

# **A literature review of methods used to measure attitudes towards disability**

**December 2022**



**Údarás Náisiúnta Míchumais  
National Disability Authority**

## Contents

<b>Introduction.....</b>	<b>4</b>
Background .....	4
<b>Methods of measuring attitudes .....</b>	<b>6</b>
Direct measures .....	7
Self-report surveys.....	7
Q-methodology .....	9
Sociometric techniques .....	11
Semantic differential technique .....	12
Single-item direct attitude scales .....	12
Vignettes .....	13
Threats to validity associated with direct measures .....	16
Indirect measures .....	17
IAT .....	18
Go/No Go Association Task (GNAT) .....	20
Implicit Relational Assessment Procedure (IRAP) .....	20
Vignettes .....	21
List Experiments.....	22
The Endorsement Experiment.....	23
The Randomised Response Method .....	24
Policy Framing.....	24
Threats to validity associated with indirect measures .....	25
<b>Comparison of indirect and direct measures .....</b>	<b>26</b>
Evaluation of the List Method compared to other indirect techniques.....	26
Relationship between indirect and direct measures .....	27
Predictive value of direct and indirect measures.....	28

<b>Modes of survey data collection.....</b>	<b>31</b>
<b>Conclusion .....</b>	<b>34</b>
<b>References .....</b>	<b>36</b>
<b>Appendix I: Additional information on list experiments .....</b>	<b>50</b>

## Introduction

Since 2001, the National Disability Authority (NDA) has conducted a time series Public Attitudes to Disability in Ireland survey, every five or six years.<sup>1</sup> These surveys have been face-to-face, interview administered surveys to a random sample of the Irish population with quotas to match the age and gender demographics of the national population. In 2022, as the next attitudes survey was due, the NDA decided to review whether a public attitudes survey to disability was the optimum approach to take in terms of measuring attitudes.

This review entailed three distinct but inter-related elements:

1. A review of the literature on attitudes and, in particular, changing negative attitudes and behaviours towards disabled people.
2. A review of the various methodologies by which attitudes are measured
3. A list and associated experiments on attitudes to disability (conducted by the Economic and Social Research Institute (ESRI) for the NDA).

This report will cover the second item, i.e. a review of common methodologies used to measure attitudes. Separate reports will present the findings from the other two elements of this project. The literature and methodological reviews and the ESRI work will inform NDA's decisions about its future research and policy work in the area of attitudes.

## Background

Individuals with disabilities are the world's largest minority group and one of the only minority groups that any individual can unexpectedly become a member of at any time.<sup>2</sup> However, people with disabilities are the most under-represented minority group, even within Equality, Diversity and Inclusion Policies.<sup>3</sup> Negative attitudes towards people with disabilities pose substantial barriers to the full social inclusion of disabled people. Research has shown that attitudes play at least some part in predicting various behaviours and indicating values.<sup>4</sup> For example, negative attitudes have been associated with a desire for increased social distance

---

<sup>1</sup> Surveys were conducted in 2001, 2006, 2011 and 2017 and are available at <https://nda.ie/awareness-raising-and-attitudes/public-attitudes-to-disability-in-ireland>

<sup>2</sup> Kallman, 2017; United Nations, 2006

<sup>3</sup> United Nations, 2006 cited in Kallman, 2017; Ball et al., 2005; Walstra et al., 2022; Wolbring and Lillywhite, 2021; Gould et al., 2020

<sup>4</sup> Cooke et al., 2016; Perugini, Richetin, and Zogmaister, 2010

from people with Schizophrenia,<sup>5</sup> developmental disabilities,<sup>6</sup> and intellectual disabilities.<sup>7</sup> Furthermore, a large body of research has indicated that employers often hold negative attitudes and stereotypes regarding people with disabilities which may lead to workplace discrimination.<sup>8</sup> As part of a systematic review, Fitzgerald and Hurst (2017) highlighted that negative bias among healthcare providers was negatively correlated with quality of care indicators for people with mental illnesses<sup>9</sup> and acquired brain injuries<sup>10</sup>. Therefore, research into attitudes is an appropriate approach towards investigating the causes of prejudice and discrimination against people with disabilities in order to identify ways of mitigating them.<sup>11</sup>

---

<sup>5</sup> Loch and Wang, 2014; Economou et al., 2012; Dabby, 2015

<sup>6</sup> Meza et al., 2019

<sup>7</sup> Pelleboer-Gunnick, Van Weeghal and Embregts, 2021

<sup>8</sup> Rohmer and Louvet, 2018

<sup>9</sup> MacKay and Barrowclough, 2005; Neauport et al., 2012, Peris, Teachman and Nosek, 2008 cited in Fitzgearld & Hurtz, 2017

<sup>10</sup> Redpath et al., 2010 cited in Fitzgerald and Hurtz, 2017

<sup>11</sup> Aaberg, 2012

## Methods of measuring attitudes

Methods of measuring attitudes have changed over the years, becoming more sophisticated as knowledge of the underlying processes involved in attitude formation and persistence has improved.<sup>12</sup> As a result, there now exists an abundance of measurement tools and approaches to measuring attitudes. The purpose of this chapter is to provide a brief overview of some of the measures that are commonly used to assess attitudes towards disability.

Although research into attitudes is continually advancing, it is currently generally accepted that attitudes consist of three components: affective, cognitive and behavioural.<sup>13</sup> The affective component reflects the emotions underpinning an attitude, such as and the amount of positive or negative feelings towards the referent.<sup>14</sup> The cognitive component refers to the individual's ideas, thoughts, perceptions, beliefs, opinions, or mental conceptualisation of the attitude referent. The behavioural component reflects the individual's intent to behave in a certain way, or their actual behaviour, towards the referent.<sup>15</sup> Studies have found that these affects, cognitions and behaviours are not necessarily congruent within individuals.<sup>16</sup> For example, Findler et al. (2007) found that participants responded more positively on the behavioural aspect of their multidimensional scale of attitudes towards people with disabilities compared to the affective and cognitive components. The authors argued that this indicated that individuals did not intend to act on their feelings or thoughts of discomfort towards people with disabilities, or they at least did not want to admit to acting on them. Findings such as these emphasise the importance of assessing all three components when measuring attitudes. Not doing so risks gathering an incomplete understanding of an individual's attitudes and behaviours.

There is also a general consensus that attitudes can be categorised into two distinct but related levels.<sup>17</sup> Those that exist on an explicit or conscious level and those that exist on an implicit or less conscious level.<sup>18</sup> Explicit attitudes are attitudes that the individual is aware of, they are deliberately formed and are easy

---

<sup>12</sup> Antonak & Liveneh, 2000

<sup>13</sup> Olson & Zanna, 1993

<sup>14</sup> Antonak and Liveneh, 1988

<sup>15</sup> Cook, 1992

<sup>16</sup> Breen, 2018

<sup>17</sup> Greenwald et al., 2009

<sup>18</sup> ibid

to self-report. Implicit attitudes on the other hand, are formed involuntarily and the individual is generally unaware of them. Explicit attitudes are generally measured by what are known as direct measures while implicit attitudes are measured by indirect measures.

### **Direct measures**

Direct measures are the most widely used method of measuring attitudes to persons with disabilities, likely because they are easier to administer since explicit attitudes can be self-reported or overtly observed.<sup>19</sup> When using direct methods of measuring attitudes, respondents are aware that their attitudes are being measured, either because they have been informed by the researcher or it becomes apparent through the process of being assessed.<sup>20</sup> There are numerous direct measures available to assess attitudes. Below is a brief overview of some of the common direct measures that have been identified.<sup>21</sup>

### **Self-report surveys**

The most common method of directly measuring attitudes is through self-report surveys in which participants are asked to express their attitudes by responding to a list of questions or statements about people with disabilities. Many surveys use validated scales to measure attitudes such as the Attitudes Toward Intellectual Disability Questionnaire,<sup>22</sup> the Revised Acceptance of Disability Scale<sup>23</sup> or the Attitude to Disability Scale.<sup>24</sup> A recent systematic review assessed the psychometric properties and overall utility of 31 identified validated scales for assessing attitudes towards disability.<sup>25</sup> The reviewed instruments included five that measured attitudes toward communication disability (used in this instance to refer to people with speech or hearing impairments), seven toward intellectual disability, four toward mental illness, and 15 toward disability in general. Target respondents included children and adults from different occupations and cultural backgrounds. Twenty-three measures were found to have adequate overall utility, while eight were found to have poor overall utility. The authors concluded that there are a number of adequate validated surveys in existence for researchers to

---

<sup>19</sup> Kallman, 2017

<sup>20</sup> Antonak and Liveneh, 2000

<sup>21</sup> There are many approaches so this is not an exhaustive list, just an introduction to some common methods.

<sup>22</sup> Morin et al., 2013

<sup>23</sup> Grooms et al., 2007

<sup>24</sup> Power et al., 2014

<sup>25</sup> Palad et al., 2016

use, however, they caution that careful consideration must be given to the limitations of the chosen scale when they are used. A more recent systematic review of self-report measures assessing workplace attitudes towards people with disabilities highlighted a number of validity issues with the nine scales identified and particularly highlighted the lack of explicit theoretical foundations underpinning the scales.<sup>26</sup>

Self-report measures have been widely used to directly assess attitudes towards people with disabilities, including those with mental illness,<sup>27</sup> <sup>28</sup> disabilities in general,<sup>29</sup> those with intellectual disabilities,<sup>30</sup> autism,<sup>31</sup> and sensory disabilities<sup>32</sup> Many studies, particularly those conducted in academic settings, have relied on convenience sampling, and recruited university students, so generalisability limitations must be considered. A small number of survey studies which did achieve representative random samples have also been identified, which provide good levels of evidence regarding public attitudes towards disability.<sup>33</sup> Many of these nationally representative surveys have relied on face-to-face interviewing, however, which can threaten the validity of the data, due to factors such as socially desirable responding. This will be discussed in more detail later.

Many self-report surveys rely on Likert scale responses in which respondents indicate their level of agreement usually on a scale of five or seven, for example, 1 (strongly disagree) to 5 (strongly agree). Likert scales are advantageous as they are widely used, easily understood, and easily quantifiable. They are also quick and inexpensive to administer. Likert scales also make it easier for the respondent to answer the question as they are not forced to give an absolute answer but can indicate a degree of agreement. Debates exist around the use of Likert scales to measure attitudes, however. For example, some argue that attitudes are multi-dimensional and exist on a vast continuum, whereas Likert scale are unidimensional and only give a small number of response options. Research also indicates that people often avoid extremes when answering, and may tend to give answers that are closer to the middle of the scale. Given the

---

<sup>26</sup> Breen, 2018

<sup>27</sup> Dabby et al., 2015

<sup>28</sup> Arora et al., 2019

<sup>29</sup> Bollier et al., 2021

<sup>30</sup> Murch et al., 2018

<sup>31</sup> Campbell et al., 2008

<sup>32</sup> Cooper et al., 2014; McDonnell et al., 2019

<sup>33</sup> NDA, 2017, 2011, 2006; British Attitudes Survey; the Eurobarometer; Bollier et al., 2021

complexity around attitudes and behaviours, some researchers argue that relying on Likert scales alone is likely an inefficient approach to assessing internal processes. Numerous alternatives to Likert-type scales exist (e.g. q-methodology, socio-metrics) that may increase the validity of attitudes data.<sup>34</sup>

### **Q-methodology**

Q-methodology is a combination of qualitative and quantitative techniques for studying subjective views on a topic.<sup>35</sup> In a Q-methodology task, participants' views about a particular topic are gathered by asking them to rank a set of statements (Q-set) about the topic according to a specified criterion, such as intensity of agreement.<sup>36</sup> Developing the Q set involves collecting all logical ideas, concepts, or statements that can be expressed about the topic until saturation is reached, this is called development of the concourse (i.e. the body of information related to the topic of information).<sup>37</sup> Developing the concourse can involve conducting a literature review, preliminary data collection (e.g. interviews or surveys), and/or searching publicly available resources or academic databases. Collected statements are then reduced and refined by the researchers (e.g. grouping similar ideas together, consulting with experts) to produce a manageable Q-set of, usually, up to 60 statements that are presented to the participant.<sup>38</sup> Each respondent's final ranking is called their 'Q-sort' and this is then analysed using factor analysis techniques to identify similarities and differences in the ways in which participants have ranked the items. Qualitative interpretation of the commonalities across participants generates the factors.<sup>39</sup> Participants are purposively sampled as they would be in qualitative research. The aim of Q-methodology is not to achieve a generalisable sample, instead the aim is to include a comprehensive range of perspectives on the topic under discussion.<sup>40</sup> A sample size of 40-60 participants is typical in Q-methodology studies.<sup>41</sup> A number of advantages of using Q-methodology have been identified. Q-methodology has been found to be useful for exploring complex topics, or topics

