrationUK	0.0478959	0.0488067	
0.9813 0.3264			
vigHighSES	-0.0916286	0.0624419	
-1.4674 0.1423			
sensitiveList	0.8159520	0.0716457	
11.3887 <2e-16 ***			
I(vigHighIntegrationUK * vigHighSES)	0.0409641	0.0857528	
0.4777 0.6329			
I(vigHighIntegrationUK * sensitiveList)	-0.0068362	0.0783641	
-0.0872 0.9305			
I(vigHighSES * sensitiveList)	-0.1064013	0.1038529	
-1.0245 0.3056			
I(vigHighIntegrationUK * vigHighSES * sensitiveList)	0.1096556	0.1357535	
0.8078 0.4192			

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

[1] "coefTest\_reg1\_basicQ\_ipw:" MODEL 4B

z test of coefficients:

value Pr(> z )	Estimate	Std. Error	z
(Intercept)	0.78110426	0.10715653	
7.2894 3.114e-13 ***			
female	-0.06328941	0.07965116	
-0.7946 0.42686			
clinPsychologist	-0.03037048	0.08182636	
-0.3712 0.71052			
as.factor(professionalTrainingLevel)qualified	0.06389383	0.08354894	
0.7647 0.44442			
as.factor(professionalTrainingLevel)trainee	0.04806432	0.09430496	
0.5097 0.61028			
someTrainingFGC	0.01465004	0.06808610	
0.2152 0.82964			
black	0.03512905	0.12159037	
0.2889 0.77265			
notwhite	-0.09808628	0.11283654	
-0.8693 0.38470			
oneOrMoreParentsFromPracticingCountry	0.13371421	0.08633765	
1.5487 0.12145			
orderForPresentation	-0.01241410	0.00585360	
-2.1208 0.03394 *			
vigHighIntegrationUK	0.01369344	0.02077806	
0.6590 0.50987			
vigHighSES	0.00089741	0.01738700	
0.0516 0.95884			
sensitiveList	0.05701811	0.06581088	
0.8664 0.38627			
I(vigHighIntegrationUK * vigHighSES)	-0.05178239	0.02448286	

results29April2019 (1)

```

-2.1150 0.03443 *
I(vigHighIntegrationUK * sensitiveList) -0.05696739 0.03407513
-1.6718 0.09456 .
I(vigHighSES * sensitiveList) 0.00155813 0.02303295
0.0676 0.94607
I(vigHighIntegrationUK * vigHighSES * sensitiveList) 0.04416466 0.03871412
1.1408 0.25396

```

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

[1] "coefTest\_reg2\_basicQ\_ipw:" MODEL 5B

z test of coefficients:

value	Pr(> z )	Estimate	Std. Error	z
(Intercept)		0.78818709	0.10709950	
7.3594	1.848e-13 ***			
female		-0.06597560	0.07929387	
-0.8320	0.40539			
clinPsychologist		-0.02605656	0.08188171	
-0.3182	0.75032			
as.factor(professionalTrainingLevel)qualified		0.06280325	0.08393019	
0.7483	0.45429			
as.factor(professionalTrainingLevel)trainee		0.03453724	0.09351844	
0.3693	0.71190			
someTrainingFGC		0.01750744	0.06807812	
0.2572	0.79705			
black		0.08744502	0.12783827	
0.6840	0.49396			
notwhite		-0.06668520	0.10751929	
-0.6202	0.53512			
orderForPresentation		-0.01240825	0.00585321	
-2.1199	0.03401 *			
vigHighIntegrationUK		0.01369266	0.02076284	
0.6595	0.50959			
vigHighSES		0.00089698	0.01737413	
0.0516	0.95883			
sensitiveList		0.05551706	0.06629949	
0.8374	0.40239			
I(vigHighIntegrationUK * vigHighSES)		-0.05178070	0.02446508	
-2.1165	0.03430 *			
I(vigHighIntegrationUK * sensitiveList)		-0.05696693	0.03405017	
-1.6730	0.09432 .			
I(vigHighSES * sensitiveList)		0.00155740	0.02301579	
0.0677	0.94605			
I(vigHighIntegrationUK * vigHighSES * sensitiveList)		0.04411173	0.03870371	
1.1397	0.25440			

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

[1] "coefTest\_reg3\_basicQ\_ipw:" MODEL 6B

z test of coefficients:

value	Pr(> z )	Estimate	Std. Error	z
(Intercept)		0.76575661	0.05113264	
14.9759	< 2e-16 ***			
orderForPresentation		-0.01247707	0.00581260	
-2.1466	0.03183 *			
vigHighIntegrationUK		0.01370178	0.02065678	
0.6633	0.50713			
vigHighSES		0.00090196	0.01728655	
0.0522	0.95839			
sensitiveList		0.05479385	0.06497131	
0.8434	0.39903			
I(vigHighIntegrationUK * vigHighSES)		-0.05180060	0.02434269	

results29April2019 (1)

