

**RESURRECTING THE ERROR CHOICE TECHNIQUE:
The premature demise of an indirect measure of attitude?**

by

RONALD DOUGLAS PORTER

A thesis submitted to the Department of Psychology
in conformity with the requirements for
the degree of Doctor of Philosophy

Queen's University
Kingston, Ontario, Canada

(April, 2010)

Copyright © Ronald Douglas Porter, 2010

Abstract

The error choice (EC) technique was among the earliest indirect attitude measures developed and was originally designed to overcome social desirability concerns (Hammond, 1948). This programme of research set out to advance EC research in several ways. First, an exploratory factor analysis examined whether participants' responses to the EC target items produced a systematic pattern of responding. The factor analysis indicated that a single underlying factor best accounted for the data. Additionally, the EC measure demonstrated good reliability across the 3 studies.

Second, these studies provided evidence that the EC measure is, at least in part, attitudinal. That is, the EC measure showed a modest positive correlation with the direct measure of attitude in all 3 studies. This suggested that participants' responses to the EC target items were, at least in part, attitudinal. Additionally, across the studies participant's EC scores did not change significantly between the high and low social desirability conditions, while participants' scores on the direct measure were significantly more positive in the high social desirability condition. These findings suggest that the EC measure is, to some degree, resistant to socially desirable responding.

Studies 2 and 3 also represent the first time that recommendations made by early EC researchers to improve the EC technique have been systematically examined. In Studies 2 and 3 the amount of time participants had to complete the EC measure was restricted. The time restriction did not improve the performance of the EC measure. The other optimal condition examined in Study 3 was the effect of the presence or absence of filler items in the EC measure. Removing the filler items from the EC measure did not negatively impact its performance.

Finally, this programme of research compared the EC measure with more contemporary indirect measures of attitude. In Study 2, the Implicit Association Test (IAT) was compared with the EC and direct measure. In Study 3, the Personalized IAT was compared to the EC and direct measure. The results revealed that neither the IAT nor PIAT correlated with the EC measure. In summary, these results suggest that the EC technique holds some promise as an approach to attitude measurement and is well worth resurrecting.

Acknowledgements

I would like to extend my sincere gratitude to Dr. Leandre Fabrigar for his endless patience, encouragement, guidance and enthusiasm (regardless of how hard I attempted to make even the simplest task). I am also indebted to my committee members Dr. Tara MacDonald and Dr. Ronald Holden for their insightful reviews and advice, which appreciably added to the overall quality of my dissertation. I was also very fortunate in being able to work with a number of highly motivated and talented research assistants. In this regard, I would like to thank Natalie Watts, Andrea Tang and Nicole Vigneault who were instrumental in arranging and conducting the experimental sessions. This is especially true of Tory Dickenson, to whom I would like to extend a very special thank you for her tireless effort over the 3 years of data collection, running session and training new research assistants. Finally, thanks to my family for their unwavering support and encouragement.

Table of Contents

Abstract	ii
Acknowledgements	iv
Table of Contents	v
List of Tables	xi
List of Figures	xi
Chapter 1	1
General Introduction	1
<i>Socially Desirable Responding</i>	1
<i>Approaches to Overcoming the Effects of Socially Desirable Responding</i>	1
<i>Direct measurement of socially desirable responding</i>	1
<i>Procedural approaches to overcoming socially desirable responding</i>	2
<i>Indirect Measures of Attitudes</i>	4
<i>The psychoanalytic roots of indirect measures</i>	4
<i>Behaviour as an indirect measure of attitude</i>	5
<i>A structured objective technique of indirect attitude measurement</i>	6
<i>Reviews and Criticisms of Indirect Measures of Attitude</i>	7
<i>Reliability of indirect measures of attitude</i>	8
<i>The validity of indirect measures of attitude</i>	8
<i>Resistance to social desirability of indirect measures of attitude</i>	9
<i>Ethical concerns relating to indirect measures of attitude</i>	10
<i>The final assessment of indirect measures of attitude</i>	12
<i>Renewed Interest in the Indirect Measurement of Attitudes</i>	12

<i>Resurrecting the Error Choice Technique</i>	16
<i>The Error Choice Technique</i>	16
<i>Overview</i>	23
Chapter 2	25
Determining Whether the Systematic Response Pattern is Related to Attitudes and is Resistant to Socially Desirable Responding.	25
Method	26
<i>Participants</i>	26
<i>Procedure</i>	26
<i>Measures</i>	28
<i>The Error Choice (EC) measure of attitude</i>	28
<i>Direct measure of attitude</i>	30
<i>Self-monitoring scale</i>	31
Results.....	31
<i>Systematic Response Pattern</i>	31
<i>Exploratory factor analysis</i>	31
<i>EC measure reliability</i>	40
<i>Interpreting the correlations between the measures</i>	42
<i>Evaluating the effect of socially desirable responding</i>	42
<i>Self-monitoring scale</i>	44
<i>Summary of Findings</i>	46
Chapter 3	47
Providing an Initial Examination of an Optimal Feature of the Error Choice (EC) Technique and Comparing the EC Technique with a Current Indirect Measure of Attitude.	47

Method	47
<i>Participants</i>	47
<i>Procedure</i>	48
<i>Social desirability.</i>	48
<i>Time restriction.</i>	48
<i>Presentation of indirect measures of attitude.</i>	49
<i>Measures</i>	49
<i>The Error-Choice (EC) Measure.</i>	49
<i>The Direct Measure of Attitude</i>	50
<i>The Implicit Association Test (IAT).</i>	50
<i>Self-monitoring scale.</i>	53
<i>The Balanced Inventory of Desirable Responding – Version 6.</i>	53
<i>Results</i>	54
<i>Systematic Response Pattern</i>	54
<i>Exploratory factor analysis</i>	54
<i>Interpreting the correlations between the measures.</i>	60
<i>Evaluating the effect of socially desirable responding.</i>	62
<i>Direct measure of attitude.</i>	62
<i>EC measure of attitude.</i>	62
<i>The Implicit Association Test (IAT).</i>	64
<i>Self-monitoring scale.</i>	64
<i>The Balanced Inventory of Desirable Responding – Version 6.</i>	65
<i>Discussion.</i>	67

Chapter 4	69
Examining Item Ratio, an Optimal Feature of the Error Choice (EC) Technique and Comparing the EC with the Personalized Implicit Association Test.....	69
Method	71
<i>Participants</i>	71
<i>Procedure</i>	71
<i>EC filler to target item ratio.</i>	72
<i>Measures</i>	72
<i>The Error-Choice (EC) Measure.</i>	72
<i>The Direct Measure of Attitude</i>	72
<i>The Personalized Implicit Association Test (PIAT)</i>	73
<i>Self-monitoring scale.</i>	74
<i>The Balanced Inventory of Desirable Responding – Version 6</i>	75
Results.....	75
<i>Systematic Response Pattern</i>	75
<i>Exploratory factor analysis</i>	75
<i>Interpreting the correlations between the measures</i>	83
<i>Evaluating the effect of socially desirable responding.</i>	85
<i>Direct measure of attitude.</i>	85
<i>EC measure of attitude</i>	86
<i>The PIAT.</i>	88
<i>Self-monitoring scale.</i>	88
<i>The Balanced Inventory of Desirable Responding – Version 6</i>	89

<i>Discussion</i>	90
Chapter 5	92
General Discussion	92
<i>Summary of Findings and Implications</i>	92
<i>Insights into the functioning of the Error Choice technique</i>	92
<i>General implications for early indirect measures of attitude</i>	97
<i>Theoretical implications of the Error Choice technique</i>	97
<i>Practical implications of the Error Choice technique</i>	99
<i>Unresolved Issues in the Current Programme of Research</i>	99
<i>Modest correlations between the EC and direct measure of attitude</i>	99
<i>Other unresolved issues</i>	105
<i>Self-monitoring (SM) and impression management (IM) scales</i>	106
<i>Directions for Future Error Choice Technique Research</i>	106
<i>Conclusion</i>	109
References	110
Appendix A: Ethics Material	119
Appendix B: Social Desirability Manipulation Scripts	124
Appendix C: Study 1 Measures	127
Appendix D: Study 1 Correlation Table	135
Appendix E: Study 2 Measures	137
Appendix F: Study 2 Correlation Table	148
Appendix G: Study 3 Measures	150
Appendix H: Study 3 Correlation Table	161

List of Tables

Table 1. <i>Parallel Analysis – Eigenvalues from the Reduced Correlation Matrix.</i>	34
Table 2. <i>RMSEA Goodness-of-Model Fit</i>	35
Table 3. <i>Factor Matrix for the 1 Factor Model.</i>	37
Table 4. <i>Pattern Matrix for the 2 Factor Model.</i>	38
Table 5. <i>Factor Pattern Matrix for the 3 Factor Model.</i>	40
Table 6. <i>Factor Matrix 1 Factor Model (8 EC Items).</i>	41
Table 7. <i>Parallel Analysis – Eigenvalues from the Reduced Correlation Matrix.</i>	55
Table 8. <i>RMSEA Goodness-of-Model Fit.</i>	56
Table 9. <i>Factor Matrix for the 1 Factor Model.</i>	57
Table 10. <i>Pattern Matrix for the 2 Factor Model.</i>	58
Table 11. <i>Factor Matrix for the Time Condition.</i>	60
Table 12. <i>Parallel Analysis – Eigenvalues from the Reduced Correlation Matrix.</i>	76
Table 13. <i>RMSEA Goodness-of-Model Fit.</i>	77
Table 14. <i>Factor Matrix for the 1 Factor Model.</i>	78
Table 15. <i>Pattern Matrix for the 2 Factor Model.</i>	79
Table 16. <i>Factor Pattern Matrix for the 3 Factor Model.</i>	80
Table 17. <i>Factor Matrix for the Time Condition.</i>	81
Table 18. <i>Factor Matrix for the Filler Item Condition.</i>	82

List of Figures

Figure 1. Number of peer-reviewed journal articles on indirect measures between 1940 and June 2009.13

Figure 2. Scree Plot of Eigenvalues from the Reduced Correlation Matrix.....35

Figure 3. Scree Plot of Eigenvalues from the Reduced Correlation Matrix.....54

Figure 4. Effects of Time on Socially Desirable Responding.64

Figure 5. Scree Plot of Eigenvalues from the Reduced Correlation Matrix.....75

Figure 6. Effects of Filler Items on Socially Desirable Responding.....87

Figure 7. Effects of Time Restriction on Socially Desirable Responding.87

CHAPTER 1

General Introduction

Socially Desirable Responding

Imagine if you will . . . you are being interviewed for a job when you are asked how you feel about working for an aboriginal; or a Black canvasser shows up at your front door and asks you your opinion regarding Barack Obama. In the event that you held an unfavourable attitude toward the group or persons in question, would you shape your responses in an attempt to be viewed favourably; or would you provide an honest response to the questions? Social scientists have found when answering questions about socially sensitive topics people will often shape their responses in the hope of creating a favourable impression.¹ This phenomenon is known as socially desirable responding and is probably one of the most well-known and studied measurement response biases in personality and attitude research (Lemon, 1973; Paulhus, 1991).