---

<sup>34</sup> Antonak and Liveneh, 2000; Haddock and Zanna, 1998

<sup>35</sup> Brown et al., 1997 cited in Chuyrruca et al., 2021

<sup>36</sup> Watts et al., 2012 cited in Chuyrruca et al., 2021

<sup>37</sup> McKenzie et al., 2011

<sup>38</sup> Paige and Morin, 2014; Chuyrruca et al., 2021

<sup>39</sup> Chuyrruca et al., 2021

<sup>40</sup> Stainton et al., 1995 cited in McKenzie et al., 2011

<sup>41</sup> Watts et al., 2012, Stainton et al., 1995 cited in Chuyrruca et al., 2021

that elicit strong, differing or under-acknowledged views.<sup>42</sup> Some researchers have argued that the Q-methodology represents a holistic approach and can obtain a fuller perspective due to the forced choice nature of the approach.<sup>43</sup> Other advantages include the systematic process involved, which researchers argue ensures the internal validity of the method, the replicability of the method,<sup>44</sup> the integration of qualitative and quantitative techniques, the ability to administer the method in a variety of modes including online or face-to-face;<sup>45</sup> the requirement for only a small number of participants, and the increased accessibility of the method to participants with certain communication difficulties due to the non-verbal nature in which views are gathered.<sup>46</sup> Q methodology has faced criticism related to a perceived lack of robust data collection methods and rigorous analysis.<sup>47</sup> Even proponents of this method concede that demonstrated evidence of the robustness of certain elements, such as concourse sampling, has not been adequately reported in the literature and that the accessibility of the method to researchers could be improved.<sup>48</sup> Additionally, the reliability of Q-methodology has also been questioned because the same result is not necessarily obtained when an individual repeats the task.<sup>49</sup> A risk of bias exists at the data collection phase, as the participant may provide socially desirable answers, and at the data interpretation phase as this task lies with the researcher.<sup>50</sup>

Q-methodology has often been used to directly assess attitudes towards disability. The method has been used to assess, for example, attitudes towards the sexuality of people with ID<sup>51</sup>, understandings of Downs Syndrome<sup>52</sup>, private therapists' attitudes towards working with people with learning disabilities<sup>53</sup>,

---

<sup>42</sup> Farrimond, 2017; Gidman et al., 2009; Grimshaw et al., 2017; Cross-Sudworth et al., 2011 cited in Chuyrruca et al., 2021; Bryant et al., 2006

<sup>43</sup> Ho et al., 2015; Daniels et al., 2013 cited in Churruca et al., 2021

<sup>44</sup> Cross, 2005

<sup>45</sup> Brown and Pritle, 2008; Churruca et al., 2021

<sup>46</sup> Churruca et al., 2021

<sup>47</sup> Kenward, 2019

<sup>48</sup> ibid

<sup>49</sup> Cross, 2005

<sup>50</sup> ibid

<sup>51</sup> Brown and Pritle, 2008

<sup>52</sup> Bryant et al., 2006

<sup>53</sup> Besika et al., 2018

practitioners' beliefs relating to parenting and psychosis,<sup>54</sup> and stakeholder beliefs about the use of coercion in mental healthcare<sup>55</sup>

### **Sociometric techniques**

Sociometric techniques are designed to measure social relationships and uncover how an individual behaves or intends to behave towards a person within the referent group (e.g. a person with a disability).<sup>56</sup> For example, a child might be asked to categorise their classmates in terms of who they most like to play with or sit with. The resulting data are then used to provide a picture of the social status and levels of acceptance of the individuals in the group. The method has been commonly used in education research in particular, with many studies focusing on inclusive education and disabled children.<sup>57</sup> There are a number of limitations to using sociometric techniques, however. For example, asking individuals to simply name people relies on free recall, which may mean that some individuals in the group are forgotten or neglected.<sup>58</sup> This approach may also require the participants to write out the names which can lead to frustration or fatigue. To address this, some researchers provide a roster of all of the names of individuals in the group, but this may produce a response bias as some may have a tendency to just pick the names at the top of the list.<sup>59</sup> There are also ethical concerns around using sociometric techniques. For example, asking children to rate their peers could have a negative impact upon those who already have difficulty with peer relationships and reinforce their pre-existing isolated or rejected position.<sup>60</sup> A further limitation relates to the assumption which underlies sociometric techniques, that all children need or want a certain amount of friends.<sup>61</sup> Child and Nind (2012) have cautioned that:

this practice of categorising individuals as popular, and so on, may construct and reinforce social difficulties, especially for disabled children. It fails to consider the role of the environment in creating these difficulties, and instead may blame the individual. (p. 1)

---

<sup>54</sup> Adderly et al., 2019

<sup>55</sup> Disetch et al., 2011

<sup>56</sup> Antonak and Liveneh, 2000

<sup>57</sup> Avramidis, 2013; Kuhne & Wiener, 2000; Nepi et al., 2013;

<sup>58</sup> Avramidis et al., 2017

<sup>59</sup> *ibid*

<sup>60</sup> Child and Nind, 2012

<sup>61</sup> *ibid*

### **Semantic differential technique**

The Semantic differential method, developed by Osgood, Suci, and Tannenbaum (1957), measures a person's attitudes and beliefs about a single concept, phenomenon, or object on a bipolar scale defined by contrasting adjective pairs, e.g. "likable-unlikable". The adjectives tend to reflect three dimensions: evaluation (i.e. whether a respondent views the attitude theme in a negative or positive light), potency (i.e. how strongly a respondent feels about the attitude topic) and activity (i.e. whether the attitude topic is viewed as active or passive).<sup>62</sup> Participants are asked to mark a point on the scale that represents their perspective.<sup>63</sup> Then, the factors underlying the multivariate data are extracted by conducting a factor analysis to explain the affective responses in terms of a limited number of specific main factors. The advantages of using Semantic Differential techniques include being comprehensive, simple to administer and relatively straightforward to complete.<sup>64</sup> Similar to other scales, disadvantages include the potential for positional response bias, whereby participants choose one response position more often than any others regardless of the item content.<sup>65</sup>

This technique has been used in a variety of contexts in relation to measuring attitudes towards disabled people. For example, Burke et al. (2019) used the technique as a direct measure of undergraduate students' attitudes towards people who have engaged in non-suicidal self-harm. Tackahashi (2018) used the technique to assess undergraduate students' attitudes to people with various disabilities such as intellectual disability, physical disability and deafness. Janero et al. (2018) used the technique to assess attitudes towards people with disabilities in general.

### **Single-item direct attitude scales**

Single item direct measures of attitudes are used to measure a targeted construct using a single targeted question.<sup>66</sup> For example, Crittenden (2018) asked participants to indicate their attitude towards disabled people on a scale of 1 (extremely unfavourable) to 100 (extremely favourable), using a visual of a

---

<sup>62</sup> Osgood et al., cited in Janero et al., 2017

<sup>63</sup> Antonak and Liveneh 2000

<sup>64</sup> Chin, Johnson, & Schwartz, 2008 cited in Jenaro et al., 2018; Russell 2011 cited in Brown et al, 2020).

<sup>65</sup> Russell 2011 cited in Brown et al, 2020

<sup>66</sup> Brown et al, 2020

thermometer as an aid. The project Implicit website,<sup>67</sup> has used also a single item to directly measure explicit attitudes. Participants were asked to indicate their preference for disabled or non-disabled people on a seven point rating scale ranging from strongly prefer disabled people to strongly prefer non-disabled people.<sup>68</sup> This was then used to compare explicit attitudes to implicit attitudes gathered by the Implicit Association Test.<sup>69</sup>

There are a number of reasons that researchers want to use single item measures, including reducing respondent burden, which has been linked to increased response rates,<sup>70</sup> reducing the time needed to complete the questionnaire, and reduced cost. Single-item ratings have been found to adequately represent the outcomes of traditional multi-item assessments. In one study, the majority of single-item assessments agreed with their lengthier multiple-item subscales and adequately represented the outcomes of traditional multi-item assessments.<sup>71</sup> Single-item scales are sometimes used to gauge constructs that, originally, were assumed to comprise several dimensions. However, while there may be benefits in relation to time, cost and respondent burden, single-item attitude scales can be problematic.<sup>72</sup> As discussed earlier, it is generally accepted that attitudes are multi-dimensional and so, many researchers argue that multi-dimensional measurement is required to obtain accurate data.<sup>73</sup> Another criticism of single item measure is the potential for acquiescence bias because participants have to decide on the spot if they agree or disagree with the statement which may lead to an overestimation in the ratings.<sup>74</sup>

### **Vignettes**

Vignettes simulate real events through short descriptive sketches about hypothetical characters in hypothetical situations which are presented to respondents to elicit their reactions and opinions according to how they state they would behave in the situation depicted.<sup>75</sup> Vignettes are a popular method used in social psychology to investigate attitudes and prejudices towards certain

---

<sup>67</sup> <https://implicit.harvard.edu/implicit/>

<sup>68</sup> Xu et al., 2014; Friedman et al., 2019

<sup>69</sup> e.g. Friedman et al, 2019; Charlesworth and Banji, 2019; Feldner et al., 2022

<sup>70</sup> Rolstad et al., 2011

<sup>71</sup> Verster et al., 2021

<sup>72</sup> Brown et al., 2020

<sup>73</sup> Antonak and Liveneh, 2000

<sup>74</sup> Fabrigar et al. 2005 cited in Brown et al., 2020

<sup>75</sup> Hughes, 2002; Lewis and Stenfert-Kroese, 2010; Shoenberg and Ravdal, 2000

groups, and can be used as a direct or indirect measure of attitudes. They are useful when observing or placing an individual in a particular context is not possible for logistical or ethical reasons.<sup>76</sup> They can be depicted in text, images or other stimuli, such as videos and may consist of single snapshots of scenarios or a series of stages as the scenarios unfold.<sup>77</sup> Vignette design generally consists of two components: an initial description of a hypothetical situation and the participant's responses to the situation.<sup>78</sup> Following the vignette, respondents may be asked either open-ended or closed questions with a set of response categories from which they may choose.<sup>79</sup>

The practical advantages of using vignettes were outlined in a review of the literature by Hughes and Hubey (2002):

- Vignettes are more efficient in terms of cost and time compared to observational studies.<sup>80</sup>
- Vignettes can lead to more uniform data as the content and staging of the vignette content can be defined and standardised.
- Vignettes do not necessarily require participants to have in-depth knowledge of the topics under study<sup>81</sup> and may be used to elicit participants' automatically generated meanings<sup>82</sup>
- Vignettes can help to desensitise sensitive topics (such as attitudes to disability) by asking participants to respond from the point of view of the vignette character rather than as themselves. This allows respondents to distance themselves from the topic to make responding easier and remove feelings of embarrassment.<sup>83</sup> This can also help reduce socially desirable responding. The situated context of vignettes in stories allows participants to respond within the context of the particular situations presented.<sup>84</sup>

---

<sup>76</sup> Fleming and Stalker, cited in Beneditti et al., 2018

<sup>77</sup> Benedetto et al., 2018

<sup>78</sup> Andersson et al., 2015

<sup>79</sup> Beneditti et al., 2018

<sup>80</sup> Gould 1996, Sumrall & West 1998, Wilson & While 1998 cited in Hughes and Hubey 2002

<sup>81</sup> Liker, 1982 cited in Hughes and Hubey, 2002

<sup>82</sup> Bendelow, 1993; Collins & Brief, 1995 cited in Hughes and Hubey, 2002

<sup>83</sup> Finch, 1987 cited in Schoenberg & Ravdal, 2000

<sup>84</sup> Hughes, 1998 cited in Hughes and Hubey, 2002

- Vignette scenarios may be constructed from unrealistic events in order to understand more about how individuals might behave. However, they can also be based on real life situations to reveal nuances that only 'insiders' are usually aware of.<sup>85</sup>
- They can provide a useful focus for discussion during individual interviews<sup>86</sup> and can act as a stimulus for group discussions.<sup>87</sup>
- From a practical point of view, where research studies are limited by available resources, vignettes can be used to generate a sound data-base relatively quickly and cost effectively.<sup>88</sup>

Additional advantages according to Schoenberg and Ravdal (2000) include:

- the storytelling nature of vignettes, which may be relaxing, pleasant, and interesting and could reduce feelings of being overburdened by the interview process<sup>89</sup>
- their adaptability to suit the topic or population of interest.

In their study, Shoenberg and Ravdal also highlighted a number of problems with vignettes in relation to data collection and analysis. For example, some participants did not want to give advice based on the vignette scenario. Shoenberg and Ravdal suggest that this may happen if respondents have a lack of faith in their own opinion or have a perceived lack of knowledge related to the subject. They suggest that this may be overcome by creative probing, encouragement and reassurance from the interviewer. A second issue was the potential threats to validity. Not rigorously pre-testing the vignettes to ensure content validity and clarity may leave them open to interpretation by both the participants and the researchers. In addition to pre-testing, Shoenberg and Ravdal recommend using standard probes to ensure participants stay focused on the topic, and ongoing debate, comparison and contextualisation of the data by the researchers. Additionally, since vignette scenarios are hypothetical, they may not predict the respondents' actual behaviour.<sup>90</sup> problems may also arise if there is a

---

<sup>85</sup> Sumrall and West 1998; Rahman 1996 cited in Hughes and Hubey, 2002

<sup>86</sup> Hughes 1998 cited in Hughes and Hubey, 2002

<sup>87</sup> Sim et al. 1998 cited in Hughes and Hubey, 2002

<sup>88</sup> Hughes and Hubey, 2002;

<sup>89</sup> Schoenberg and Ravdal, 2000 cited in Benedetti et al., 2018; Kayser-Jones & Koenig, 1994 cited in Schoenberg & Ravdal, 2000

<sup>90</sup> Hughes and Hubey, 2002; Shoenberg and Ravdal, 2000

large disconnection between the actual experiences of the participants and the vignette characters they are asked to assume.<sup>91</sup>

In recent research, vignettes have been used to explore the attitudes of various groups towards people with various types of disabilities. Sanhala & Theodore (2015) used a vignette as part of an online survey to investigate public attitudes towards the sexuality of people with intellectual disabilities within a UK sample and to compare attitudes between people from White Western and South Asian backgrounds. Anderson et al. (2015) also used vignettes as part of an online survey to investigate Swedish employers' experiences and attitudes toward hiring people with various disabilities. In a review of health professionals' stigma towards people with mental illness, Carrera et al., found that four of the reviewed studies used vignettes to assess attitudes or stigma towards mental illness. Dabby et al., (2015) used vignettes as part of the Social Distance Scale to assess attitudes towards people with Schizophrenia and diabetes mellitus and compared the responses to the vignettes to responses on another direct measure as well as an indirect measure (the IAT).