-2.1280 0.03334 \*  
I(vigHighIntegrationUK \* sensitiveList) -0.05697227 0.03387647  
-1.6818 0.09261 .  
I(vigHighSES \* sensitiveList) 0.00156603 0.02289914  
0.0684 0.94548  
I(vigHighIntegrationUK \* vigHighSES \* sensitiveList) 0.04454566 0.03844170  
1.1588 0.24654

---  
signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

[1] "vigHighSES: 0"  
[1] "vigHighIntegrationUK: 0"  
[1] "sensitiveList: 0"  
[1] "Mean (numYes, list exp): 2.03488372093023"  
[1] "Std dev (numYes, list exp): 0.284818803849029"  
[1] ""  
[1] "Mean (report, basic Q): 0.734939759036145"  
[1] "Std dev (report, basic Q): 0.444048375451825"  
[1] ""  
[1] ""  
[1] "vigHighSES: 0"  
[1] "vigHighIntegrationUK: 0"  
[1] "sensitiveList: 1"  
[1] "Mean (numYes, list exp): 2.854166666666667"  
[1] "Std dev (numYes, list exp): 0.632108695960412"  
[1] ""  
[1] "Mean (report, basic Q): 0.791208791208791"  
[1] "Std dev (report, basic Q): 0.408696668732407"  
[1] ""  
[1] ""  
[1] "vigHighSES: 0"  
[1] "vigHighIntegrationUK: 1"  
[1] "sensitiveList: 0"  
[1] "Mean (numYes, list exp): 2.08139534883721"  
[1] "Std dev (numYes, list exp): 0.350297899409993"  
[1] ""  
[1] "Mean (report, basic Q): 0.746987951807229"  
[1] "Std dev (report, basic Q): 0.437380603749181"  
[1] ""  
[1] ""  
[1] "vigHighSES: 0"  
[1] "vigHighIntegrationUK: 1"  
[1] "sensitiveList: 1"  
[1] "Mean (numYes, list exp): 2.89473684210526"  
[1] "Std dev (numYes, list exp): 0.721607015854438"  
[1] ""  
[1] "Mean (report, basic Q): 0.747252747252747"  
[1] "Std dev (report, basic Q): 0.436994955307943"  
[1] ""  
[1] ""  
[1] "vigHighSES: 1"  
[1] "vigHighIntegrationUK: 0"  
[1] "sensitiveList: 0"  
[1] "Mean (numYes, list exp): 1.9438202247191"  
[1] "Std dev (numYes, list exp): 0.435878177356246"  
[1] ""  
[1] "Mean (report, basic Q): 0.734939759036145"  
[1] "Std dev (report, basic Q): 0.444048375451825"  
[1] ""  
[1] ""  
[1] "vigHighSES: 1"  
[1] "vigHighIntegrationUK: 0"  
[1] "sensitiveList: 1"  
[1] "Mean (numYes, list exp): 2.65217391304348"  
[1] "Std dev (numYes, list exp): 0.732822071994357"  
[1] ""  
[1] "Mean (report, basic Q): 0.791208791208791"  
[1] "Std dev (report, basic Q): 0.408696668732407"

MEANS FOR BOTH  
THE LIST EXPERIMENT  
& THE FGM-SENSITIVE  
QUESTION.

results29April2019 (1)

```
[1] " "  
[1] " "  
[1] "vigHighSES: 1"  
[1] "vigHighIntegrationUK: 1"  
[1] "sensitiveList: 0"  
[1] "Mean (numYes, list exp): 2.03488372093023"  
[1] "Std dev (numYes, list exp): 0.284818803849029"  
[1] " "  
[1] "Mean (report, basic Q): 0.698795180722892"  
[1] "Std dev (report, basic Q): 0.461570483454048"  
[1] " "  
[1] " "  
[1] "vigHighSES: 1"  
[1] "vigHighIntegrationUK: 1"  
[1] "sensitiveList: 1"  
[1] "Mean (numYes, list exp): 2.84210526315789"  
[1] "Std dev (numYes, list exp): 0.624257677867289"  
[1] " "  
[1] "Mean (report, basic Q): 0.739130434782609"  
[1] "Std dev (report, basic Q): 0.441515004595909"  
[1] " "  
[1] " "  
>
```