Approaches to Overcoming the Effects of Socially Desirable Responding.

Direct measurement of socially desirable responding. Socially desirable responding has been a long standing topic for social scientists, who for over 80 years have been concerned with accounting for its effects on the validity of self-report

¹ Although there are two types of socially desirable responding discussed in personality and attitude research: impression management (when people respond to socially sensitive questions in order to create a positive image) and self-deceptive positivity (a person's tendency to give overly positive but honest self-reports), this research will focus on the former.

measures (Paulhus, 1991). Over the years, social scientists have developed several different types of approaches to account for the effects of social desirability. One such approach is the development of instruments that directly measure socially desirable responding. These types of measures allow the researcher to assess whether the content of the instrument is being confounded by socially desirable responding, by either administering a separate socially desirable responding measure (e.g., the Marlowe-Crowne Scale, the Balanced Inventory of Desirable Responding, etc.) or by incorporating a socially desirable responding scale into the instrument (e.g., the Minnesota Multiphasic Personality Inventory K-Scale and Minnesota Multiphasic Personality Inventory Lie Scale). This approach enables social scientists to account for the effects of socially desirable responding by either establishing a cutscore so that outliers can be identified and their scores discarded from the data, or by statistically partialling out the effects of socially desirable responding (Paulhus, 1991).

Procedural approaches to overcoming socially desirable responding. The research situation itself can influence the behaviour of participants causing their responses not to be representative of their true attitudes. As a result, researchers have employed a variety of procedural methods to control for social desirability. The simplest approaches to controlling for socially desirability responding involve increasing the participant's perceptions of anonymity. Often, this can be accomplished by providing the participant verbal and written assurances of anonymity. More elaborate techniques for overcoming socially desirable responding involve the use of deception, sometimes entailing elaborate techniques that require either informing the participant that a device is capable of accurately detecting their attitudes or is capable of

lie detection. Perhaps one of the best known of these types of techniques is the bogus pipeline (Jones & Sigall, 1971). The basic premise of this technique is that the participant is led to believe that researcher has the ability to detect their true attitude regardless of whether or not it is overtly expressed. This was accomplished by first having the participant complete a pre-screening measure sufficiently far enough in the past that they would not recognise the questions or their responses.

For example, when conducting the bogus pipeline procedure the participants arrive at the laboratory to participate in the experiment and they are attached to an apparatus that appears to be capable of recording their physiological responses. They are then informed that the measurements produced by the apparatus accurately capture their true attitudes. As proof of the accuracy of the device, the participant is asked to intentionally give incorrect answers to questions, which (without the participant's knowledge) the researcher had already obtained the correct answers to as part of the pre-screening measure. With each incorrect response the machine provides visual and/or audible signals that clearly indicate that the participant's attempt at deception had been detected. This demonstration of the machine's "efficacy" decreases the likelihood that participants misreport their attitudes, even when answering socially sensitive questions because they do not want to appear to be a liar or out of touch with themselves. Although these types of approaches have met with some success in reducing socially desirable responding in the laboratory (Roese & Jamieson, 1993), their elaborate methodological requirements have not always made them practical.

Indirect Measures of Attitudes.

Another approach used to control for socially desirability responding is to disguise the true purpose of the measure so that the participant is unaware of the construct being studied by the researcher. To that end, indirect methods of measurement control for socially desirability responding by measuring a construct without the participant being aware of the true purpose of the measure.² The underlying premise of this approach is that if the person is unaware of the construct being measured then they will not engage in socially desirable responding. There were a number of early attempts to develop indirect measures of attitude including projective techniques (Proshansky, 1943; Riddleberger & Motz, 1957; Vaughan & Thompson, 1961), behavioural measurement (Milgram, Mann, & Harter, 1965; Webb, Campbell, Schwartz & Sechrest, 1966), and objective semi-structured approaches (Coffin, 1941; Hammond, 1948; Newcomb, 1940).

The psychoanalytic roots of indirect measures. Early attempts to develop indirect measures of attitude evolved from psychoanalytic personality theory and its use of projective techniques to identify unconscious aspects of personality (Coffin, 1941, Fromme, 1941; Murray, 1938). Beginning in the late 1930's a number of attitude researchers hypothesized that differences between verbally expressed attitudes and unconscious thoughts and feelings could be identified by presenting ambiguous partially structured attitudinal stimuli (e.g., Thematic Appreciation Test). For example,

² When referring to attitude measures that disguise the topic of investigation some attitudes researchers will use either "implicit" and "indirect" interchangeably. In order to avoid any confusion, throughout this dissertation these types of measures will be referred to as "indirect" measures of attitude.

Proshansky (1943) examined attitudes towards organised labour using an attitude scale and an adapted Thematic Appreciation Test (TAT) and found evidence that this technique was measuring attitudes. In this approach, the participant's responses to the TAT pictures were coded by experts for their evaluative content. Proshansky (1943) developed a TAT that was comprised of ambiguous pictures relating to organised labour. He found that participants' written responses to the TAT did produce a significant correlation with an attitude survey (developed by Newcomb, 1940) for both the anti and pro labour groups ($r = .87$; and $r = .67$, respectively). Although similar results were reported by other researchers (Coffin, 1941; Proshansky & Murray, 1942; Riddleberger & Motz, 1957) none of these studies provided any evidence that this attitude measurement technique was less susceptible to socially desirable responding than direct measures of attitude.

Behaviour as an indirect measure of attitude. Another indirect approach to measuring attitude is through the examination of a person's observable behaviour. The underlying premise of this approach is that, if someone has a favourable or unfavourable attitude towards an attitude object, then it would presumably be reflected in their behaviour towards that attitude object. One of the best known examples of this approach is the Lost Letter Technique (Milgram, Mann, & Harter, 1965). In this technique, a specific attitude object is identified (i.e., abortion). Preaddressed and stamped letters are left in a variety of public locations. Each envelope is addressed to an organisation that someone could clearly identify as being either for-or-against a specific attitude object. The researcher then tracks how many letters are delivered to each addressee. The underlying assumption of this technique is that when a letter was

found, people would be more likely to place it in a mailbox if it was addressed to an organisation that represented their attitude, thereby providing a rough estimate of the popularity of each position. In this way, the person is not affected by social desirability because no one, other than themselves, would be aware of their actions. Although the Lost Letter Technique provides an overall estimate of attitudes in a specific region (i.e., the area where the letters were distributed) it does not provide any information on an individual's specific attitudes (Fabrigar, Krosnick, & MacDougall, 2005).

A structured objective technique of indirect attitude measurement. Another early indirect approach to attitude measurement involved the use of structured objective questionnaires. That is to say, participants were given what they believed was an objective information test that would assess their knowledge on a particular subject; however, some of the questions were not objective and had no correct response. Rather, researchers developed a number of questions whose responses were intentionally weighted for-or-against an attitude object and randomly dispersed within the information test (Coffin, 1941; Hammond, 1948; Kubany, 1953; Newcomb, 1940; Newcomb, 1946; Smith, 1947; Weschler, 1950a; Weschler, 1950b). The underlying premise of this approach is that there is a relationship between a person's attitudes and how they interpret information presented as fact. These researchers believed that when a person is presented with a question for which they do not know the correct response their guessing would be "weighted towards the subject's attitude" (Newcomb, 1946; p.293).

For example, Newcomb (1946) examined students' attitudes toward the Spanish Civil War at three American universities (each university was located in a community

that had demonstrated strong support for either the Loyalists, or the Nationalists). Newcomb contended that within these special communities the Spanish Civil War was an important topic (self-determination by the majority versus foreign aggression/oppression by the minority) that formed a “close attitude-information relationship” (Newcomb, 1946; p.291). That is, the student bodies at two of the colleges had been exposed to numerous lectures, movies and discussions that were pro-Loyalist (Bennington and Williams), while the other college (a Catholic university) had received comparable exposure to pro-Nationalist material. As expected, Newcomb (1946) found that there were correlations between the participants’ responses (from their respective colleges) and their attitudes towards the Spanish Civil War in the expected directions (Bennington, $r = -.57$; Williams, $r = -.34$; Catholic university, $r = .51$). The reliability coefficient for the measures was 0.7 for the Catholic university and 0.8 for Bennington College. Unfortunately, like the projective and behavioural techniques mentioned earlier these studies of structured objective indirect measures (Coffin, 1941; Hammond, 1948; Kubany, 1953; Newcomb, 1940; Newcomb, 1946; Smith, 1947; Weschler, 1950a; Weschler, 1950b) provide no evidence to suggest that they are less susceptible to socially desirable responding.

Reviews and Criticisms of Indirect Measures of Attitude.