### **Threats to validity associated with direct measures**

Research on attitudes towards people with disabilities tends to indicate increasingly positive attitudes towards this group.<sup>92</sup> However, such generally positive attitudes are contrary to the well-documented experiences of discrimination and stigma that are commonly reported by people with disabilities.

One potential reason for this mismatch in reported public attitudes and reported experiences of people with disabilities is that attitudes towards people with disabilities have most commonly been examined through direct measures, such as those listed above. However, direct measures of attitudes are subject to a number of validity issues.<sup>93</sup> Simply responding to an attitude measure may create an attitude that did not exist prior to responding, or could create a transient attitude that the researcher may mistake as stable.<sup>94</sup> Respondents may intentionally distort their answers for a number of reasons including, wanting to please the researcher by telling them what they think they want to hear, giving the referent groups the benefit of the doubt, trying to sabotage the study by deliberately providing inaccurate representations of attitudes, or failing to give thorough answers due to a lack of interest<sup>95</sup>. Another threat to the validity of

---

<sup>91</sup> Hughes and Hubey, 2002

<sup>92</sup> e.g. NDA, 2017

<sup>93</sup> Antonak and Levneh, 2000

<sup>94</sup> *ibid*

<sup>95</sup> *ibid*

direct measures is the social desirability bias. There is ample evidence that survey questions asking about sensitive topics, such as attitudes towards people with disabilities, often generate inaccurate data because they are distorted by this bias.<sup>96</sup> Respondents often misrepresent their true attitudes and behaviours if, for example, they feel they conflict with prevalent social norms, or if they want to give a good impression of themselves. So, due to self-preservation, respondents often underreport socially undesirable behaviours and attitudes and overrepresent socially desirable ones.<sup>97</sup> There is also the possibility that incorrect data could be generated because some participants could lack the necessary introspective ability when answering self-report surveys, especially if there is a disconnect between who the person is and who they think they are.<sup>98</sup>

Styles of responding on direct measures can also threaten validity. For example, some respondents may answer all items affirmatively, also known as the acquiescence style. Some respondents may choose responses based on where in the continuum of the scale they are placed e.g. choosing the middle or neutral response for each item. Another response style is when respondents give similar responses to items that they perceive to be related. Finally, some respondents may only give responses that are unusual or uncommon.<sup>99</sup> Considering these threats to validity, some argue that it is unlikely that measuring explicit attitudes alone provides a sufficient basis for a holistic understanding of attitudes towards people with disabilities.<sup>100</sup>

### **Indirect measures**

In recent decades, more attention has been paid to the measurement of implicit or unconscious attitudes using indirect measures. Some argue that they represent a more accurate basis for the measurement and description of values and behaviours because, due to the time pressure often employed, participants have less opportunity to manipulate their answers as compared to when responding to direct measures.<sup>101</sup> Indirect methods of measuring attitudes include nonobtrusive behavioural observations in which the respondents do not know that they are being observed or measured; projective techniques in which respondents are aware that they are being observed or measured, but are not necessarily aware

---

<sup>96</sup> Krumpal, 2013; Gnamb and Kasper, 2015; Rohmer and Louvet, 2018 cited in Friedman, 2019

<sup>97</sup> Krumpal, 2013

<sup>98</sup> Hoskin, 2012

<sup>99</sup> Antonak and Livneh, 2000

<sup>100</sup> e.g. Antonak and Livneh, 2000

<sup>101</sup> Aaberg, 2012; Amodia and Mendoza cited in Robb, 2015; Gawronski and De Houwer, 2014 cited in Cabrera et al.,2014

of, or clear about, the purpose of the measurement; disguised techniques in which respondents are deliberately deceived regarding the true purpose of the measurement situation; and physiological methods in which the respondents are aware they are being measured but are inactive participants in the measurement process.<sup>102</sup> The assumption of physiological indirect measures is that physiological reactions, such as heart rate, blood pressure, pupil dilation or electrical conduciveness of the skin, are linked to autonomic arousal or the intensity of an underlying attitude.<sup>103</sup> Other indirect measures do exist in which participants are aware that they are being measured or observed and are aware of the purpose of the measurement situation but that responding is so automatic that this awareness is assumed not to pose a threat to validity. The Implicit Association Test, the Go/No Go Association Test (GNAT) and the Implicit Relational Assessment Procedure (IRAP) are examples of this. Although such techniques may be useful for avoiding certain threats to validity such as social desirability biases, they have also been criticised for a lack of reliability and psychometric validation.<sup>104</sup>

## **IAT**

The Implicit Association Test (IAT), originally developed by Greenwald, McGhee, and Schwartz (1998), is one of the most widely used indirect measures for assessing attitudes. The instrument is computer-based and determines the strength of participants' unconscious biases by the speed at which participants classify different groups of words or images into subordinate categories.<sup>105</sup> Two of these groups are target concepts (e.g. disabled and non-disabled) and two are attribute concepts (e.g. pleasant and unpleasant).<sup>106</sup> The assumption is that participants will categorise stimuli faster when the categories align with the participant's internally held associations or attitudes. The IAT has been used to assess a whole host of implicit cognitions including racial attitudes,<sup>107</sup> gender attitudes<sup>108</sup> and attitudes towards criminal violence.<sup>109</sup>

---

<sup>102</sup> Livene & Antonak, 1995 cited in Antonak and Liveneh, 2000

<sup>103</sup> Cacioppo et al., 1990 cited in Antonak and Liveneh, 2000;

<sup>104</sup> Yaker, 1994 cited in Preutt, 2006

<sup>105</sup> Teige-Mocigemba, Klauer, & Sherman, 2010 cited in Rob, 2015

<sup>106</sup> McCaughey, 2005

<sup>107</sup> Ahadinezhad et al., 2022

<sup>108</sup> Parker et al., 2018

<sup>109</sup> Simane-Vigante, et al., 2015

A large body of research has now used the IAT to examine attitudes towards people with disabilities. Some researchers have developed and used their own disability versions of the IAT<sup>110</sup>, however, the validated Disability Attitude Implicit Association Test (DA-IAT)<sup>111</sup> is the most widely used disability-related IAT.<sup>112</sup> The DA-IAT is similar to the original IAT except the target-concept discriminations are ‘disabled-persons’ and ‘abled-persons’<sup>113</sup> and the attribute dimensions are ‘good’ and ‘bad’.<sup>114</sup>

Research has provided support for the reliability and validity of the IAT and the DA-IAT.<sup>115</sup> However, some challenges have also been identified, such as the ambiguity in interpretation caused by comparing two categories. By design, the IAT can only be used as a relative comparison of a person’s implicit biases towards two different groups (e.g. disabled and non-disabled people) and not their absolute implicit attitude towards one of the groups alone. In response to such limitations, other versions of the IAT have been developed. These include the Single Category Implicit Association Test (SC-IAT);<sup>116</sup> and the Single Target Implicit Association Test (ST-IAT)<sup>117</sup> which measure absolute implicit attitudes to a single target (i.e. disabled person) and concepts (e.g. good/bad).

Many studies have used secondary data gathered from Project Implicit, a website where anyone can test their implicit attitudes and biases to a variety of groups, including people with disabilities.<sup>118</sup> Although using such data allow for analyses on large samples, caution must be exercised as all participants are self-selected and so a selection bias may exist. Nevertheless, regardless of the type of IAT used,

---

<sup>110</sup> e.g. Chen et al., 2011; Enea-Drapeau et al. 2012; Hein et al., 2011 cited in Wilson and Scior 2014

<sup>111</sup> Pruett & Chen, 2006

<sup>112</sup> Friedman et al., 2015; Wilson and Scior, 2014

<sup>113</sup> These are the terms used in the official tool

<sup>114</sup> Wilson & Scior, 2014

<sup>115</sup> Greenwald and Nosek, 2001; Aaberg, 2012; Thomas et al., 2013; White et al., 2006 cited in Friedman and VanPuymbrouck, 2021b

<sup>116</sup> Karpinski & Steinman, 2006

<sup>117</sup> Wigboldus, Holland, & van Knippenberg, 2004

<sup>118</sup> e.g. Friedman & van Puymbrouck, 2019; Friedman et al., 2019; Charlesworth and Banaji, 2019; Lum et al., 2021

overall, findings indicate negative implicit attitudes toward the disabled population.<sup>119</sup>

### **Go/No Go Association Task (GNAT)**

The GNAT<sup>120</sup> is similar to other indirect measures such as the IAT but examines only one category. Respondents categorise stimuli (e.g. disabled person) depending on whether or not they are associated with an evaluative category (e.g. positive or negative) and respond by pressing a key (“go”) when stimuli belong to the category or to give no response (“no go”) when stimuli do not belong to the category.<sup>121</sup> Like other indirect measures, the strength of associations in the GNAT is determined by the speed of responding.

The GNAT has been used less commonly than the IAT in disability related research but has been used to assess attitudes towards mental illness<sup>122</sup> and intellectual and developmental disabilities.<sup>123</sup> James et al. (2018) found in their study, however, that the GNAT had unacceptable internal consistency and suggested that this tool may not be as good a choice for assessing implicit attitudes to mental illness compared to the IAT.

### **Implicit Relational Assessment Procedure (IRAP)**

Relational Frame Theory (RFT) is a psychological theory of behaviour, language and cognition.<sup>124</sup> RFT focuses more on how stimuli are related rather than the extent to which associations exist between them which is what the IAT is based on.<sup>125</sup> Proponents of using RFT in the measurement of implicit cognition argue that this allows for greater specificity and context sensitivity compared to associative approaches such as the IAT or GNAT. The Implicit Relational Assessment Procedure (IRAP),<sup>126</sup> an indirect measure of implicit cognitions, based on RFT, is a computerised assessment tool that requires participants to respond in ways that are either consistent or inconsistent with their prior learning histories by choosing between a pair of relational response options (e.g. “similar”

---

<sup>119</sup> Wilson & Scior, 2006; Robb et al., 2016; Fitzgearld & Hurst, 2017

<sup>120</sup> Nosek and Banji, 2001

<sup>121</sup> James, 2018

<sup>122</sup> James et al., 2018; Kopera et al., 2015; Young et al., 2019

<sup>123</sup> Keith et al., 2015

<sup>124</sup> Hughes et al., 2011; Hughes, Barnes-Holmes and Vahey, 2012 cited in Golijani-Maghaddam et al., 2013

<sup>125</sup> Golijani-Maghaddam et al., 2013

<sup>126</sup> Barnes-Holmes, et al., 2006

and “opposite”) when presented with a label stimulus (e.g. “Pleasant”) and a target stimulus (e.g. “disabled person”). Similar to other measures of implicit bias, the IRAP relies on response-latency to determine the relations between the two targeted concepts.<sup>127</sup> In other words, it is assumed that respondents will be slower to respond when the required response goes against their internally held cognitions.<sup>128</sup>

The IRAP has been found to demonstrate levels of predictive validity comparable to the IAT<sup>129</sup> as well as a resilience to faking strategies.<sup>130</sup> Examples of use of the IRAP in disability-related research included assessing attitudes toward people with visible and invisible intellectual disabilities,<sup>131</sup> and people on the autism spectrum.<sup>132</sup>

### **Vignettes**

As noted previously, Vignettes are often used as a direct measure of attitudes. However, they can also be used as an indirect measures. A systematic review by Fitzgeraald and Hurst (2017) examining implicit bias among healthcare professionals included studies that employed clinical vignettes to assess bias. In these studies, participants were given a number of vignettes in which one characteristic (e.g. ethnicity) varied by vignette. The differences in responding to the different vignettes were measured. If a statistically significant difference in the diagnosis or treatment prescribed was correlated with certain characteristics such as ethnicity or gender of the patient, the researchers inferred that it was partly a result of implicit processes in the physicians’ decision-making. Fitzgerald and Hurst labelled this ‘the assumption method’ because it is based on the assumption that most individuals will be motivated to explicitly disregard factors such as ethnicity. It is not possible, however, to say with certainty whether implicit or explicit processes are involved. Of the 14 studies included that used the assumption method, one study assessed attitudes towards people living with AIDS, two studies assessed attitudes towards people with acquired brain injury, and four assessed attitudes towards people with mental illness.