In the wake of the rapidly growing body of research on early indirect measures of attitude, a number of reviews emerged (Campbell, 1950; Kidder & Campbell, 1970; McNemar, 1946). Of these, perhaps the most detailed and comprehensive review was that of Kidder and Campbell (1970), who provided a systematic review of the various types of indirect attitude measures that had been developed at that time. Their review

provided a number of detailed criticisms of indirect measures of attitude, which fell into four general categories: poor validity, poor reliability, lack of evidence regarding resistance to social desirability, and ethical concerns.

Reliability of indirect measures of attitude. One of Kidder and Campbell's (1970) criticisms focused on the reliability of indirect attitude measures. To support this concern, they cited numerous studies in which indirect attitude measures had demonstrably low levels of reliability. For example, they reported that in one study (Burwen & Campbell, 1957; cited in Kidder & Campbell, 1970) a TAT had low reliability (.55), which was comparable to other projective techniques used in the study (e.g., photo judging task .46, an adjective check list .24 and an autobiographical inventory .56). As a result, of these and other similar results they concluded that although the reported reliabilities for these measures were "among the highest of their kind . . . they are generally below the levels considered respectable" (p. 369). That being said, earlier in their review Kidder and Campbell acknowledged that Johnson (1949; cited in Kidder and Campbell, 1970), using a TAT methodology had obtained a very high reliability (.90). Additionally, some of the other indirect measures of attitude they discussed in their review (i.e., structured objective techniques) had shown very respectable levels of reliability (e.g., ranging from .78 to .87; Hammond, 1948).

The validity of indirect measures of attitude. Kidder and Campbell were also critical of the criterion validity of indirect measures. As evidence of poor criterion validity, Kidder and Campbell stated that in the very few studies where both indirect and direct measures of attitude were compared with a criterion variable, the indirect measures failed to reach "the levels of the direct tests" (p.369). They stated that this

may have been the result of researchers selecting inappropriate criterion variables. That is, Kidder and Campbell suggested that each type of attitude measure may be tapping into a different type of attitude; consequently, they would each require a different type of criterion measure. Kidder and Campbell stated that many of our attitudes (those best suited for indirect measurement) were a result of learning and past experience, which represent the people's true attitudes. Additionally, people have other attitudes which are "artificial and molded by social consideration" (p. 370) and because of this are susceptible to normative influence, which would be more appropriately measured with an indirect approach. As a result of these differences, one would expect indirect measures to correlate very well with other indirect measures, but not necessarily with direct measures of attitude. However, as described by Kidder and Campbell, past research revealed non-significant or very low correlations between indirect and direct measures of attitude, and between indirect measures when more than one was used in the study.

Resistance to social desirability of indirect measures of attitude. Perhaps most ironically, Kidder and Campbell (1970) were harshly critical of early indirect measure researchers for not having demonstrated that these types of measures were more resistant to socially desirable responding than direct measures of attitude, which had originally been purported as the primary advantage of this approach to attitude measurement. Kidder and Campbell pointed out, in the approximately 20 years leading up to their review social scientists who had been studying indirect measures of attitude had not provided any evidence that these measures were less susceptible to socially

desirable responding than direct measures of attitude, but rather had shown the exact opposite.

The authors described several studies in which participant's anonymity was varied (i.e., signing of names to test sheets or seeing how others were responding) and found that only the scores from the indirect attitude measures were adversely affected by the manipulation (Cook, Johnson, & Scott, 1969; cited in Kidder and Campbell, 1970; Kidder, 1969). That is, participants' scores on the indirect attitude measures varied significantly depending on whether they were in the anonymous or identified condition, while the scores on the direct attitude measures did not change significantly. Based on these findings Kidder and Campbell concluded that indirect measures were significantly affected by social desirability and that indirect measures were not adding any value with regards to attitude measurement and research.

Ethical concerns relating to indirect measures of attitude. Kidder and Campbell also voiced strong ethical concerns regarding the development and use of indirect attitude measures (although on closer examination these ethical concerns were not unique to indirect measures of attitude). Their first concern focused on the possible unethical application of indirect attitude measures of attitude. For example, the authors suggested that within the area of industrial psychology the application of indirect attitude measures of attitude could lead to unfair personnel selection practices on the part of management. They suggested that indirect attitude measures could be used to select out applicants whose attitudes might be anti-management or otherwise incongruent with the organisation, or limit someone's opportunities for employment or advancement.

These concerns were also raised by Weschler (1951) who concluded that indirect attitude measures were very susceptible to misuse, and consequently proposed that a special procedural code needed to be developed. In his experiment Weschler (1950b) found that mediators who scored in the mid-range (i.e., “neutral zone”) of the attitude measure were assessed by their peers as being better mediators than those whose scores reflected either a more pro-labour or pro-management attitude. Upon seeing the results of the study, a number of participants complained that they had been taken advantage of because they had been misled about the true purpose of the test and that the published results violated their confidence and limited their future employment possibilities. Fortunately this was not the case, because the results of the Weschler (1950b) study were only ever reported in aggregate form so no individual or their corresponding score could be identified.

Another ethical concern raised by Kidder and Campbell was that some social scientists were motivated to conduct indirect measures of attitude research because it afforded these researchers the opportunity to deceive and exploit people. In this criticism, Kidder and Campbell’s attempted to impart motive to indirect measure researchers. It was their position that, for some indirect attitude measure researchers, it was the act of deception and taking advantage of a “gullible sucker” that were the very pleasurable ends of their studies. Again, no specific examples of this unethical behaviour on the part of indirect measure researchers were provided by Kidder and Campbell. Consequently, it is unclear why they would conjecture that indirect measure researchers might be more susceptible to this type of unethical behaviour than other social scientists. Certainly some deception was used by researchers in their studies of

indirect measures (i.e., being informed that they were completing an “information test”) but there is no evidence to suggest that it was unusually or unnecessarily higher than one might find in attitude studies using direct measures.

The final assessment of indirect measures of attitude. Kidder and Campbell (1970) concluded their review of indirect measures by describing them as “novel and creative” (p. 375), and stated that the novelty of the indirect measures might allow them to play a supporting role in attitude measurement, as a member of multi-method approaches. That being said, Kidder and Campbell made it very clear that indirect measures are immensely fallible, as reflected in the disappointing nature of the results from the studies they reviewed, and given the ethical concerns, were simply not worth the risk. In the years that followed the Kidder and Campbell review, interest in conducting research in indirect measures of attitude appears to have waned.

Renewed Interest in the Indirect Measurement of Attitudes. Although no one can attribute any clear trends to a single review, nevertheless there does seem to have been a sharp decrease in indirect attitude measurement research in the wake of the Kidder and Campbell (1970) review. This is supported by the graph in Figure 1, which is based on the results of an informal keyword search conducted on the PSYCInfo data base using *projective/attitude*, *indirect/attitude* and *implicit/attitude*. As represented in the graph, after this area of research emerged in the late 1930’s, interest continued to grow for the next 20 years at which time the number of articles remained fairly constant until the late 1960s, after 1970 there was a sharp drop for more than two decades. When interpreting the graph it should be remembered that the data reflect the raw

numbers of articles at that time. Given the expanded number journals and total research output

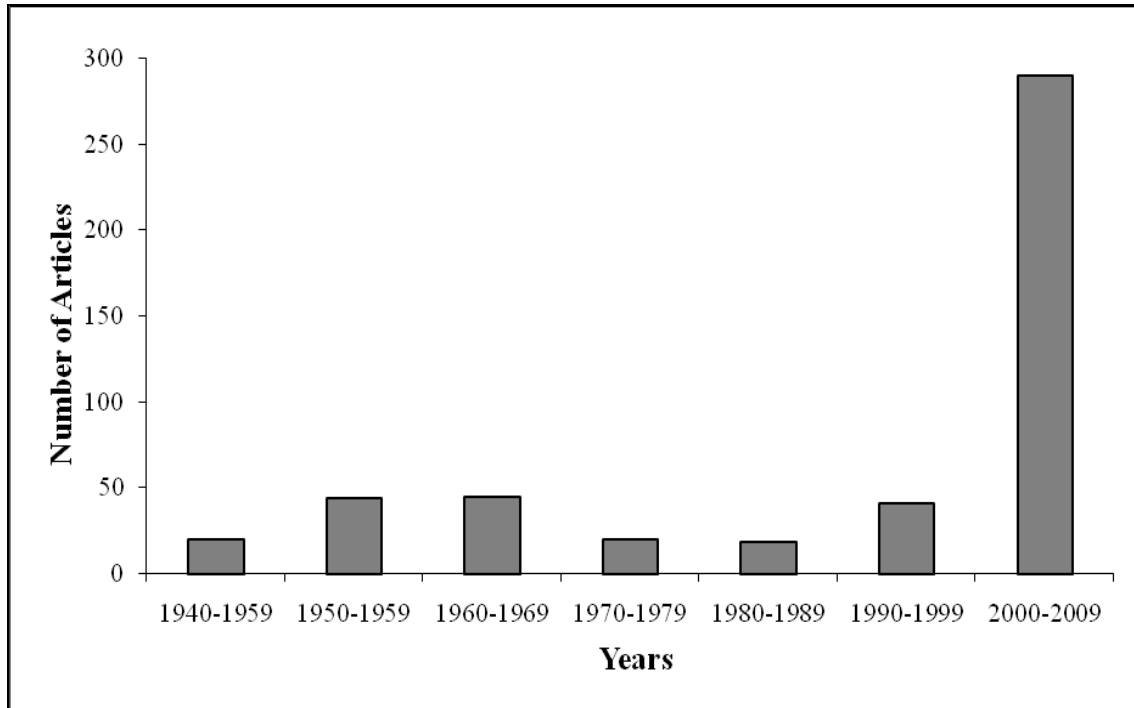


Figure 1. Number of peer-reviewed journal articles on indirect measures between 1940 and June 2009.

that currently exist, if there was a method of gauging the percentage of total articles the drop might be even steeper. Then, as can be seen in the graph, in the 1990s there was a strong surge in interest that seems to have reached an almost unprecedented level as of the last 10 years (Wittenbrink, 2007). This raises the very intriguing question of why this sudden surge of interest?