---

<sup>127</sup> Hughes and Barnes Homes, 2011

<sup>128</sup> Lund, 2019; Hughes and Barnes Homes, 2011

<sup>129</sup> Barnes-Holmes et al., 2010; Barnes-Holmes, et al., 2009

<sup>130</sup> McKenna et al., 2007 cited in Hughes and Barnes-Holmes, 2011

<sup>131</sup> Lund, 2019

<sup>132</sup> Kelly et al., 2013

Another use of vignettes as an indirect measure of attitudes is as part of the Multi-Dimensional Attitudes Scale (MAS).<sup>133</sup> This indirect measure uses a social scenario vignette describing an interaction between non-disabled and disabled people. Respondents read the vignette and are then asked to indicate how they think the non-disabled person would feel (affect), think (cognitions), or act (behaviour) in the situation. Responses are marked on a 5-point Likert-type scale ranging from 1 (not at all) to 5 (very much). The original MAS depicted a situation involving a wheelchair user, but other studies have adapted the MAS to depict people with other types of disabilities including intellectual disability,<sup>134</sup> autism,<sup>135</sup> hearing and vision impairments,<sup>136</sup> and ADHD.<sup>137</sup>

### **List Experiments**

A List Experiment is an experimental design in which some subjects are induced to speak their mind freely, without fear of being blamed for their "socially deviant" opinions, while other subjects are exposed to social normative influences. Between-group comparisons and recent refinements in the list experiment method allows one to estimate the aggregate proportion of "liars" as well as conducting a multivariate analysis of their profile – social desirability responding moderators.<sup>138</sup> The simple idea behind the list experiment is that respondents are not directly asked about the sensitive issue. Instead, they are asked to provide enumerators with an aggregate count of the items on a list that they agree with.

The list experiment randomises a sample of respondents into two groups. A list of control items are presented to a control group and a list of the same control items plus one sensitive item are presented to the treatment group. Respondents are not asked which of the items they agree with, just how many items they agree with. As the groups are randomly assigned and both samples are presented the same control list items, any difference between the average response in the control and treatment groups is due to the sensitive item. Respondents in the control group are asked about the sensitive topic with a direct question. The difference between responses from the direct survey question and responses from the list is interpretable as a measure of the extent to which, for example, support for a particular group is over-stated or negative attitudes are hidden or

---

<sup>133</sup> Findler et al., 2007

<sup>134</sup> e.g. Domagała-Zys'k, 2021; Morin, et al., 2013

<sup>135</sup> Mathews et al., 2015

<sup>136</sup> De Lat et al., 2013

<sup>137</sup> Barr et al., 2015

<sup>138</sup> Marquis, 2017

'masked'. Respondents have the options of "Refuse to Answer" and "Don't Know." The protection of privacy in the list experiment is designed to increase willingness of respondents to provide truthful answers to sensitive questions. The format of list experiments can be confusing, however; as indicated by studies which have found inconsistent findings in terms of direct and indirect measures, especially among some groups, such as those who are less well educated.<sup>139</sup> Thus, while the list experiment can reduce social bias it can introduce different forms of error associated with question complexity and difficulty. Efforts are being made to address these issues including the use of images and clearer instructions for enumerators.<sup>140</sup>

The list experiment design has been used in surveys, for example, to elicit vote preferences,<sup>141</sup> views on undocumented migration,<sup>142</sup> prevalence of the use of micro-finance loans,<sup>143</sup> and opinions on topics such as gay marriage<sup>144</sup> and racism<sup>145</sup> and many other issues. No list experiment focused on measuring attitudes towards people with disabilities has been identified. In an effort to address this gap in the evidence base, the National Disability Authority (NDA) is working through a research partnership with the Economic and Social Research Institute (ESR) to undertake research on attitudes to disability in Ireland using a range of indirect measurement methods, including a list experiment. The details and findings of the list experiment are published in full elsewhere. For more information about list experiments see Appendix I.

### **The Endorsement Experiment**

Endorsement experiments offer an indirect way of asking sensitive questions other than list experiments. It obscures individual responses by exploiting evaluation bias in human judgment. It draws on psychology, which demonstrates that people tend to evaluate identical objects positively (or negatively) when paired with favourable (or unfavourable) entities. It asks respondents non-sensitive questions but randomises whether these questions are paired with the sensitive object. Like list experiments, a sample of respondents is randomly divided into control and treatment groups. In the control group, respondents are

---

<sup>139</sup> Kramon and Weghorst, 2019; Kao and Lust, 2022

<sup>140</sup> Kramon and Weghorst, 2019; Kao and Lust, 2022

<sup>141</sup> Gonzalez-Ocantos, et al 2012; Holbrook, and Krosnick, 2010 cited in Lépine, Treibich, and D'Exelle, 2020

<sup>142</sup> McKenzie and Siegel, 2013 cited Lepine, et al., 2020

<sup>143</sup> Karlan and Zinman, 2012 cited in Lepine, et al., 2020

<sup>144</sup> Lax, Phillips, and Stollwerk, 201 cited in Lepine et al., 2020

<sup>145</sup> Krumpal, 2013 cited in Lepine, et al., 2020

asked to rate their level of support, for example, for a particular policy. For those in the treatment group, the same question is asked, except that the policy is said to be endorsed by an actor of interest. The main idea is to take advantage of subtle cues induced by endorsements (or names) and interpret the difference in responses between the treatment and control groups as evidence of support (or lack thereof) for the actor of interest. Typically, several policies are selected so that measurement does not rely on a single instrument and statistical power is increased by analysing them together. Respondents might be asked for example to indicate their level of support for the various policies on a 5-point Likert scale from strongly agree to strongly disagree.<sup>146</sup>

### **The Randomised Response Method**

This method asks respondents to use a randomisation device, such as a coin flip, whose outcome is unobserved by the interviewer. By introducing random noise, the method obscures or conceals individual responses and protects respondent privacy. Under a standard design, it asks respondents to use a randomising device e.g. coin flip and truthfully answer the sensitive question only when the randomisation results in a certain outcome. In other cases, respondents simply give a pre-determined response. Because enumerators do not know the outcome of the randomisation, they have no way of knowing whether respondents are answering the sensitive question or providing a pre-set response.<sup>147</sup>

### **Policy Framing**

A further method of indirectly assessing attitudes is the use of policy framing questions. In this method, participants are presented with a number of policies and asked to indicate their level of support for each policy. The policies differ slightly in terms of the trade-offs that would be required. For example, participants could be asked: 1) to indicate their support on a policy to increase social welfare payments for disabled people with no explicit trade-off mentioned; 2) to indicate their support for increased social welfare payments for disabled people that would be funded by budget re-allocation; or 3) to indicate their support for increased social welfare payments for disabled people that would be funded by increased general taxation. It could be expected that most people would indicate support for the first policy, but that less support might be elicited when trade-offs, that would impact the respondent, are made explicit. Thus, revealing attitudes that may otherwise may have been concealed. The NDA have recently commissioned the ESRI to indirectly examine attitudes towards disabled

---

<sup>146</sup> Blair, Imai, and Lyall, 2014

<sup>147</sup> Rosenfeld, Imai, and Shapiro, 2016

people in Ireland. Policy framing questions were incorporated into the experiment and the results will be reported elsewhere.

### **Threats to validity associated with indirect measures**

Although indirect measures have the potential to address some of the threats to validity associated with direct measures, such as reducing the occurrence of socially desirable responding, they are not without criticism. Although earlier work found that the IAT demonstrated high test-retest validity,<sup>148</sup> later work has challenged this.<sup>149</sup> Indirect measures of attitudes have since been shown to have low test-re-test reliability, exhibit low predictive validity and scores can vary due to context or situational factors. In addition, various indirect measures often correlate weakly with one another.<sup>150</sup> This may be a result of issues with reliability within the instruments used, or it may be the result of different task demands; leading some researchers to recommend that any conclusions about implicit attitudes should be based on multiple indirect measures.<sup>151</sup> Other researchers question the overall value of indirect measures as, in some cases, they have demonstrated similar, or poorer abilities to predict discriminatory behaviours than direct measures.<sup>152</sup> Some researchers have also argued that no indirect measures are completely free of a conscious component and may be amenable to voluntary control.<sup>153</sup> Fazio and Olsen (2003) state:

Nothing about our current implicit measurement procedures, be it a priming method, the IAT, or one of the other techniques mentioned earlier, guarantees that participants are unaware of their attitudes. p 9.

---

<sup>148</sup> Nosek, Greenwald, and Banaji, 2007

<sup>149</sup> Machery, 2017

<sup>150</sup> e.g. Golijani-Maghaddam et al., 2013; Machery, 2017 Oswald et al., 2013; Burke et al., 2019

<sup>151</sup> Nosek & Banaji, 2001; Kopera et al., 2015

<sup>152</sup> Oswald et al., 2013

<sup>153</sup> Fiedler and Bluemke, 2005; Klauer and Teige-Mocigemba, 2007; Steffens, 2004; Teige-I4 Mocigemba and Klauer, 2008; Teige-Mocigemba, Penzl, Becker, Henn, and Klauer, 2016 cited in Corneile and Huter, 2020

## Comparison of indirect and direct measures

### Evaluation of the List Method compared to other indirect techniques

A 2016 study empirically evaluated the performance of the list experiment, the randomised response method, the endorsement method and the direct questioning method against true information available by county for the 2011 Mississippi General Election, which included a sensitive personhood amendment.<sup>154</sup> The Mississippi study found that direct questioning led to underestimation of casting a “no” vote, which was the socially undesirable behaviour in the context of that referendum, by more than 20 percentage points in most counties. In contrast, all three indirect techniques provided estimates closer to the actual vote count. The randomised response method yielded a higher non-response rate than other indirect methods but was lower than the non-response rate with direct questioning, which is consistent with previous studies. While indirect methods reduced non-response and social desirability bias relative to the direct question, it produced estimates that were less efficient than direct questioning. Recent research has sought to address this issue by developing models that combine experimental techniques, such as list and endorsement, in order to recoup this loss of efficiency.

The research highlighted some of the strengths and weaknesses of each method and trade-offs to be considered when choosing a survey method for eliciting responses to sensitive questions might include the following:<sup>155</sup>

- Cost considerations/resource constraints: indirect questions involve using more space on a survey than a simple direct question. In applied work, the endorsement experiment is likely to be the most expensive, as gaining statistical power requires offering each respondent multiple endorsement questions. Both randomised response methods and list experiments require explanatory time, though each is more efficient than the endorsement experiment in terms of total responses for a given level of power.
- A randomised controlled study on the impact of interventions on sensitive attitudes may be less interested in bias than in statistical efficiency, because, under the assumption that the response bias is identical across treatment arms (and not so large that floor or ceiling effects interfere with estimates of attitudes in each group), the study will still recover a valid estimate of the

---

<sup>154</sup> *ibid*

<sup>155</sup> *ibid*

treatment effect. They might therefore choose direct questions even though the levels will be biased on sensitive traits.

- Researchers seeking to measure attitudes to evaluate claims about the prevalence of sensitive attitudes and behaviours may be more concerned with avoiding any bias and so choose randomised response.
- Cognitive difficulties for minimally educated populations may make some modes of indirect questioning less efficacious. Endorsement experiments have worked well with such populations in Afghanistan and Pakistan and the randomised response technique was successfully applied to an environment where gambling with dice is common.