Obviously concerns about social desirability and other response biases had not gone away, but that cannot account for this increase because those very same concerns existed in the prior decades. Nor was the increase in indirect measures fuelled by

renewed interest in the old indirect attitude measurement techniques. Rather, researchers over the past 15 or so years have begun to propose a number of new methods for the indirect measurement of attitudes. Some of the increase might be the emergence of new technologies, which have suggested new approaches to attitude measurement. Certainly some of the new physiological measures like functional Magnetic Resonance Imagery (Cunningham, Johnson, Gatenby, Gore, & Banaji, 2003) and Event Related Potential (Crites, Cacioppo, & Gardner, 1995) were not available in the 1940s through 1960s, but in point of fact the physiological articles account for only a very small portion of the increase in indirect measurement articles. So it appears as though it is not the new physiological measures that drove the upswing in the indirect measurement of attitude research.

Many of these new indirect measurement methods relied upon response latency techniques and cognitive priming effects drawn from social cognition (e.g., Neely, 1976). The premise underlying these techniques is that evaluations attached to one stimulus facilitate judgements of a second stimulus if the evaluations are shared. So, if a particular stimulus has a particular evaluation (i.e., positive or negative) associated with it, then it will facilitate the speed of responding to a second stimuli that shares a congruent evaluations and inhibit incongruent or contradictory evaluations (Fazio, Sanbonmatsu, Powell, & Kardes, 1986). Within the area of social psychology, probably the best known of these techniques are the Implicit Association Test (Greenwald, McGee, & Schwartz, 1998) and Evaluative Priming (Fazio, Jackson, Dunton, & Williams, 1995).

One might infer from this that the upswing in indirect measurement of attitude research is related to the response latency methodology, however, this approach had existed for many years before it was employed in attitude measurement research (Neely, 1976). Indeed, if this was the case, then the increased research interest should have commenced following Fazio et al. (1995) evaluative priming research, but in fact in the three years that followed publication of this article there were only about fifteen additional articles (using the same search terms as Figure 1). Rather, it was not until after Greenwald et al. (1998) Implicit Association Test (IAT) article that the spike in research appeared, as reflected in the number of articles ($n = 54$) published in the three subsequent years. On closer examination of all the data for the graph in Figure 1, the vast majority of the published articles were related to the Implicit Association Test (IAT). It appears then, that the IAT is almost certainly responsible for this sharp increase.

The IAT was put forth not just as an indirect measure of attitude, but Greenwald et al. (1998) also posited that this measure tapped into a fundamentally different kind of construct from self-report measures, which they called an implicit attitude. Compared to an explicit attitude (those attitudes of which people are consciously aware), implicit attitudes are defined as attitudes that are unconscious and people are unaware they possess, but nonetheless influence their behaviour. Whether it is in fact the case that these types of indirect measures provided a direct and clear pipeline into this construct is unclear and remains a matter of great debate among social psychologists. Interestingly, the idea of implicit attitudes bears striking similarity to the speculations advanced by Kidder and Campbell (1970) and earlier still by others (Coffin, 1941;

Proshansky, 1943; Riddleberger & Motz, 1957). These social psychologists suggested that the poor correlations that existed between direct and indirect measures of attitude might be a result of indirect measures tapping into a different kind of construct than direct measures, but this was never really followed up on or embraced by researchers at the time.

Resurrecting the Error Choice Technique.

Notwithstanding the aforementioned explosion in indirect measurement research, the traditional indirect measures have largely been abandoned and the real focus has been on the development of new indirect measures of attitude. This might intuitively lead one to a couple of conclusions. First, that it has been established that the old indirect measures of attitude are highly problematic and have no value. Secondly, that these new indirect measures are clearly superior. But in fact, a close examination of the literature suggests it is unclear whether either of these conclusions is entirely supportable. For example, in the Kidder and Campbell (1970) review, it appears as though some indirect measures that had shown some promise were lost in their overall condemnation of the indirect approach to attitude measurement. Probably the best example of this is provided by the Error Choice Technique (Hammond, 1948).

The Error Choice Technique. In an attempt to overcome concerns about socially desirable responding, Hammond (1948) developed an indirect measure of attitude called the Error-Choice (EC) Technique. The EC technique rests on two basic premises. First, when people are faced with a knowledge-based question for which they do not know the answer, their guess will not be random; and one factor that they might rely upon in such guessing is their attitude. So for example, when faced with a question where there are

two factual possible answers they will tend to pick the answer that best fits with their attitude. Thus across a series of such questions, one might expect to find a systematic guess pattern that is consistent with their attitude. The second premise of the measure is that, because each of the items is presented as a factual question, people will not be aware that their attitude is being assessed. To test his theory, Hammond conducted a series of studies examining the EC technique using the known groups paradigm. For example, in one study Hammond constructed an EC measure that consisted of 20 Error Choice questions (target items) and 20 factual questions (filler items) concerning attitudes towards Labour and Management.

The target (EC) items for this study were partially structured questions (i.e., multiple choice), each of which had in principle an objective answer, but it was extremely unlikely that participants would know the correct answer. Furthermore, each target item provided participants with two response options, where the chosen response might tend to imply something either positive or negative about the attitude object. For example, in one EC target question, Hammond asked participants to estimate the percentage of each dollar companies earned that is profit and then presented them response options of 16¢ and 3¢. If the pattern of guessing for EC items was systematic and not random, then pro-labour participants would choose a response with a higher percentage/value would also select similar responses on the other EC target items. The participant's responses could then be summed and the resulting value would reflect their attitudes towards the attitude object.

The filler (factual) items used in the EC measure were also partially structured; however, these questions were not evaluative in nature. Rather, the filler items were

knowledge based questions about the topic where the correct answers were knowable. For example, in one filler item Hammond asked participants to identify the average weekly wage of the US war workers in 1945, and then presented them with response options of \$37 and \$57. It is certainly possible that participants in this 1948 study (i.e., members of a labour union and businessmen) might have known the correct answer to these questions, but knowing the correct responses would not necessarily be informative of the person's attitude. Hammond contended that the filler items played a crucial role in the EC measure because they helped disguise the true intent of the measure.

In Hammond's (1948) original EC study, he used a known groups paradigm (management versus labour), and developed an instrument that would take advantage of a participant's selective recall and distortion of perception. The fundamental assumptions of the known groups paradigm are that attitude objects exist on a continuum, and that each of the chosen groups represents diametric positions along that continuum. Therefore, based on group membership one should be able to predict the overall response patterns for participants (i.e., businessmen selecting more pro-business responses on EC target items and labour union participants selecting more pro-labour responses on EC target items). Participants were told that they would be completing an information test, and that they would probably not know the answers to all of the questions, but should make their best guess. Finally, it was emphasised to the participants that they should work as rapidly as possible when answering the questions.

Consistent with his hypothesis, Hammond found that when responding to the EC target items, participants did reveal a systematic guess pattern that was consistent with their group membership. In other words, those participants who worked for a labour

organisation selected responses that were favourably biased toward labour, while participants from two groups of businessmen selected responses that viewed labour groups less favourably. The reliability coefficient for the EC measure was .78. Notwithstanding these initially promising results, Hammond made a number of suggestions for further improving the performance of the EC technique.

Hammond believed that for the EC technique to function properly it would be necessary to disguise the purpose of the EC measure. In order to ensure that participants were unaware of the true purpose of the EC measure Hammond suggested that the EC measure be represented as an information or knowledge test. Additionally, he suggested that there should be at least two filler items for every target item in order to further disguise the EC measure. Hammond also recommended that each item on the EC measure have more than two possible responses. He suggested that if there were four possible responses then it might be possible for the EC technique to provide a measure of the intensity of the participant's attitude. Although not a recommendation per se, the amount of time participants spent answering the questions appears to have been an important concern for Hammond, as reflected in the EC instructions to participants to "work as rapidly as possible" when they answered the questions of the EC measure. Of note, other EC researchers (Antonak & Livneh, 1995a; Weschler, 1950b) have also identified a number of principles for developing EC instruments, although there was significant overlap with those put forward by Hammond (1948).

Surprisingly, given the initial promise of Hammond's original EC study, there appears to have been only sporadic interest in the EC technique over the years. Interestingly, only 14 EC articles investigating the indirect measurement of attitudes

have been published in journals in the last 60 years (Antonak, 1994; Antonak & Livneh, 1995b; Balma, Maloney & Lawshe, 1958; Bishop & Slevin, 2004; Blumfeld, 1966; Buttery, 1978; Clarke & Crewe, 2000; Hepperlen, Clay, Henly & Barké, 2002; Jacobs, 1972; Kubany, 1953; Livneh & Antonak, 1994; Weschler, 1950a; Weschler, 1950b; Wilde & Fortuin, 1969; Wilde & de Wit, 1970) and none of the articles were published in mainstream social psychology journals, but rather in more applied and/or specialised journals. All of these studies have replicated Hammond's basic finding that people's guess patterns systematically differ across groups in theoretically expected ways. This has been demonstrated across a host of attitude objects including labour and management (Brenberg, 1951; Weschler, 1950a; Weschler, 1950b,), socialised medicine (Kubany, 1953), doctors (Blumfeld, 1966), gifted children (Buttery, 1978), people with epilepsy (Antonak & Livneh, 1995b; Bishop & Slevin, 2004), Attention-Deficit/Hyperactivity Disorder (Hepperlen, Clay, Henly & Barké, 2002) and persons with disabilities (Clarke & Crewe, 2000; Livneh & Antonak, 1994). Additionally, the EC technique also demonstrated some promise as an alternate form of personality measurement (Wilde & Fortuin, 1969; Wilde & de Wit, 1970).