Even with a common statistical framework, the relative performance of these indirect approaches may be context specific, in terms of the bias-efficiency trade-off as well as their ability to limit the influence of experimenter demand effects on survey responses. Ideally, applied researchers would have a body of validation studies to consult in choosing the method to use for their own research in a specific context.<sup>156</sup>

### **Relationship between indirect and direct measures**

Many studies have examined both implicit and explicit biases towards people with disabilities in tandem to assess whether or not they align and if and how they interact.<sup>157</sup> Research has consistently found that respondents tend to report positive attitudes towards people with disabilities on direct measures while their scores on indirect measures indicate more negative attitudes. For example, Ferrarea et al. (2015) compared scores on an indirect measure to scores on a direct measure to assess attitudes to people with Intellectual Disability. Overall, participants (college students) exhibited very positive attitudes towards disabled people on the direct measure but scores on the indirect measure indicated a bias in favour of non-disabled people. In another study, Friedman et al. (2019) examined explicit and implicit attitudes of family members of people with disabilities and also found that, although family members reported no preference between disabled and non-disabled people on the direct measures, scores on the indirect measures indicated a strong preference for non-disabled people. Similar findings have been observed elsewhere in relation to, for example, people on the autism spectrum,<sup>158</sup> and people with mental illness.<sup>159</sup>

---

<sup>156</sup> Rosenfeld, Imai and Shapiro, 2016

<sup>157</sup> Vazalli & Capozza, 2011 cited in Kallman, 2017

<sup>158</sup> Bast et al., 2020

<sup>159</sup> Kopera et al., 2015

In addition to how people score on direct and indirect measures, many studies have assessed whether statistical correlations exist between measures.<sup>160</sup> A majority of studies report weak or non-existent correlations between direct and indirect measures. In a systematic review, for example, Wilson and Scior (2014) found that, of the 11 studies that assessed relationships between scores on the IAT and scores on direct measures, five found no correlations while the remaining six found only weak correlations. Potential reasons for weak and non-existent correlations include:

- that implicit and explicit attitudes are two distinct but related constructs<sup>161</sup>
- factors related to the task procedure; for example, presenting direct measures before indirect measures have been found to produce higher correlations, potentially because people are primed by the answering question on the direct measure.<sup>162</sup>
- The social sensitivity of the topic – less sensitive topics tend to demonstrate larger correlations between direct and indirect measures than more sensitive topics.<sup>163</sup>
- The type of indirect and direct measures used - some may correlate more highly with each other than others. For example, explicit affective measures have been found to have stronger correlations with the IAT than other types of measures<sup>164</sup>
- The extent to which people have thought about the concepts.<sup>165</sup>
- The degree to which people view their opinions about the concepts to be distinct from others.<sup>166</sup>

### **Predictive value of direct and indirect measures**

Some evidence suggests that indirect measures may be better than direct measures at predicting unconscious, spontaneous, and automatic behaviours such

---

<sup>160</sup> e.g. González-Sanguino et al., 2020; James, 2018;

<sup>161</sup> Charlesworth and Banaji, 2019; Scior & Wilson, 2015

<sup>162</sup> Hoffman et al., 2005

<sup>163</sup> *ibid*

<sup>164</sup> Hoffman et al., 2005; Forscher et al., 2019

<sup>165</sup> Forecher et al., 2019

<sup>166</sup> Cameron et al., 2012; Greenwald et al., 2009; Hofmann et al., 2005; Nosek, 2005 cited in Forscher et al., 2019

as eye contact or verbal slips,<sup>167</sup> particularly in relation to socially sensitive topics.<sup>168</sup> Direct measures, on the other hand, may be better than indirect measures at predicting more deliberate, controlled behaviours.<sup>169</sup> This may be because when assessing more sensitive topics using direct measures, there is a greater likelihood that motivational factors (e.g. self-preservation) will exert some influence on responses.<sup>170</sup>

Greenwald and colleagues (2009), found that racially discriminatory behaviour was better predicted by scores on an indirect measure (i.e. the IAT) compared to self-report measures. Similarly, Green et al. (2007) found that while explicit measures of prejudice had no impact on doctors' behaviours towards Black patients, the more strongly the clinicians held implicit prejudices against Black people, the less likely they were to recommend a preferred treatment to their Black patients compared to White patients who presented the same symptoms.

Later research has challenged some of these findings however. In a meta-analysis of the predictive value of the IAT and direct measures, both were found, on average, to be weak predictors of various behaviours (e.g. micro-behaviour, interpersonal behaviour, person perceptions, and reaction times) in relation to race and ethnicity. While in some cases, the explicit measures performed similarly, and sometimes better than, the IATs on these criteria.<sup>171</sup> The authors concluded that had the researchers in the analysed studies taken certain steps, such as minimising reactivity bias, then the direct measures may have performed much better than the indirect measures at predicting behaviours.<sup>172</sup>

Examining the predictive value of direct and indirect measures on behaviour towards people with disabilities has also produced mixed results. In a recent study<sup>173</sup> examining employers' attitudes towards hiring people with disabilities, implicit attitudes towards disabled people were depicted as automatic and unconscious, while explicit attitudes were linked to controllability and self-reflectiveness. Similar to previous research, the authors argued that implicit

---

<sup>167</sup> Dabby et al., 2015; Stome and Motowski, 2011; Friese, Hofmann, & Schmitt, 2008; Hofmann, Gawronski, Gschwendner, Le, and Schmitt, 2005),

<sup>168</sup> Greenwald et al., 2009 cited in Forscher et al., 2019

<sup>169</sup> Devine, 1989; Fazio & Olson, 2014; Kurdi et al., 2018; cf. Greenwald et al., 2009 cited in Forscher et al., 2019; Dabby et al., 2015; Fazio and Olsen, 2003

<sup>170</sup> Fazio and Olsen, 2003

<sup>171</sup> Oswald et al., 2013

<sup>172</sup> *ibid*

<sup>173</sup> Florence and Marc, 2021

attitudes were more sensitive for assessing negative beliefs or veiled prejudice that people would not explicitly endorse or would prefer not to reveal.

In another study<sup>174</sup>, which examined non-autistic college students' face-to-face interactions with a confederate who they believed to be either autistic or non-autistic depending on the experimental condition, scores on indirect measures were associated with some non-verbal behaviours. For example, participants who demonstrated more implicit bias on the indirect measure smiled more when the confederate was labelled autistic but behaved in ways that were perceived to be neurotypical, and smiled less when the label and behaviour were associated with autism. Additionally, participants who demonstrated more bias on indirect measures leaned away from the confederate when the label and behaviour was associated with autism. No associations were found between scores on direct measures and any of the assessed behaviours.

---

<sup>174</sup> Lipson et al., 2020

## Modes of survey data collection

Various survey modes are used to measure attitudes towards disabilities. These include fact-to-face interviews, telephone surveys, paper and pen surveys, computerised surveys, and web-based surveys. A wealth of literature examining the effect of the mode of data collection in survey research has shown that the mode can impact on the quality or accuracy of the data gathered, particularly in relation to sensitive topics, often due to misreporting.<sup>175</sup> Misreporting had been shown to increase as the questions become more sensitive and decrease as the method of data collection becomes more private.<sup>176</sup> The main distinction between different modes of data collection is whether the survey is administered by an interviewer or is self-administered.

Surveys that are delivered by an interviewer can be subject to a number of interviewer effects, which are the undesirable active or passive influences of an interviewer on the answers obtained.<sup>177</sup> Passive influences include the interviewer's characteristics (e.g. gender, ethnicity, socio-economic status) and their assumed or perceived expectations. For example, if the respondent thinks that some of the response alternatives would be insulting, offensive or embarrassing to the interviewer they may adjust their answers to avoid this.<sup>178</sup> Active influences include instances when an interviewer might skip a questions or changes the wording because they feel uncomfortable asking it.<sup>179</sup> Similarly, an interviewer might suggest a particular response alternative to a respondent who asks for clarification.<sup>180</sup> Interviewer effects can be variable (i.e. the effect differs within and between interviewers) or systematic (i.e. there is no difference between the interviewers in how they influence the data). Two types of systematic effects can be identified: 1) all interviewers alter the same aspect of the interview with all respondents in the same way (i.e. pure bias); and 2) all interviewers alter the same aspect of the interview for all respondents but in various ways.<sup>181</sup> In a recent analysis of the impact of interviewer effects using data

---

<sup>175</sup> e.g. Langhaug et al., 2010; Des Jarlais et al. 1999; Holbrook et al. 2003; Metzger et al. 2000; Okamoto et al. 2002; Tourangeau et al. 1997; Tourangeau and Yan 2007; Turner et al. 1998, 2005 cited in Krumpal, 2013

<sup>176</sup> Ong and Weiss 2000 cited in Krumpal, 2013

<sup>177</sup> Loosdvelt, 2022

<sup>178</sup> Krumpal, 2013

<sup>179</sup> Krumpal, 2013

<sup>180</sup> Loosdvelt, 2022

<sup>181</sup> Loosdvelt, 2022

from several rounds of the European Social Survey, Loosdvelt (2022) found that interviewers had systematic effects on the answers and that some of the variance was explained by differences between interviewers. A cross-national analysis also indicated that interviewer effects were higher in some countries than others. The authors argue that such was the impact on the quality of data obtained, including on questions assessing attitudes, that evaluations of interviewer impact should be an essential part of data quality assessments and should be routinely reported along with other elements such as sampling design and response rates.

Methods of self-administration which minimize the presence of the interviewer, seem to reduce feelings of jeopardy, decrease the potential for emotions like shame and embarrassment, and increase respondents' sense of privacy, which may lead to more honest answers to sensitive questions such as discriminatory attitudes.<sup>182</sup> Computerised and web-based surveys have become a popular mode of self-administration over recent decades, with a number of reviews highlighting the advantages of computerised surveys on sensitive issues such as sexual or delinquent behaviours.<sup>183</sup> Web-based surveys are becoming ever more popular in social science research due to increasing internet usage and declining response rates to other survey modes.<sup>184</sup> Compared to other survey modes, they have many advantages including being cheaper and faster to conduct and being less susceptible to social desirability bias than other modes such as face-to-face and telephone surveys.<sup>185 186</sup>

In addition to the growth of the internet, the advent of the Covid-19 pandemic, has encouraged some researchers to investigate the advantages and trade-offs of moving from a face-to-face interview survey to a web-based survey.<sup>187</sup> In a recent study, May et al. (2022) investigated the psychometric properties of a new survey for measuring attitudes to people with Epilepsy and compared results from a web-based and face-to-face survey. The face-to-face survey was performed as part of a representative Omnibus Survey in Germany and was conducted as a Computer Assisted Personal Interview (CAPI) with a stratified, multi-level random sample of participants. The web-based survey was performed as Computer Assisted Web Interviews (CAWI) and participants were recruited

---

<sup>182</sup> Krumpal, 2013; D'Ancona, 2017

<sup>183</sup> Langhaug et al., 2010; Kleck and Roberts, 2012 cited in Gnamb & Kasper, 2016

<sup>184</sup> D'Ancona, 2017; May et al., 2022

<sup>185</sup> Henderson et al, 2012

<sup>186</sup> Kreuter et al., 2008; Sakshaug, Yan, and Tourangeau, 2010; Heerwegh and Loosveldt, 2008; Kreuter et al., 2008; Lindhjem and Navrud, 2011; Simmons and Bobo's, 2015

<sup>187</sup> British Social Attitudes Survey; May et al., 2022

through an online survey panel. The results found that the psychometric properties of the scales used varied only slightly by mode and that, in both surveys, fears and concerns were more pronounced than stereotypes and social distance.

Researchers from NatCen Social Research<sup>188</sup> have also investigated the advantages and trade-offs associated with moving from a face-to-face interview survey to a push-to-web or mixed modes<sup>189</sup>. They assessed the sample quality by comparing the achieved sample profile, the response rates<sup>190</sup> and the weighting efficiency. Little difference was observed between the two modes in terms of age, sex and region of the respondents. The web-based survey was found to be scalable, reduced respondent burden due to shorter time required to complete the survey, reduced fieldwork time by about 10 weeks and reduced the cost by about 50%. A number of trade-off were observed, however. People with higher education and those who were more politically engaged were over-represented in the web-based survey and under-represented in the face-to-face survey. The face-to-face survey was better at reaching harder to reach groups such as ethnic minorities, people with no qualifications, those who were economically inactive and social renters. The web-based survey achieved a much lower response rate and although the researchers suggest that the demographic representativeness was unlikely to have been unduly damaged, there were concerns related to the representativeness of behaviour or attitudes of respondents compared to the population. The web-based survey was also found to have lower levels of 'don't know' and 'prefer not to say' responses, especially for sensitive questions. The researchers concluded that the web-based survey provided a robust sample with comparable estimates at a lower cost and shorter time period to conduct.

---

<sup>188</sup>

[https://www.youtube.com/watch?v=pHr257x0vGA&t=3028s&ab\\_channel=EuropeanSocialSurvey](https://www.youtube.com/watch?v=pHr257x0vGA&t=3028s&ab_channel=EuropeanSocialSurvey)

<sup>189</sup> The push to web mode included an option for people to opt to take part by telephone in order to reach the offline population. The comparisons were not truly experimental, as, due to the Covid-19 pandemic, they were unable to run parallel face-to-face and web-based surveys, instead, they compared the web-based survey to the face to face survey from the previous year

<sup>190</sup> Household level response rate.

## Conclusion

There now exists an abundance of methods for measuring attitudes towards people with disabilities, both direct and indirect. This review has indicated, however, that no single approach is wholly reliable or uncontested. Measuring attitudes towards people with disabilities is not a straightforward task due to the complexity of the concept. A general consensus is that approaching the measurement of attitudes in a simplistic manner will not produce a holistic picture of attitudes. Antonak & Liveneh (2000) recommended that measuring attitudes towards people with disabilities:

requires innovative experimental methods and psychometrically sound instruments that are reliable, valid and multi-dimensional. Without such instruments, it will not be possible to obtain conclusive answers to important research questions concerning the relationship between these attitudes and the acceptance and integration of persons with disabilities into society.

Self-report survey research remains the most common method of assessing attitudes to disability as it is the most accessible and straightforward approach for many researchers. The main issues with self-report surveys and other direct measures, however, are the threats to validity such as social desirability bias. Thus, relying on direct measures alone may only provide information around how a person thinks they should feel and behave but may not necessarily reflect how they actually behave or intend to behave.<sup>191</sup> An examination of the modes of data collection indicates, however, that self-administered survey modes may help with eliciting more honest responses on direct measures regarding sensitive topics. Recent research investigating the trade-offs between the various survey modes has shown promising results, however, there are important caveats that need to be considered, particularly in terms of response rates and overall representativeness of the sample achieved.