Thus the core findings of Hammond's original EC study have generally been replicated; however, this literature has made little progress beyond these original findings. For example, EC researchers have historically used reliability indices to assess whether people's guessing revealed a systematic response pattern. This is certainly a reasonable approach to determining if the people's responses to the EC target items were systematic. If people's responses were completely random, then the EC measure would have had low reliability. In fact, the majority of EC studies produced

very respectable levels of reliability ranging from .78 to .89 (Antonak, 1994; Balma, Maloney & Lawshe, 1958; Blumfeld, 1966; Buttery, 1978; Hammond 1948; Hepperlen, Clay, Henly & Barké, 2002; Jacobs, 1972; Kubany, 1953; Livneh & Antonak, 1994; Weschler, 1950a;). This suggests that in most of the studies the responses to the EC target items did reveal a systematic responses pattern; however, this still left open the question of the EC measure's dimensionality. Although EC researchers have implicitly assumed the measure was unidimensional (i.e., items have always been summed or averaged to a single score), this was not formally tested and reliability indices do not test dimensionality (Cortina, 1993; John & Benet-Martinez, 2000). Instead, dimensionality of a measure should be determined using factor analysis.

A second issue that has not been advanced by EC researchers is determining the attitudinal nature of the EC measure. Typically early EC researchers inferred the attitudinal nature of the EC measure based on the known groups paradigm. That is, attitudes were inferred by determining if people scored differently in their guess patterns based on their group affiliation, then the known groups should differ attitudinally. Although this is a reasonable strategy, interestingly, very few early EC researchers directly correlated the EC with direct measures of attitude and often these analyses provided ambiguous results. Consequently, these researchers had not formally gauged the precise magnitude of the overlap between the EC and direct measure on the construct of interest.

Another substantive issue that the Error Choice literature has never addressed is whether these measures actually combat social desirability. This was of course the primary rationale for researchers to construct EC measures, but so far no one has ever

demonstrated that the EC technique is less susceptible to socially desirable responding than direct measures of attitude. In fact, out a total of 14 empirical EC articles only one attempted to assess the measure's resistance to socially desirable responding. In this study Antonak and Livneh (1995b) developed an EC measure that examined attitudes towards people with epilepsy. In their study they included a measure of social desirability (i.e., the Marlowe-Crowne Social Desirability Scale), and both a direct (the Scale of Attitudes Toward Persons with Epilepsy) and an EC measure of attitude. They conducted a correlational analysis to examine the relationships between the three measures. The results of their study showed that there was no significant correlation between the social desirability scale and the EC measure. However, because neither the direct nor indirect measure revealed a significant correlation with the social desirability measure, it is not clear whether they used a sufficiently sensitive topic for socially desirable responding to even exist.

A final way in which the EC literature has failed to progress is that researchers have proposed a host of features that they think are important to the effective functioning of this technique (Antonak & Livneh, 1995a; Hammond, 1948; Weschler, 1950b). However, no one has ever systematically investigated whether these features are necessary. For example, Hammond's recommended a 2:1 filler/target item ratio. The majority of EC measure developed in subsequent studies; however, used a number of different item ratios (e.g., 2:1, 1:1 and 25:1) (Antonak, 1994; Balma, Maloney & Lawshe, 1958; Bishop & Slevin, 2004; Blumfeld, 1966; Hepperlen, Clay, Henly & Barké, 2002; Jacobs, 1972; Kubany, 1953; Livneh, 1994; Weschler, 1950b).

Finally, subsequent to the introduction of the EC technique a host of indirect measures of attitude have been developed (e.g., Implicit Association Test, Evaluative Priming, Personalized Implicit Association Test). However, even the more recent EC studies (Bishop & Slevin, 2004; Hepperlen, Clay, Henly, & Barké, 2002) have not examined the extent to which EC measure produce similar or very different results from these new indirect measures of attitude.

In summary, what has been done in the way of EC research is moderately encouraging, but so far it appears as though very little progress has been made since Hammond's original study.

Overview.

The goals of this programme of research are to provide a more thorough assessment of the EC technique. In order to accomplish this, I first conducted a factor analysis to examine whether responses to EC target items revealed a systematic response pattern, and determine the dimensionality of the response pattern. Second, I attempted to more definitively determine whether the guess pattern produced by the EC technique was reflective of attitudes. This was accomplished by directly correlating the EC measure with a direct measure of attitude. Third, I more thoroughly examined whether the EC technique was more resistant to socially desirable responding than a direct measure of attitude. In this programme of research, I attempted to directly manipulate social desirability in order to examine the EC measure's resistance to socially desirable responding. Fourth, I assessed whether some of the procedural features (i.e., time pressure and item ratio) that had been recommended by EC researchers were important for the success of the measure. Hammond suggested that

placing participants under time pressure when responding to EC items accentuates the effects of socially desirable responding. Additionally, Hammond recommended that there should be a 2:1 ratio between neutral and nonfactual items to disguise the true purpose of the EC measure. Finally, I examined how the EC technique related to other more contemporary indirect measures of attitude (i.e., Implicit Association Test and Personalized Implicit Association Test).

Chapter 2

Determining Whether the Systematic Response Pattern is Related to Attitudes and is Resistant to Socially Desirable Responding.

The goals of this study were fourfold. First, to extend Hammond's original findings by going beyond the use of reliability to examine the systematic response pattern to EC target items. In this study, rather than using the EC measure's reliability, which had been used in past EC research, an exploratory factor analysis was used to more formally investigate whether participants' responses revealed a systematic pattern. Second, the exploratory factor analysis was also used to determine the dimensionality underlying the participants' response pattern to the EC target items. Specifically, it is hypothesised that peoples' guess patterns will be systematic, such that, if they tend to guess in a positive direction on one item, they will tend to guess positively across other EC target items. Conversely, if they tend to guess in a negative direction on one EC target item, they will tend to guess negatively across other target items. Therefore, I expect to find a single dimension underlying peoples' responses to the EC target items.

Third, I will use an alternative approach to assess the attitudinal nature of the EC measure. Having established that there was an underlying pattern of guessing to the EC target items I wanted to use a different approach to test the attitudinal nature of the dimension different than the known groups paradigm. Specifically, it is hypothesised that the systematic guess pattern people engage in when responding to EC target items will positively correlate with a more direct measure of attitude.

Finally, I wanted to test the EC measure's resistance to socially desirable responding for participant's attitudes towards people who are overweight. This was accomplished by experimentally creating a condition whereby social desirability bias was high and a condition where social desirability was low, and then testing whether the EC scores differed across these conditions. It is predicted that the EC measure will be less susceptible to socially desirable responding than more direct measures of attitude.

Method

Participants

A total of 127 undergraduate students were recruited from either introductory psychology or other university courses at Queen's University. All participants received either partial course credit or five dollars for their participation. A total of 5 participants were dropped from the study because they did not complete all of the attitude measures. As a result, the data from 122 participants were included in the analyses.

Procedure

This study used a one-way design with 2 levels (social desirability: high vs. low). Each of the experimental sessions was comprised of 7–10 participants. Before the beginning of each session the participants were randomly assigned to either the high or low social desirability condition. Regardless of condition, all participants were given a letter of information, a consent form, and following the study, a debriefing form (See Appendix A for all ethics documents). After providing their consent to participate in the study they were given a booklet that contained both the EC and direct attitude measures. The EC measure was described by the research assistant as a general knowledge test about eating habits, weight and exercise. The coversheet of the EC

measure clearly identified the measure as a general knowledge test. The EC measure was always completed before the direct measure in order to disguise the purpose of the EC measure. Participants were instructed to select the correct answer for each question and, if they were unsure, then they should make their best guess. Finally, participants were told that the EC measure only took approximately five minutes to complete, but to do so they would have to answer the questions as quickly as possible.

In the low social desirability condition, participants were seated as far apart from each other as possible, which helped ensure no one other than the participant would be able to read their responses to the measures. Participants were then read a standardised script by the research assistant. The script for this condition described the booklet they were to complete as a general knowledge test about eating habits, weight and exercising (for the script to this condition and all conditions see Appendix B). The research assistant emphasised the anonymity of the research procedure by stating that the booklet assignment had been random, consequently there would be no way to match surveys with specific participants. Once the participants had completed the survey they were to ensure that there were no identifying marks on the booklet and were then directed to place the booklet face down in a box, which already contained a number of facedown booklets, as they left the room.

In the high social desirability condition participants were seated as close together as possible. This created the perception that other participants, as well as the research assistant, might be able to observe their responses to the measures. They were then read a standardised script, which was the same as the low social desirability condition with the following exceptions. Participants were instructed to write their

name and contact information (i.e., address, telephone number and email address) on the coversheet of the booklet, and to hand their completed booklets to the research assistant as they left the room at the end of the session. The participants were told that the research assistant would verify their contact information, which would enable them to ‘put a face to the name’ in the event that the participant needed to be contacted in the future. In order to ensure participant anonymity, this information was never coded or retained with the primary data file. Additionally, to further heighten the social desirability manipulation the research assistant went on to describe herself as someone who had struggled with being overweight and was very familiar with the stigmas encountered by people who are obese. As a result, this “attitudes towards people who are overweight” was a very important issue for her and why she became involved in this research project.

Measures

The Error Choice (EC) measure of attitude. The EC measure (See Appendix C for all study 1 measures) was developed on an attitude object that pre-testing had shown to be susceptible to socially desirable responding (i.e., attitudes towards people who are overweight).³ It was important to select an issue for whom many people would see a negative opinion as undesirable. If this was not the case, there would not be the need for people to misreport their attitudes because they could just report their true opinions.

³ Participants were provided with a number of issues (e.g., homosexuality, overweight, etc.), which they were asked to rate based on whether they believe that if someone was to express “X” negative attitude, to what extent did they think other people would see that attitude as undesirable. So for example, if someone were to express negative opinions about people who are overweight, to what extent do you think other people would view that as an undesirable or inappropriate opinion?

For this reason a set of issues was selected and a separate set of judges were asked to indicate if expressing negative opinions on those issues would be seen as undesirable by others. Our analysis indicated that attitudes towards people who are overweight was one of the highest scoring issues, consequently it was selected.

The EC measure was comprised of 23 filler (neutral) items and 10 target (attitudinal) items, which provided an approximately 2:1 item ratio as recommended by Hammond (1948). Additionally, each of the EC items had four possible responses from which the participant could select their answer. Consistent with EC theory the target items had clearly evaluative implications. The EC target items were structured such that the four possible responses reflected either an extremely negative, a slightly negative, slightly positive, or extremely positive attitude towards overweight people. For example, one of the target items asked participants what percentage of Canadians they thought were embarrassed by the weight of an obese family member, and then provided them with varying percentage levels of agreement (i.e., 16%, 38%, 61%, or 74%).

Total EC measure scores were then calculated for each participant by summing their responses across all EC items (negatively coded EC items were reverse coded). The EC scoring was structured such that lower scores represented more negative attitudes and higher scores represented more positive attitudes towards people who are overweight.

The EC filler items were objective knowledge questions that related to eating habits, weight and exercising. It was possible that participants might have known the correct answers to some of these questions. Indeed, Hammond (1948) suggested that being able to answer some of the questions correctly may assist in disguising the true purpose of the EC measure. Additionally, the objective knowledge questions had no

obvious evaluative component. For example, one of the filler items asked participants how many servings of meat or meat alternatives the Canadian Food Guide recommended an adult should consume daily. The filler items provided the participant with 4 possible response options, one of which was the correct answer; however, providing the correct response would not in and of itself reveal anything about their attitude. The participants' EC scores were calculated by summing their responses to the EC target items (negatively keyed items were reverse coded), whereby lower scores represented more negative attitudes and higher scores represented more positive attitudes towards people who are overweight ($M = 20.32$, $SD = 4.54$, Range = 22.00, Skewness = -.29, Kurtosis = -.26).

Direct measure of attitude. Participants' attitudes were also assessed using an 8 item semantic differential scale. Participants indicated their attitudes towards overweight people by selecting a rating for each pair of semantic differential scale items (e.g., bad/good, dislike/like, disgust/delight) based on how overweight people made them feel using a seven point rating scale (i.e., -3 to 3). The scale points were chosen to facilitate the association between the scale item and the associated numeric value (i.e., negative word item with corresponding negative numeric value). The participants' scores were rescaled from 1 to 7 and then summed ($M = 31.14$, $SD = 7.98$, Range = 38.00, Skewness = 0.64, Kurtosis = 0.64). Total direct attitude measure scores were then calculated for each participant, whereby lower scores represented more negative attitudes and higher scores represented more positive attitudes towards people who are overweight ($\alpha = .90$).

Self-monitoring scale. Finally, for exploratory purposes a measure of self-monitoring was included in the study (Snyder, 1974). Self-monitoring is defined as a person's tendency to monitor and control their behaviour in response to self-presentation concerns. Because self-monitoring can serve a number of purposes, like concealing inappropriate emotions and behaviours, one might postulate that high self-monitors may well be more prone to socially desirable responding than low self-monitors. The Self-Monitoring (SM) scale was comprised of 25 true/false items. Participant SM scores were computed by summing items (negatively keyed items were reverse coded) such that higher scores reflected higher levels of self-monitoring ($M = 10.39$, $SD = 3.76$, Range = 17.00, Skewness = $-.13$, Kurtosis = $-.48$). The Cronbach alpha coefficient for the self-monitoring scale used in the present study was $\alpha = .75$.

Results

Systematic Response Pattern

*Exploratory factor analysis*⁴. If participants' answers to the EC target items reveal a systematic pattern of responding, then a factor analysis should reveal a single latent construct. Although not examined directly, one might infer from the generally positive results and the good reliability of measures developed and used in past EC research that a single factor might account for EC target item data. However, previous EC research had not examined the underlying structure of the EC target items. One approach to examining the underlying structure is to conduct an exploratory factor

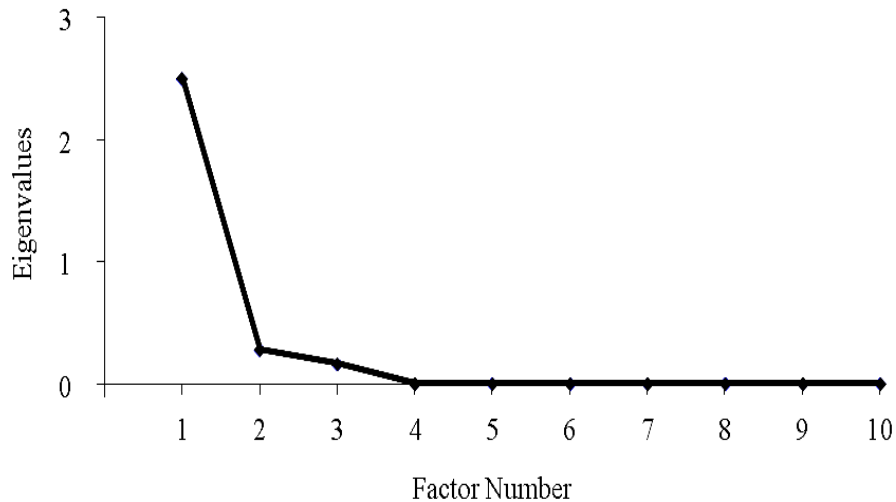
⁴ In study 1, and all subsequent studies, the dimensionality of the EC target items was also examined using Principal Components Analysis. Additionally, the parallel analysis and scree plot were calculated with eigenvalues from the original correlation matrix. These analyses produced very similar results to what is presented here in all three studies.

analysis, which fits a common factor model to the data. The ideal model is one that balances parsimony with explanatory power. Although there are a number of different methods for determining the appropriate number of factors none of them are infallible. Consequently, three separate statistical methods were used to identify the number of factors.

The first method was the scree test of the eigenvalues from the reduced correlation matrix (i.e., the amount of variance accounted for by each of the factors). This procedure involves constructing a graph from eigenvalues that are calculated from the correlation matrix with communalities in the diagonal (Tucker, Koopman, & Linn, 1969; Fabrigar, Wegener, MacCallum, & Strahan, 1999). These values are then plotted in descending order. The plot is then examined to determine the number of factors that precedes the last ‘substantial’ drop. The challenge that one faces when interpreting a scree test is that there are no firm criteria established for determining what constitutes a ‘substantial’ drop. As reflected in the scree plot for the EC data (see Figure 2) it does appear as though there is one dominant factor; however, without firm criteria for what constitutes a ‘substantial’ drop one cannot unequivocally rule out a possible 3 factor interpretation of the scree plot. Although the drop for the second and third factor is fairly modest they do stand out from the remainder of the factors depicted in the graph. Consequently, the results from this technique suggest either a 1 or 3 factor solution.

A second method used to determine the number of factors was parallel analysis (Horn, 1965). This method is based on the premise that the eigenvalues calculated from the reduced correlation matrix should be substantially larger than the eigenvalues from a reduced matrix based on completely random data (given the same sample size and

Figure 2. Scree plot of eigenvalues from the Reduced Correlation Matrix



number of variables). The number of factors is determined by comparing both sets of eigenvalues, whereby the number of factors corresponds to the number of eigenvalues from the actual data set that are greater than eigenvalues expected from random sets of data. In this analysis, the eigenvalues expected from random data were calculated using an SPSS macro developed by O'Connor (2000). The parallel analysis suggests that a 1 factor solution best fits the data (see Table 1) because only the first eigenvalue from the actual data set is greater than those produced by random data.

Another approach to assessing factor structure is to examine model fit (Browne & Cudeck, 1989) which uses maximum likelihood (ML) for factor estimation and provides goodness-of-fit information that can be used to establish the correct number of factors. One of these approaches for examining model fit is Root Mean Square Error of Approximation (RMSEA), which calculates the discrepancy between the model and the

Table 1. *Parallel Analysis – Eigenvalues from the Reduced Correlation Matrix.*

Factor number	Eigenvalues from actual data set	Eigenvalues from random data set
1	2.499	.458
2	.286	.299
3	.169	.178
4	.008	.078
5	-.077	-.009
6	-.135	-.091
7	-.173	-.183
8	-.247	-.278
9	-.325	-.347

data per degree of freedom for the model. The goal of this approach is to select a model that fits substantially better than one with one fewer factor, while fitting as well as, or nearly as well as a factor model with one additional factor. Ideally, the preferred model should also have at least acceptable fit (i.e., $RMSEA < .08$) and preferably good fit (i.e., $RMSEA < .05$). An additional feature of RMSEA is that it also provides a confidence interval, which allows the researcher to assess the precision of the sample value of RMSEA. Thus, the confidence intervals provide further assistance in evaluating model fit (MacCallum, Brown, & Sugawara, 1996).

In RMSEA, each of the factor models of interest to the researcher are examined, starting with the simplest and continuing until all of the models have been examined. In the present study, 3 separate factor models (i.e., 1, 2, and 3) were examined for the EC measure (see Table 2). After examining the RMSEA results for the 1 factor model (see Table 2, row 1) it appears as though 1 factor fits the model quite well. The RMSEA value for this model indicates good fit (.032) and the confidence interval indicates that even at the upper boundary is below the .05 heuristic for good model fit. When the 2 factor model is examined (see Table 2, row 2), there is a slight improvement in the RMSEA value (< .01) and a corresponding increase in the confidence interval. This pattern of result is repeated when an additional factor is added to the model. In the 3 factor model (see Table 2, row 3), there is a slight improvement in the RMSEA value (< .01) but the confidence interval is broader than the proceeding factor models. The 2 and 3 factor models provided only slightly lower RMSEA point estimates and there was also a corresponding increase in the size of their respective confidence intervals, which suggests that the 1 factor model was the most parsimonious and unambiguous.