Indirect measures may address some of the threats to validity associated with direct measures but they only provide part of the picture and are subject to their own challenges.<sup>192</sup> For example, some traditional indirect measures are lab-based and require specific equipment or computer programmes, making them less practical, especially if the aim is to conduct a nationally representative study. List experiments have shown promising results for the indirect measurement of

---

<sup>191</sup> Kopera, 2015; Kallman, 2017

<sup>192</sup> Girod, 2020

attitudes and this method is appropriate for conducting nationally representative studies with large samples. No studies were identified in which list experiments were used to examine attitudes towards people with disabilities, however, which is a clear gap in the literature. The NDA has commissioned the ESRI to conduct research which examines attitudes towards disabled people in Ireland using indirect methods including a list experiment which be a valuable addition to the evidence-base.

This literature review has indicated that there are various useful methods and approaches that can be used to measure attitudes towards people with disabilities. The measurement of attitudes is complex, however, and each methodology has advantages and drawbacks, so careful consideration should be given to the type of measurement that would be most suitable in various contexts.

## References

- Aaberg, V.A. (2012) A path to greater inclusivity through understanding implicit attitudes toward disability. **Journal of Nursing Education**, 51(9), pp.505-510. <https://doi.org/10.3928/01484834-20120706-02>
- Adderley, H. et al. (2020) Adult mental health practitioner beliefs about psychosis, parenting, and the role of the practitioner: A Q methodological investigation. **Psychology and Psychotherapy: Theory, Research and Practice**, 93(4), pp.657-673. <https://doi.org/10.1111/papt.12249>
- Ahadinezhad, B. et al. (2021) Implicit racial bias among medical graduates and students by an IAT measure: a systematic review and meta-analysis. **Irish Journal of Medical Science** (1971-), 191, pp.1941-1949. <https://doi.org/10.1007/s11845-021-02756-3>
- Andersson, J. et al. (2015) Employer attitudes toward hiring persons with disabilities: A vignette study in Sweden. **Journal of Vocational Rehabilitation**, 43(1), pp.41-50. <https://doi.org/10.3233/jvr-150753>
- Antonak, R.F. and Livneh, H. (1988) **The measurement of attitudes toward people with disabilities: methods, psychometrics and scales**. Springfield, IL: Thomas.
- Antonak, F.R. and Livneh, H., 2000. Measurement of attitudes towards persons with disabilities. *Disability and rehabilitation*, 22(5), pp.211-224. <https://doi.org/10.1080/096382800296782>
- Aronow P.M. et al (2015) Combining List Experiment and Direct Question Estimates of Sensitive Behaviour Prevalence. **Journal Survey Statistics and Methodology** 3(1):43-66. <https://doi.org/10.1093/jssam/smu023>
- Arora, A., Sandhu, H.S. and Brasch, J. (2019) The potential effect of the psychiatric clerkship and contact-based hypothesis on explicit and implicit stigmatizing attitudes of Canadian medical students towards mental illness. **Academic Psychiatry**, 43(6), pp.605-609. <https://doi.org/10.1007/s40596-019-01090-2>
- Avramidis, E. (2013) Self-concept, social position, and social participation of pupils with SEN in mainstream primary schools. **Research Papers in Education**, 28(4), pp.421-442. <https://doi.org/10.1080/02671522.2012.673006>

- Avramidis E. et al. (2017) Using sociometric techniques to assess the social impacts of inclusion: Some methodological considerations. **Educational Research Review**, 20, pp.68-80. <https://doi.org/10.1016/j.edurev.2016.11.004>
- Ball, P. et al. (2005) Disability as diversity in Fortune 100 companies. **Behavioral Sciences & the Law**, 23(1), pp.97-121. <https://doi.org/10.1002/bsl.629>
- Barnes-Holmes, D. et al. (2006) Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. **The Irish Psychologist**, 32(7), pp.169-177.
- Barr, J.J. and Bracchitta, K. (2015) Attitudes toward individuals with disabilities: The effects of contact with different disability types. **Current Psychology**, 34(2), pp.223-238. <https://doi.org/10.1007/s12144-014-9253-2>
- Bast, D.F. et al. (2020) The effect of educational messages on implicit and explicit attitudes towards individuals on the autism spectrum versus normally developing individuals. **The Psychological Record**, 70(1), pp.123-145. <https://doi.org/10.1007/s40732-019-00363-4>
- Benedetti, A., Jackson, J. and Luo, L. (2018) Vignettes: Implications for LIS research. **College & Research Libraries**, 79(2), p.222. <https://doi.org/10.5860/crl.79.2.222>
- Besika, A., Collard, P.C. and Coogan, J. (2018) Attitudes of therapists towards people with learning disabilities. **Counselling and Psychotherapy Research**, 18(1), pp.89-101. <https://doi.org/10.1002/capr.12139>
- Blair, G., Imai, K., and Lyall, J. (2014) Comparing and Combining List and Endorsement Experiments: Evidence from Afghanistan. **American Journal of Political Science**, 58 (4), 1043-1063. <https://doi.org/10.1111/ajps.12086>
- Bollier et al. (2021) Attitudes Matter. Findings from a national survey of community attitudes toward people with disability in Australia. Available from: <https://mspgh.unimelb.edu.au/research-groups/centre-for-health-equity/disability-and-health-unit/community-attitudes> [last accessed November, 2022].
- Breen, J.S. (2018) Attitudes toward employees with disabilities: A systematic review of self-report measures. **The Australian Journal of Rehabilitation Counselling**, 24(2), pp.67-87. <https://doi.org/10.1017/jrc.2018.8>
- Brown, T. et al. (2020) 'Measuring Attitudes: Current Practices in Health Professional Education', in Nestel, D. et al. (eds.), **Clinical Education for the**

**Health Professions**, Springer Nature Singapore Pte Ltd.  
[https://doi.org/10.1007/978-981-13-6106-7\\_76-1](https://doi.org/10.1007/978-981-13-6106-7_76-1)

Bryant, L.D., Green, J.M. and Hewison, J. 2006. Understandings of Down's Syndrome: A Q methodological investigation. **Social Science & Medicine**, 63(5), pp.1188-1200. <https://doi.org/10.1016/j.socscimed.2006.03.004>

Cabrera, I. et al. (2021) Development and validation of an Implicit Relational Assessment Procedure (IRAP) to measure implicit dysfunctional beliefs about caregiving in dementia family caregivers. **The Psychological Record**, 71(1), pp.41-54. <https://doi.org/10.1007/s40732-020-00445-8>

Campbell, J.M. (2008) Brief report: reliability and validity of the Shared Activities Questionnaire as a measure of middle school students' attitudes toward autism. **Journal of Autism and Developmental Disorders**, 38, pp.1598-1604. <https://doi.org/10.1007/s10803-007-0534-6>

Charlesworth, T.E. and Banaji, M.R. (2019) Patterns of implicit and explicit attitudes: I. Long-term change and stability from 2007 to 2016. **Psychological Science**, 30(2), pp.174-192. <https://doi.org/10.1177/0956797618813087>

Child, S. and Nind, M. (2013) Sociometric methods and difference: A force for good—or yet more harm. **Disability & Society**, 28(7), pp.1012-1023. <https://doi.org/10.1080/09687599.2012.741517>

Churrua, K. et al. (2021) A scoping review of Q-methodology in healthcare research. **BMC Medical Research Methodology**, 21(1), pp.1-17. <https://doi.org/10.1186/s12874-021-01309-7>

Cook, D. (1992) 'Psychological impact of disability', in R. M. Parker and E.M. Szymanski (eds.) **Rehabilitation Counseling Basics and Beyond**. Austin: TX: PRO-ED, pp. 249–272.

Cooke, R. et al. (2016) How well does the theory of planned behaviour predict alcohol consumption? A systematic review and meta-analysis, **Health Psychology Review**, 10(2), pp.148-167. <https://doi.org/10.1080/17437199.2014.947547>

Cooper, A., Rose, J. and Mason, O. (2004) Measuring the attitudes of human service professionals toward deafness. **American Annals of the Deaf**. 48(5). Pp.385-389. <http://www.ncbi.nlm.nih.gov/pubmed/15132018>

Corneille, O. and Hütter, M. (2020) Implicit? What do you mean? A comprehensive review of the delusive implicitness construct in attitude research. **Personality and Social Psychology Review**, 24(3), pp.212-232. <https://doi.org/10.1177/1088868320911325>

Crittenden, S.N.R (2018) The impact of supercrip representations on attitudes towards people with physical disabilities. Undergraduate Honors Thesis, 79. Available from: [https://ir.lib.uwo.ca/psychK\\_uht/79](https://ir.lib.uwo.ca/psychK_uht/79) [last accessed November 2022]

Cross, R.M. (2005) Exploring attitudes: the case for Q methodology. **Health Education Research**, 20(2), pp.206-213. <https://doi.org/10.1093/her/cyg121>

Cross-Sudworth F, Williams A, and Herron-Marx S. (2011) Maternity services in multi-cultural Britain: Using Q methodology to explore the views of first- and second-generation women of Pakistani origin. **Midwifery**, 27(4), pp.458–68. <https://doi.org/10.1016/j.midw.2010.03.001>

Cui, W. et al. (2015) Recent changes in attitudes of US adults toward people with epilepsy - Results from the 2005 SummerStyles and 2013 FallStyles surveys. **Epilepsy & Behavior**, 52(PtA), pp.108–118. <https://doi.org/10.1016/j.yebeh.2015.08.040>

Dabby, L., Tranulis, C. and Kirmayer, L.J. (2015) Explicit and implicit attitudes of Canadian psychiatrists toward people with mental illness. **The Canadian Journal of Psychiatry**, 60(10), pp.451-459. <https://doi.org/10.1177/070674371506001006>

D'Ancona, M.Á.C. (2017) Measuring multiple discrimination through a survey-based methodology. **Social Science Research**, 67, pp.239-251. <https://doi.org/10.1016/j.ssresearch.2017.04.006>

Dean R, et al. (2018) Staff perceptions of borderline personality disorder and recovery: A Q-sort method approach. **British Journal of Clinical Psychology**. 57(4):473–90. <https://doi.org/10.1111/bjc.12180>

DeCastellarnau A. (2018) A classification of response scale characteristics that affect data quality: a literature review. **Quality and Quantity**, 52(4), pp.1523–59. <https://doi.org/10.1007/s1135-017-0533-4>

De Laat, S., Freriksen, E. and Vervloed, M.P. (2013) Attitudes of children and adolescents toward persons who are deaf, blind, paralyzed or intellectually

disabled. **Research in Developmental Disabilities**, 34(2), pp.855-863.  
<https://doi.org/10.1016/j.ridd.2012.11.004>

Diseth, R.R., Bøgwald, K.P. and Høglend, P.A. (2011) Attitudes among stakeholders towards compulsory mental health care in Norway. **International Journal of Law and Psychiatry**, 34(1), pp.1-6.  
<https://doi.org/10.1016/j.ijlp.2010.11.001>

Domagała-Zyśk, E. (2021) Attitudes of different age groups toward people with intellectual disability during the COVID-19 pandemic. **Frontiers in Psychiatry**, p.1121. <https://doi.org/10.3389/fpsy.2021.591707>

Economou, M. et al. (2012) Medical students' beliefs and attitudes towards schizophrenia before and after undergraduate psychiatric training in Greece. **Psychiatry and Clinical Neurosciences**, 66(1), pp.17-25.  
<https://doi.org/10.1111/j.1440-1819.2011.02282.x>

Farrimond H. (2017) A typology of vaping: Identifying differing beliefs, motivations for use, identity and political interest amongst e-cigarette users. **International Journal of Drug Policy**, 48, pp.81–90.  
<https://doi.org/10.1016/j.drugpo.2017.07.011>.