Table 2. RMSEA Goodness-of-Model Fit

Number of Factors	RMSEA	Confidence Interval (90%)
1	0.032	0.012-0.046
2	0.024	0.000-0.040
3	0.016	0.000-0.036

In order to make a final determination on the most appropriate number of factors, it is necessary to overlay the results of all 3 factor extraction procedures. The results of the scree test support 1 or 3 factor solution. The parallel analysis is less ambiguous and indicates that 1 factor model is most appropriate. Finally, the RMSEA also suggested that the 1 was most likely the correct number of factors. It appears then, that based on the preponderance of the evidence a 1 factor model best accounts for the data. That being said, there is one final criterion for estimating the number of factors that is at least as important as the 3 procedures already used is factor interpretability. Although all 3 procedures seem to favour the 1 factor model, nonetheless to be further confident the interpretability of the 1, 2 and 3 factor models were examined.

The factor interpretability is determined by examining the factor loadings of the EC items for a specified model and, where necessary, examining the individual EC items to ensure that the factor(s) make sense conceptually. An exploratory factor analysis was carried out on the EC measure (all 10 EC target items) using a maximum likelihood factor analysis (with a direct quartimin rotation for the 2 and 3 factor models). In the 1 factor model, all of the EC target items did load on a single factor; however, 6 of the EC target items were strongly influenced by the underlying factor demonstrating factor loadings greater than .5, while the factor loadings for the remaining 4 EC target items hovered around the .4 cutoff heuristic (see Table 3 for the factor matrix for the 1 factor model).

The interpretation of the 2 factor model for the 10 EC target items was very difficult for several reasons. First, all 6 EC target items that loaded stronger in the 1 factor model retained substantial loadings in the 2 factor model; however, factor 2

separates out into 3 items. The structure of these EC target items did not make sense conceptually (e.g., personal discomfort watching obese people on television, and lack of self-control as a causal factor for being overweight and being embarrassed by an obese

Table 3. Factor Matrix for the 1 Factor Model.

EC Target Items	Factor Loadings
EC 9	.72
EC 7	.67
EC 8	.63
EC 6	.57
EC 5	.54
EC 10	.52
EC 1	.38
EC 4	.37
EC 2	.34
EC 3	.31

Note: Extraction Method: Maximum Likelihood.

family member's weight). Other than the primary theme of obesity it is difficult to determine what other underlying construct these items have in common. Additionally,

it was very difficult to interpret this second factor because 2 of the EC target items (i.e., 1 and 4) are positive and the third item (i.e., 2) is negative. Yet the keying of those EC target items was not opposite; furthermore, the loading of EC item 4 was somewhat marginal and it too loaded somewhat on both factors (see Table 4 for the factor pattern matrix for the 2 factor model). Given its more complex nature and lack of interpretability, the 2 factor model is not strongly supported.

Table 4. Pattern Matrix for the 2 Factor Model.

EC Target Items	Factor 1 Loadings	Factor 2 Loadings
EC 9	.68	.14
EC 7	.68	.05
EC 8	.61	-.07
EC 6	.58	.04
EC 5	.54	-.05
EC 10	.52	.16
EC 2	.47	-.40
EC 3	.27	.07
EC 1	.20	.57
EC 4	.24	.39

Note: Extraction Method: Maximum Likelihood.
Rotation Method: Oblimin with Kaiser Normalization.

The 3 factor model of the 10 EC target items was even more difficult to interpret than the 2 factor model (see Table 5 for the factor pattern matrix for the 3 factor model). The 6 EC target items that had moderate to strong factor loadings in the 1 and 2 factor models no longer formed a single factor. In fact the first factor for this model was comprised on only 1 EC target item and it was a Heywood case (EC target item 5 produce a communality > 1). This can indicate that either the specified common factor model was not appropriate (i.e., too many or too few factors). Given that the 1 factor model for the 10 EC target items provided good communalities and neither the 1 nor 2 factor models produced any Heywood cases, it seems as though the problem may be that too many factors have been specified in this model.

The second factor in this model was comprised of 6 EC target items, similar to the 1 and 2 factor models; however, item 5 (i.e., personal comfort level being in the company of obese family members) had been dropped and item 2 (e.g., lack of self-control as a causal factor for being overweight) had been added. This target item again moderately cross-loaded on 2 different factors (i.e., 2 and 3). The third factor was comprised on 3 EC target items, and mirrored all of the target items and their respective concerns described for the second factor in the 2 factor model. That is to say, EC target item 2 cross-loaded on both factors, EC target item 3 did not load on any of the factors and EC item 5 negatively loaded on the third factor. Finally, the factor structure of the EC target items did not make sense conceptually. Given the more complex nature and difficulty interpreting the factors contained in the 3 factor model the 1 factor solution provides the most parsimonious and interpretable solution for the data.

Table 5. Factor Pattern Matrix for the 3 Factor Model.

EC Target Items	Factor 1 Loadings	Factor 2 Loadings	Factor 3 Loadings
EC 5	1.01	-.06	.08
EC 9	-.08	.778	.16
EC 8	-.12	.69	.06
EC 7	.22	.54	.02
EC 10	.03	.52	-.05
EC 6	.22	.47	-.10
EC 3	.06	.24	.05
EC 1	.07	.21	.52
EC 2	.08	.38	-.43
EC 4	.14	.18	.37

Note: Extraction Method: Maximum Likelihood.
Rotation Method: Oblimin with Kaiser Normalization.

EC measure reliability. The next thing that would be expected if the EC measure is functioning properly is that it would be fairly reliable. The 10 item EC measure did demonstrate good levels of reliability ($\alpha = 0.76$). However, the initial exploratory factor analysis revealed that 2 of the 10 EC target items (i.e., EC items 2

and 3) had marginal factor loadings and were 2 of the items that split off in the 2 and 3 factor models. This suggests the EC measure might perform better if these target items were removed entirely (see Table 6). Once these 2 items were removed the reliability of the measure improved ($\alpha = 0.78$). Given this improvement in the reliability once the poorer performing items were removed this raised the question about the effect of retaining only the highest loading target items in the EC measure. When only the 6 highest loading EC target items were analysed the reliability of the measure weakened somewhat ($\alpha = 0.74$). As a result, the 8 item EC measure was used in the remainder of the analyses for this study.

Table 6. Factor Matrix 1 Factor Model (8 EC Items).

EC Item	Factor Loadings
EC 9	.74
EC 7	.67
EC 5	.59
EC 8	.57
EC 6	.55
EC 10	.51
EC 4	.48
EC 1	.40

Note: Extraction Method: Maximum Likelihood.

Interpreting the correlations between the measures.

When the target items were created, it was believed that the response pattern should be a somewhat evaluative response pattern, because all the EC target items were designed and selected based on the premise that answers to these items had an evaluative content. As a result, the most obvious interpretation of the EC target items would be their evaluative component (see Appendix D for complete correlation tables of all dependent variables). Nonetheless, it would be valuable to more directly confirm this by examining the correlation between the EC and direct measure of attitude. That is to say, given that a participant's responses to the EC items does reveal a systematic guess pattern, are participants using their attitudes to guide their responses? In fact, the correlation between measures was significant, $r(120) = .30, p < .01$. So, it appears that whatever is governing the response pattern it does have some evaluative content.

Evaluating the effect of socially desirable responding.

The next theoretical question is whether the EC and direct attitude measures were affected by socially desirable responding. To test this, it was first necessary to determine whether the social desirability manipulation was successful. Given that direct measures of attitude tend to be vulnerable to socially desirable responding, its successful manipulation would be reflected in participants' differentially responding across conditions (social desirability: low vs. high). As hypothesized, the analysis revealed that participant's scores on the direct measure were significantly different across social desirability conditions (Low: $M = 29.48, SD = 8.36$; High: $M = 32.39, SD = 7.49$); $t(119) = 2.01, p = .05$ and in the correct direction. That is to say, participants' reported more favourable attitudes toward people who are overweight in

the high social desirability condition. Thus, the social desirability manipulation did successfully create differential responses by condition.

These results now permit the EC measure's resistance to socially desirable responding to be determined. If the EC mean is resistant to socially desirable responding then the participant's scores for the EC measure should not vary across conditions. This is in fact what was found. Participants' EC scores in the low social desirability condition ($M = 20.06, SD = 4.59$) and the high social desirability condition ($M = 20.52, SD = 4.53$) were not significantly different, $t(120) = .56, p = .73$. Given the comparatively good reliability of the 8 item EC measure ($\alpha = 0.78$) these null results are not simply due to random error (low reliability of the measure) and represent encouraging initial support for the EC measure's resistance to social desirability⁵. These results are also significant in that they represent one of the very few times that an EC measure's resistance to social desirability has been directly tested.

An additional question that might arise from these analyses relates to whether the correlations of participants' scores on the direct measure vary between social desirability conditions. If participants are induced in the high social desirability condition to respond in a desirable ways to the direct measure, but not the EC measure,

⁵ In the present study the effects of socially desirable responding were examined by looking at its effects on each attitude measure separately. Another approach would be to examine the effects of social desirability on the 2 attitude measures together. In order to accomplish this a mixed-model analysis of variance was conducted with the type of measure (i.e., EC and the direct measure) as the within-subjects factor and Social Desirability (i.e., High and Low) as the between-subjects factor. The results suggested that the effects of social desirability were weaker on the EC measure than on the direct measure of attitude ($F(1,121) = 2.93, p = .11$). Note, before conducting this analysis both the EC and direct measure were re-scaled to a common metric, where scores ranged from 0 to 1.

then the correlations between these measures would be weaker in this condition and stronger in the low social desirability condition, when presumably both measures are more accurate reflections of the person's attitudes. That being said, this may not be the case because if socially desirable responding just adds a constant to the direct measure (e.g., everyone is just responding 2 scale points higher) as this would not alter the associations.