Fazio, R.H. and Olson, M.A. (2003) Implicit measures in social cognition research: their meaning and use. **Annual Review of Psychology**, 54(1), pp.297-327.  
<https://doi.org/10.1146/annurev.psych.54.101601.145225>

Findler, L., Vilchinsky, N. and Werner, S. (2007) The multidimensional attitudes scale toward persons with disabilities (MAS) construction and validation. **Rehabilitation Counseling Bulletin**, 50(3), pp.166-176. <https://doi.org/10.1257/rct.8944-1.0>

FitzGerald, C. and Hurst, S. (2017) Implicit bias in healthcare professionals: a systematic review. **BMC Medical Ethics**, 18(1), pp.1-18.  
<https://doi.org/10.1186/s12910-017-0179-8>

Florence, Y. and Marc, C. (2021) Implicit and explicit attitudes of employers toward hiring people who have experienced depression. **Journal of Occupational Rehabilitation**, 31(4), pp.903-915.  
<https://doi.org/10.1007/s10926-021-09977-4>

Forscher P.S. et al. (2019) A meta-analysis of procedures to change implicit measures. **Journal Personality and Social Psychology**, 117(3), pp.522-559.  
<https://doi.org/10.1037/pspa0000160>

- Friedman C. (2019) Family members of people with disabilities' explicit and implicit disability attitudes. **Rehabilitation Psychology**, 64(2), pp.203-211 <https://doi.org/10.1037/rep0000265>
- Friedman, C. and VanPuymbrouck, L. (2021). Impact of occupational therapy education on students' disability attitudes: a longitudinal study. **The American Journal of Occupational Therapy**, 75(4). <https://doi.org/10.5014/ajot.2021.047423>
- Gidman, W. et al. (2009) Delivering health care through community pharmacies: Are working conditions deterring female pharmacists' participation? **Journal of Health Services Research and Policy**, 14(3), pp.141–149. <https://doi.org/10.1258/jhsrp.2009.008077>.
- Girod, R., Thompson, P. and Davis, C. (2022) Genetic counseling students' and recent graduates' attitudes toward psychiatric illness. **Journal of Genetic Counseling**, 31(3), pp.713-721. <https://doi.org/10.1002/jgc4.1538>
- Gnambs, T. and Kaspar, K. (2015). Disclosure of sensitive behaviors across self-administered survey modes: a meta-analysis. **Behavior Research Methods**, 47(4), pp.1237-1259. <https://doi.org/10.3758/s13428-014-0533-4>
- Golijani-Moghaddam, N., Hart, A. and Dawson, D.L.. (2013) The implicit relational assessment procedure: Emerging reliability and validity data. **Journal of Contextual Behavioral Science**, 2(3-4), pp.105-119. <https://doi.org/10.1016/j.jcbs.2013.05.002>
- Gonzalez-Ocantos, E. et al. (2012) Vote buying and social desirability bias: experimental evidence from Nicaragua, **American Journal Political Science**. 56 (1), 202-217. <https://doi.org/10.1111/j.1540-5907.2011.00540.x>
- González-Sanguino, C. et al. (2019) Study of the relationship between implicit and explicit stigmas associated with mental illness. **Psychiatry Research**, 272, pp.663-668. <https://doi.org/10.1016/j.psychres.2018.12.172>
- Gould, R. et al. (2020) Disability, diversity, and corporate social responsibility: learning from recognized leaders in inclusion. **Journal of Vocational Rehabilitation**, 52(1), pp.29-42. <https://doi.org/10.3233/jvr-191058>

- Green, A.R., et al. (2007) Implicit bias among physicians and its prediction of thrombolysis decisions for black and white patients. **Journal of General Internal Medicine**, 22, pp.1231–8. <https://doi.org/10.1007/s11606-007-0258-5>
- Greenwald, A.G., Banaji, M.R., and Nosek, B.A. (2015) Statistically small effects of the Implicit Association Test can have societally large effects. **Journal of Personality and Social Psychology**, 108(4), 553–561. <https://doi.org/10.1037/pspa0000016>
- Grimshaw, P., McGowan, L. and McNichol E. 2017. An extra care community's perceived priorities for 'whole system' relationships: a Q-methodological study. **Health and Social Care in the Community**, 25(3), pp.1169–80. <https://doi.org/10.1111/hsc.12417>
- Groomes, D.A. and Linkowski, D.C. (2007) Examining the structure of the revised acceptance disability scale. **Journal of Rehabilitation**, 73(3), p.3. <https://www.proquest.com/openview/28a49fdc372bb195e73d29ed3fd2a88e/1?pq-origsite=gscholar&cbl=37110>
- Henderson, C. et al. (2012) Responses to mental health stigma questions: the importance of social desirability and data collection method. **Canadian Journal of Psychiatry**, 57(3), pp.152–60. <https://doi.org/10.1177/070674371205700304>
- Hofmann, W. et al. (2005) A meta-analysis on the correlation between the implicit association test and explicit self-report measures. **Personality and Social Psychology Bulletin**, 31(10), pp.1369–1385. <https://doi.org/10.1177/0146167205275613>
- Hughes, S. and Barnes-Holmes, D. (2011) On the formation and persistence of implicit attitudes: New evidence from the Implicit Relational Assessment Procedure (IRAP). **The Psychological Record**, 61(3), pp.391-410. <https://doi.org/10.1007/bf03395768>
- Hughes, R. and Huby, M. (2002). The application of vignettes in social and nursing research. **Journal of Advanced Nursing**, 37(4), pp.382-386. <https://doi.org/10.1046/j.1365-2648.2002.02100.x>
- James, J.S. (2018). Explicit, Implicit, and Behavioral Stigmatization of Mental Illness (Doctoral dissertation, The University of Southern Mississippi). <https://www.proquest.com/openview/59915191b0f729cf13c4d85d62c02a59/1?pq-origsite=gscholar&cbl=18750>

Jenaro, C., Flores, N. and López-Lucas, J. (2018). Using the semantic differential technique to assess stereotypes toward individuals with disabilities: The relevance of warmth and competence. **Universitas Psychologica**, 17(4), pp.1-12 <https://doi.org/10.11144/javeriana.upsy17-4.usdt>

Kallman, D. (2017) Integrating disability: Boomerang effects when using positive media exemplars to reduce disability prejudice. **International Journal of Disability, Development and Education**, 64(6), pp.644-662. <https://doi.org/10.1080/1034912X.2017.1316012>

Kao, K. and Lust, E. (2022) Do List Experiments Run as Expected? Examining Implementation Failure in Kenya, Zambia, and Malawi Working Paper No. 57 The Program on Governance and Local Development University of Gothenberg <https://www.gld.gu.se/media/2684/gld-working-paper-57.pdf>

Karpinski, A., Steinman, R.B. and Hilton, J.L.. (2005) Attitude importance as a moderator of the relationship between implicit and explicit attitude measures. **Personality and Social Psychology Bulletin**, 31(7), pp.949-962. <https://doi.org/10.1037/e633912013-681>

Keith, J.M., Bennetto, L. and Rogge, R.D. (2015) The relationship between contact and attitudes: Reducing prejudice toward individuals with intellectual and developmental disabilities. **Research In Developmental Disabilities**, 47, pp.14-26. <https://doi.org/10.1016/j.ridd.2015.07.032>

Kelly, A. and Barnes-Holmes, D. (2013) Implicit attitudes towards children with autism versus normally developing children as predictors of professional burnout and psychopathology. **Research in Developmental Disabilities**, 34(1), pp.17-28. <https://doi.org/10.1016/j.ridd.2012.07.018>

Kenward, L. (2019) A literature review to guide novice researchers using Q methodology in the development of a framework for concourse management. *Nurse Researcher*, 27 (1). pp. 17-21 <https://doi.org/10.7748/nr.2019.e1616>

Krull, J., Wilbert, J. and Hennemann, T. (2014) The social and emotional situation of first graders with classroom behavior problems and classroom learning difficulties in inclusive classes. **Learning Disabilities: A Contemporary Journal**, 12(2), 169-190. <https://eric.ed.gov/?id=EJ1047460>

Krischler, M., Pit-ten Cate, I.M. and Krolak-Schwerdt, S. (2018) Mixed stereotype content and attitudes toward students with special educational needs and their

- inclusion in regular schools in Luxembourg. **Research in Developmental Disabilities**, 75, pp.59-67. <https://doi.org/10.1016/j.ridd.2018.02.007>
- Krumpal, I., 2013. Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & quantity*, 47(4), pp.2025-2047.
- Kramon, E., Weghorst, K (2019) (Mis)measuring sensitive attitudes with the list experiment: Solutions to list experiment breakdown in Kenya. **Public Opinion Quarterly** ,83(S1), pp. 236- 63 <https://doi.org/10.1093/poq/nfz009>
- Kopera, M. et al. (2015) Evaluating explicit and implicit stigma of mental illness in mental health professionals and medical students. **Community Mental Health Journal**, 51(5), pp.628-634. <https://doi.org/10.1007/s10597-014-9796-6>
- Kuhne, M. and Wiener, J. (2000). Stability of social status of children with and without learning disabilities. **Learning Disability Quarterly**, 23(1), pp. 64-75. <https://doi.org/10.2307/1511100>
- Langhaug, L.F., Sherr, L. and Cowan, F.M. (2010) How to improve the validity of sexual behaviour reporting: systematic review of questionnaire delivery modes in developing countries. **Tropical Medicine & International Health**, 15(3), pp.362-381. <https://doi.org/10.1111/j.1365-3156.2009.02464.x>
- Lépine, A., Treibich, C. and D'Exelle, B. (2020) Nothing but the truth: Consistency and efficiency of the list experiment method for the measurement of sensitive health behaviours, **Social Science & Medicine**, 266, 1-10 <https://doi.org/10.1016/j.socscimed.2020.113326>
- Lipson, J. et al. (2020) Perceptions of and behavior toward university students with autism. **Basic and Applied Social Psychology**, 42(5), pp.354-368. <https://doi.org/10.1080/01973533.2020.1785468>
- Loch, A.A. et al. (2014) Patterns of stigma toward schizophrenia among the general population: a latent profile analysis. **International Journal of Social Psychiatry**, 60(6), pp.595-605. <https://doi.org/10.1177/0020764013507248>
- Loosveldt, G. (2022) Interviewers and their impact on survey quality: lessons for EU-SILC from the European Social Survey. Available from: <https://lirias.kuleuven.be/retrieve/670504> [last accessed November 2022]
- Lum, J. et al. (2021) Implicit associations related to physical disability among nursing students. **Disability and Health Journal**, 14(4), p.101150. <https://doi.org/10.1016/j.dhjo.2021.101150>

Lundy, C. (2019) **The Stigmatization of Concealable and Apparent Intellectual Disabilities**. Unpublished PhD thesis. University of Mississippi. Available from: [https://egrove.olemiss.edu/hon\\_thesis/1090/](https://egrove.olemiss.edu/hon_thesis/1090/) [last accessed November 2022]

Lust, E. et al. (2019) 'The Local Governance and Performance Index (LGPI) 2019: Kenya, Malawi, Zambia'"The Program on Governance and Local Development, University of Gothenburg. Available from: <https://gld.gu.se/media/2265/lgpi-concept-note.pdf> [last accessed November 2022]

Machery, E. (2016) 'De-Freuding implicit attitudes' in M. Browstein and J. Saul (eds.), **Implicit Bias and Philosophy: Metaphysics and Epistemology** Oxford, UK: Oxford University Press, pp. 104–129

Matthews, N.L., Ly, A.R., and Goldberg, W.A. (2015) College students' perceptions of peers with autism spectrum disorder. **Journal of Autism and Developmental Disorders**, 45(1), 90-99. <https://doi.org/10.1007/s10803-014-2195-6>

Marquis, L. (2017) Conformism in opinions about the welfare state in Switzerland: results from a List Experiment no. 69 IEPHI (Institute of Political, Historical and International Studies, University of Lausanne) Working Paper Series. Available from: <https://isidore.science/document/10670/1.obltne> [last accessed November 2022]

May, T.W. et al. (2022) Attitudes toward epilepsy assessed by the SAPE questionnaire in Germany—Comparison of its psychometric properties and results in a web-based vs. face-to-face survey. **Epilepsy & Behavior**, 130, p.108665. <https://doi.org/10.1016/j.yebeh.2022.108665>

McCaughey, T.J. and Strohmer, D.C. (2005) Prototypes as an indirect measure of attitudes toward disability groups. **Rehabilitation Counseling Bulletin**, 48(2), pp.89-99. <https://doi.org/10.1177/00343552050480020301>

McDonnall, M.C., Crudden, A. and O'Mally, J. (2015) Predictors of employer attitudes toward people who are blind or visually impaired as employees. **Journal of Vocational Rehabilitation**, 42(1), pp.41-50. <https://doi.org/10.3233/jvr-140722>

- McDonnall, M.C. et al. (2019) Professionals' implicit attitudes about the competence of people who are blind. **Journal of Visual Impairment & Blindness**, 113(4), pp.341-354. <https://doi.org/10.1177/0145482X19865391>
- Mckenzie, J. et al. (2011) A case for the use of Q-methodology in disability research: lessons learned from a training workshop. **Disability and Rehabilitation**, 33(21-22), pp.2134-2141. <https://doi.org/10.3109/09638288.2011.560327>
- Meza, J.I. et al. (2019) Stigma and attention-deficit/hyperactivity disorder: negative perceptions and anger emotional reactions mediate the link between active symptoms and social distance. **ADHD Attention Deficit and Hyperactivity Disorders**, 11(4), pp.373-382.. <https://doi.org/10.1007/s12402-019-00302-x>
- Morin, D. et al. (2013) Public attitudes towards intellectual disability: A multidimensional perspective. **Journal of Intellectual Disability Research**, 57(3), pp.279-292. <https://doi.org/10.1111/jir.12008>
- Morin, D., Crocker, A.G., Beaulieu-Bergeron, R. and Caron, J. (2013) Validation of the attitudes toward intellectual disability–ATTID questionnaire. **Journal of Intellectual Disability Research**, 57(3), pp.268-278. <https://doi.org/10.1111/j.1365-2788.2012.01559.x>
- Morin, D. et al. (2013) Public attitudes towards intellectual disability: A multidimensional perspective. **Journal of Intellectual Disability Research**, 57(3), pp.279-292. <https://doi.org/10.1111/jir.12008>
- Murch, A.J. et al. (2018). Explicit and implicit attitudes towards people with intellectual disabilities: The role of contact and participant demographics. **Journal of Applied Research in Intellectual Disabilities**, 31(5), pp.778-784. <https://doi.org/10.1111/jar.12429>
- Nepi, L. D. et al. (2015). Social acceptance and the choosing of favourite classmates: A comparison between students with special educational needs and typically developing students in a context of full inclusion. **British Journal of Special Education**, 42(3), 319-337. <https://doi.org/10.1111/1467-8578.12096>
- Nosek, B.A. and Banaji, M.R. (2001) The go/no-go association task. **Social cognition**, 19(6), pp.625-666.
- Nowicki, E.A. (2003) A meta-analysis of the social competence of children with learning disabilities compared to classmates of low and average to high achievement. **Learning Disability Quarterly**, 26(3), 171-188 <https://doi.org/10.2307/1593650>

- Olson, J.M., and Zanna, M.P. (1993) Attitudes and attitude change. **Annual Review of Psychology**, 44, 117–154. <https://doi.org/10.1146/annurev.ps.44.020193.001001>
- Oswald, F.L. et al. (2013) Predicting ethnic and racial discrimination: a meta-analysis of IAT criterion studies. **Journal of Personality and Social Psychology**, 105(2), p.171. <https://doi.org/10.1037/a0032734>
- Paige, J.B. and Morin, K.H. (2016) Q-sample construction: A critical step for a Q-methodological study. **Western Journal of Nursing Research**, 38(1), pp.96-110. <https://doi.org/10.1177/0193945914545177>
- Palad, Y.Y. et al. (2016) Scoping review of instruments measuring attitudes toward disability. **Disability and Health Journal**, 9(3), pp.354-374. <https://doi.org/10.1016/j.dhjo.2016.01.008>
- Parker, R.B., Larkin, T. and Cockburn, J. (2018) Gender bias in medical images affects students' implicit but not explicit gender attitudes. **AERA Open**, 4(3), <https://doi.org/10.1177/2332858418798832>
- Pelleboer-Gunnink, H.A., van Weeghel, J. and Embregts, P.J. (2021) Public stigmatisation of people with intellectual disabilities: a mixed-method population survey into stereotypes and their relationship with familiarity and discrimination. **Disability and Rehabilitation**, 43(4), pp.489-497. <https://doi.org/10.1080/09638288.2019.1630678>
- Perugini, M., Richetin, J., & Zogmaister, C. (2010) 'Prediction of behavior', In B. Gawronski & B. K. Payne (eds.), **Handbook of Implicit Social Cognition: Measurement, Theory, and Applications**, NY: The Guilford Press, pp. 255–277
- Power, M.J., Green, A.M. and WHOQOL-DIS Group (2010) The Attitudes to Disability Scale (ADS): development and psychometric properties. **Journal of Intellectual Disability Research**, 54(9), pp.860-874. <https://doi.org/10.1111/j.1365-2788.2010.01317.x>
- Pruett, S.R. and Chan, F. (2006). The development and psychometric validation of the Disability Attitude Implicit Association Test. **Rehabilitation Psychology**, 51(3), p.202. <https://doi.org/10.1037/0090-5550.51.3.202>
- Robb, J. and Stone, J. (2016) Implicit bias toward people with mental illness: A systematic literature review. **Journal of Rehabilitation**, 82(4), p.3-13

<https://www.proquest.com/openview/2575bb4369b73611a8e5e0b667bec286/1?pq-origsite=gscholar&cbl=37110>

Rohmer, O. and Louvet, E. (2018) Implicit stereotyping against people with disability. **Group Processes & Intergroup Relations**, 21(1):127-40.

<https://doi.org/10.1177/1368430216638536>

Rolstad, S., Adler, J. and Rydén, A. (2011) Response burden and questionnaire length: is shorter better? A review and meta-analysis. **Value in Health**, 14(8), pp.1101-1108. <https://doi.org/10.1016/j.jval.2011.06.003>

Rosenfeld, B., Imai, K., and Shapiro, JN (2016) An empirical validation study of popular survey methodologies for sensitive questions. **American Journal of Political Science**, 60(3), 783-802. <https://doi.org/10.1111/ajps.12205>

Schoenberg, N.E. and Ravdal, H. (2000) Using vignettes in awareness and attitudinal research. **International Journal of Social Research Methodology**, 3(1), pp.63-74. <https://doi.org/10.1080/136455700294932>

Simane-Vigante, L., Plotka, I. and Blumenau, N. (2015) Research of attitudes towards criminal violence with implicit and explicit measures of cognition. **Journal of Education, Psychology and Social Sciences**, 3(2), pp.72-77.

Stone, J. and Moskowitz, G.B. (2011) Non-conscious bias in medical decision making: what can be done to reduce it?. **Medical Education**, 45(8), pp.768-776. <https://doi.org/10.1111/j.1365-2923.2011.04026.x>

Takahashi, J. and Haga, A. (2016) Evaluation of impressions toward people with intellectual and developmental disabilities using semantic differential method. **International Journal of Humanities and Social Science**, 6, pp.257-262.

Thomas, A., Vaughn, E.D., Doyle, A. and Bubb, R., 2014. Implicit association tests of attitudes toward persons with disabilities. **The Journal of Experimental Education**, 82(2), pp.184-204.

United Nations. (n.d.). Factsheet on Persons with Disabilities. Available from: <https://www.un.org/development/desa/disabilities/resources/factsheet-on-persons-with-disabilities.html> [last accessed November 2022]

Verster, J.C. et al. (2021) The Use of Single-Item Ratings Versus Traditional Multiple-Item Questionnaires to Assess Mood and Health. **European Journal of Investigation in Health Psychology and Education**, 11(1), pp.183-198  
<https://doi.org/10.3390/ejihpe11010015>

Vezzali, L. and Capozza, D. (2011) Reducing explicit and implicit prejudice toward disabled colleagues: Effects of contact and membership salience in the workplace. *Life Span and Disability*, 14, 139-162

Walstra, R. and Chukwuma, E. (2022) Inclusion of Disability Within the Spectrum of Diversity and the Implications for Accounting Education. **Issues in Accounting Education**. <https://doi.org/10.2308/issues-2021-007>

Wilson, M.C. and Scior, K. (2014) Attitudes towards individuals with disabilities as measured by the Implicit Association Test: A literature review. **Research in Developmental Disabilities**, 35(2), pp.294-321. <https://doi.org/10.1016/j.ridd.2013.11.003>

Wolbring, G. and Lillywhite, A. (2021) Equity/Equality, Diversity, and Inclusion (EDI) in Universities: The Case of Disabled People. **Societies**, 11(2), p.49. <https://doi.org/10.3390/soc11020049>

Wigboldus, D. H. J., Holland, R. W., & van Knippenberg, A. (2004). Single target implicit associations. Unpublished manuscript

Xu, K., Nosek, B., and Greenwald, A.G. (2014). Psychology data from the Race Implicit Association Test on the Project Implicit Demo website. **Journal of Open Psychology Data**, 2(1), e3. DOI: <http://doi.org/10.5334/jopd.ac>

Young, R.E., et al. (2019) The subtle side of stigma: Understanding and reducing mental illness stigma from a contemporary prejudice perspective. **Journal of Social Issues**, 75(3), pp.943-971. <https://doi.org/10.1111/josi.12343>

## Appendix I: Additional information on list experiments

The protection of privacy in the list experiment is designed to increase willingness of respondents to provide truthful answers to sensitive questions. To ensure that the list experiment works, individuals allocated to each group must be similar such that, on average, they agree with the same number of non-sensitive statements. The addition of the sensitive item must not change the sum of affirmative answers on the control items. Individuals may respond untruthfully if they no longer benefit from the privacy of their responses because they either agree or disagree with all the non-sensitive items (ceiling and floor effects). To eliminate this problem, there should be one non-sensitive item that most participants would agree with and another non-sensitive item that most participants would disagree with. Some people advise that one should choose non-sensitive items related to the topic of the behaviour or opinion investigated in the list experiment to avoid any suspicion on the part of respondents. The choice of the non-sensitive items is key to implementing the list experiment method successfully. Several studies advise that the non-sensitive items should be familiar to the respondent and similar in nature and specificity to the sensitive item so as not to introduce bias in the answers. The non-sensitive items should not themselves be susceptible to social desirability bias. In summary, the effectiveness of the list experiment methodology is based on assumptions that include the following:<sup>193</sup>

1. The successful randomisation of the treatment.
2. The absence of design effects.
3. The absence of ceiling and floor effects.

There are many conditions that list experiments must satisfy to be accurate.<sup>194</sup> There is some debate on whether the list experiment is reliable for obtaining accurate prevalence estimates. Several studies report challenges in terms of consistency of the prevalence estimated using list experiment techniques.<sup>195</sup> The technique can be inefficient and it requires a high number of respondents to detect effects.<sup>196</sup> The list experiment adds random noise<sup>197</sup> to the data so a trade-

---

<sup>193</sup> Lépine, Treibich, D'Exelle, 2020

<sup>194</sup> Blair and Imai, 2012 cited in Kao and Lust, 2022

<sup>195</sup> Bell et al. 2019 cited in Lepine, A et al. 2020

<sup>196</sup> Blair, Coppock, Moor, 2020 cited in Kao, and Lust, 2022

<sup>197</sup> The notion behind noisy data is that it is not necessarily precise and might not be able to be duplicated if the same information was collected or calculated again. Noise can be introduced by

off arises between potential bias reduction and the efficiency of the estimates. When the estimated prevalence of a sensitive behaviour is higher when obtained with a list experiment than with direct reports, it suggests that the higher estimated prevalence rates are more accurate, but there is not enough evidence that they are free from bias.<sup>198</sup>

The format of list experiments can be confusing<sup>199</sup> and decades of survey research had found that complicated survey questions are difficult to answer, particularly for certain subgroups.<sup>200</sup> Thus, while the list experiment can reduce social bias it can introduce different forms of error associated with question complexity and difficulty.<sup>201</sup> Efforts are being made to address these issues including the use of images and clearer instructions for enumerators. Kramon and Weghorst (2019) showed that 40 percent of respondents provided inconsistent responses between list experiment and direct question formats with simple, non-sensitive lists about food consumption and daily activities. These errors concentrated among the less numerate and less educated respondents, offering evidence that errors are driven by the complexity and difficulty of list experiments. Using the standard list experiment to measure attitudes about political violence resulted in lower rates of support for political violence compared to asking directly about this sensitive attitude, which the authors interpreted as list experiment breakdown. They evaluated modifications to the list experiment to reduce its complexity, which enhanced list experiment performance, especially among respondent subgroups where the standard procedure is most problematic.<sup>202</sup>

Kao and Lust (2022) investigated list experiment implementation failure and the drawbacks and accuracy of the technique.<sup>203</sup> The experiment was part of a larger survey on governance and local development in Kenya, Malawi, and Zambia (N~20,000) run in April to June of 2019<sup>204</sup> and analysed a list experiment across three countries. The experiment asked respondents: “Do you feel that the last question was too complicated to answer properly”? They asked the enumerator

---

measurement tools and random errors introduced by processing or in data gathering. [Noisy data](#) is less meaningful due to the existence of a lot of variation.

<sup>198</sup> Lépine, Treibich, D’Exelle, 2020

<sup>199</sup> Kramon and Weghorst, 2019

<sup>200</sup> Krosnick, J.A., 1991 cited in Kao and Lust, E., 2022

<sup>201</sup> Kramon and Weghorst, 2019

<sup>202</sup> *ibid*

<sup>203</sup> Kao and Lust, E., 2022

<sup>204</sup> Lust, E. et al., 2019

two questions: “Was the respondent able to count up the number of items without your help?” and “Did the respondent reveal to you which items they had done in the past?” The first question to the respondent indicated cognitive overload while the questions to the enumerators revealed disclosure of responses. The research found that respondents disclosed their answers more than half (55%) of the time and that the majority (79%) required help counting the answers. This problem is particularly prevalent among the less educated and older respondents, and the complicated nature of the question mediates the relationships between age and education and revealing answers. Response disclosure undermines the purpose of a list experiment. The authors reflected:

Although we do not claim that all list experiments suffer from implementation errors to the extent that those in our experiment seem to, our findings indicate the need for further investigation of these issues. Recognising the implementation problems of list experiments is important. A census of list experiments led scholars to conclude that sensitivity bias is rather limited.<sup>205</sup> However, our results suggest that this finding may reflect the prevalence of implementation errors more than reality. Apparent differences in attitudes or reported behaviours may result from different abilities to engage with the experiment. We encourage scholars to follow advice on reducing the complexity of list experiments<sup>206</sup> and include questions about implementation problems as standard follow-ups to understand if they worked properly in the field. Doing so will help scholars track and potentially account for subgroup differences in faculties leading to list experiment implementation errors.<sup>207</sup>

Researchers are increasingly taking into consideration best practices in list experiment. Some recommend that researchers use direct questioning as well as the list experiment:

Direct questions reveal reliable information about those who answer “Yes.” Among those who answer “No,” we cannot directly distinguish those who withhold from those who do not engage in the sensitive behaviour – for these subjects, list experiments may provide a workaround. Combining these two techniques into a single estimator yields more precise estimates of prevalence, and

---

<sup>205</sup> Blair and Imai, 2012.

<sup>206</sup> Kramon, and Weghorst, 2019

<sup>207</sup> Kao and Lust, 2022

employing direct and list questions in tandem also enables the researcher to test crucial identifying assumptions.<sup>208</sup>

---

<sup>208</sup> Aronow et al. 2015