The only way social desirability would be detrimental to associations between attitude measures is if there was substantial heterogeneity in the amount that participants were altering their responses in the high social desirability condition. The correlations between the EC and direct measure were slightly higher in the low social desirability condition ($r(52) = .41, p < .01$) than in the high social desirability condition ($r(68) = .20, p = .11$). It should be noted, however, that the difference between the correlations was not statistically significant, $z = -1.05, p = .15$. At least in part this homogeneity in participant responses may have been a result of the social desirability manipulation, which was designed to be quite strong and to create a situation where participants would not be highly variable in their responding.

Self-monitoring scale. As discussed earlier, because self-monitoring can serve a number of purposes, like concealing inappropriate emotions and behaviours, one might postulate that high self-monitors may be more prone to socially desirable responding than low self-monitors. If participants who are high self-monitors engage in more socially desirable responding, than those who are low self-monitors, it might be expected that there would be a positive correlation between their attitude reports and

self-monitoring scores. The analysis revealed that there were no significant correlations between the self-monitoring (SM) scale and direct measure of attitude in either the high ($r(68) = -.09, p = .45$) or low ($r(52) = .11, p = .45$) social desirability conditions. As well, there were no significant correlations between the SM scale and the EC measure in either social desirability condition (high: $r(68) = -.04, p = .76$; low: $r(52) = .16, p = .25$). The nature of the socially desirable responding in this study was probably issue specific, while the self-monitoring scale may capture more general forms of socially desirable responding. This difference in specificity might account for the lack of strong correlations between the self-monitoring scale and the direct measure of attitude.

Although the correlations between the measures were not significant, it is possible that self-monitoring may moderate the relationship between the social desirability manipulation, and the EC and/or direct measure of attitude. To examine this possibility two General Linear Model (GLM) analyses were carried out. Before performing these analyses participant's SM scores were centred. The first GLM was specified with a dummy variable representing the social desirability manipulation, a continuous variable representing self-monitoring, and an interaction term (self-monitoring x social desirability) as predictor variables for the direct measure of attitude. The first GLM analysis revealed that there was no significant interaction ($F(1,120) = 1.17, p = .28$) between the SM scale and social desirability manipulation. This analysis was repeated with EC as the dependent variable, and again there was no significant interaction ($F(1,120) = 1.09, p = .30$) between SM scale with the social desirability manipulation. These analyses suggest that the self-monitoring scale was not

functioning as a moderator between the social desirability manipulation and the attitude measures.

Summary of Findings.

The purpose of Study 1 was fourfold. First, to directly examine the systematic guess pattern for EC target items using factor analysis. Second, determine the dimensionality of the response pattern. Third, to determine whether the systematic guess pattern people engaged in when responding to EC target items would be positively correlated with a more direct measure of attitude. Fourth, I tested whether the EC measure was resistant to socially desirable responding. The results from this study were very encouraging, in that there was support for all four of the stated goals.

That is to say, participants' responses to the EC target items did reveal a single underlying dimension. As well, the systematic guess pattern was, at least in part, attitudinal. Finally, the EC measure was resistant to socially desirable responding, as previous EC researchers had suggested but never tested. Exploratory analyses revealed that the relationship between socially desirable responding and the attitude measures was not moderated by self-monitoring. These encouraging results suggest that it may be the case that the broad condemnation Kidder and Campbell (1970) made regarding indirect measures of attitude was premature. Overall, this study has demonstrated that the Error Choice technique can be very reliable, resistant to socially desirable responding and provided some evidence that it is somewhat attitudinal in nature.

Chapter 3

Providing an Initial Examination of an Optimal Feature of the Error Choice (EC) Technique and Comparing the EC Technique with a Current Indirect Measure of Attitude.

The purpose of Study 2 was fourfold. First, to provide a replication of the systematic response pattern and dimensionality of the EC target items. Second, test the effects of social desirability responding on the EC measure and a direct self-report measure of attitude found in Study 1. Next, to examine one of the conditions presumed to be optimal for the functioning of an EC measure (i.e., time restriction). EC researchers had suggested that restricting the amount of time participants had to the complete an EC measure was a necessary condition for the EC technique function properly (Antonak & Livneh, 1995a; Hammond, 1948). It was their contention that the restriction of time would ensure that participants did not over-think their responses and thus rely on their attitudes when guessing. Fourth, to compare the EC technique to a more current and widely used indirect measure of attitude (i.e., Implicit Association Test, IAT). The IAT (Greenwald, McGhee, & Schwartz, 1998) was selected because it is probably the best known and researched indirect measure of attitude.

Method

Participants

A total of 243 undergraduate students were recruited from either introductory psychology or other university courses at Queen's University. All participants received either partial course credit or five dollars for their participation. A total of 9 participants

were dropped from the study because they did not complete all of the attitude measures or the response latencies had an excessive error rate (i.e., 30% or greater as recommended in Greenwald, Nosek, & Banaji, 2003). As a result, the data from 234 participants were included in the analyses.

Procedure

The design of this study was a 2 (social desirability: high social desirability versus low social desirability) x 2 (time: restricted versus unrestricted) x 2 (order of indirect measures: IAT/EC versus EC/IAT). The sessions were conducted in groups of up to 5 participants, and before each session of the study, the groups were randomly assigned to one of the social desirability conditions. The EC/IAT order and time restriction conditions were assigned at the individual level within sessions. By and large the procedures for this study were the same as Study 1; however, because the optimal method for IAT delivery is computer based, digital versions of all measures used in the study were created. Following their consent to participate in the study, participants were told by the research assistant they would be completing several computer-based tasks including a categorisation task and a general knowledge test about eating habits, weight and exercise. The EC and IAT measures were always completed before the direct measure of attitude in order to disguise the purpose of the other measures.

Social desirability. Other than a number of minor procedural changes to accommodate the use of computers, rather than written booklets, the procedure followed the same manipulation of social desirability as used in Study 1.

Time restriction. In the time condition, participants were either given an unlimited amount of time (Unrestricted), or they were told that they would only have 3

minutes and 30 seconds to complete all 31 questions (Time Restricted) of the EC measure. The latter condition placed participants under time pressure, but still allowed them sufficient time to complete the entire EC measure. The parameters for the time condition were determined by pre-testing the EC measure and determining the average time participants required to complete the measure. The amount of time that participants were told they had to complete the EC measure during the study was set slightly below the pre-tested average. Because the IAT has a very specific protocol, time restriction was not placed on this measure in either of the conditions.

Presentation of indirect measures of attitude. Both indirect and direct measures of attitude were computer-based to facilitate the use of the IAT and to capture response latency data for all participants (See Appendix E for all study 2 measures). In the indirect measures of attitude condition, the EC and IAT presentation were counterbalanced. Regardless of the order of presentation for the indirect measures, they were always given before the direct measure of attitude. This helped ensure that participants were not aware that their attitudes towards overweight people were being assessed, which if known might influence their performance on the EC and IAT measures.

Measures

The Error-Choice (EC) Measure. The EC measure used in this study was the same as Study 1 except that the 2 poorer performing EC target items (i.e., 2 and 3) were removed. As a result, the EC measure consisted of 23 neutral (filler) and 8 target items ($M = 20.29$, $SD = 4.48$, $Range = 22.00$, $Skewness = 0.22$, $Kurtosis = -.41$). Participants were instructed to select the correct answer for each question and, if they were unsure,

then they should make their best guess. In the computerised version of the EC measure each of the items was presented individually on a computer monitor with its 4 response options. After the participant selected a response the next EC target item appeared. This process was repeated until all of the EC items had been presented. The computer program did not allow participants to go back to previously answered questions and review or change their responses.

Direct measure of attitude. Participants' attitudes were also assessed using the same 8 item semantic differential scale as Study 1 ($M = 28.03$, $SD = 6.29$, Range = 38.00, Skewness = $-.17$, Kurtosis = 1.10). The Cronbach alpha coefficient for the direct measure used in the present study was $\alpha = 0.90$.

The Implicit Association Test (IAT). The IAT was chosen because it is currently the most widely used indirect measure of attitude and has generated a substantial literature since it was first introduced over 10 years ago by Greenwald et al. (1998). The IAT is a response latency technique which is based on the premise that people's responses will be faster and more accurate when performing mental tasks that are easy compared to those that are more difficult. More specifically, participants completing an IAT were presented with words from 4 distinct categories, which were paired to form either an easy mental task (i.e., congruent trials) or a more difficult mental task (i.e., incongruent trials). Two of the IAT categories tap into the dimension that represents the attitude object of interest (i.e., attitudes towards people who are overweight) and the other 2 categories represent an evaluative dimension (i.e., good and bad).

Each of the categories for the IAT used in this study was comprised of a set of 12 representative words Greenwald et al. (1998) recommended that the number of IAT

items should be between 5 and 25. The sets of words were determined by providing a group of raters (n=6) a master list that contained 25 representative words for each of the categories (see Appendix E for a complete list of IAT stimuli). The raters were instructed to identify those words from the respective master lists that they believed most strongly represented that category. The category sets were chosen from those words where there was unanimous agreement between all raters. The final set of items for the 'overweight' category included words such as chubby, hefty, obese and pudgy, while the 'slim' set of items included words such as buff, skinny, slender and trim. This process was repeated for the evaluative dimension categories 'good' (e.g., sunshine, rainbow, pleasant, etc) and 'bad' (e.g., nasty, spoiled, sinful, etc.).

The IAT protocol used in this study was based on the procedure outlined in Greenwald, et al. (2003), which required participants to complete 3 practice and 4 experimental trials. In each of the experimental trials all of the words from the 4 categories were presented randomly (without replacement). For the experimental trials 2 pairs of category labels consisting of an attitudinal and evaluative dimension appeared at the top of the computer monitor. In the first 2 experimental trials the paired category labels were 'Overweight/Bad' and 'Slim/Good'. Participants were asked to correctly identify, as quickly and accurately as possible, which category the presented word belonged to by hitting the appropriate computer response key. Although there were 4 categories there were only 2 assigned response keys for the participants to correctly categorise each of the words. That is to say, participants had to hit the same response key if the word belonged to either of the 2 category labels that constituted that pair of

labels. For example, in the first set of experimental trials if the random word sequence was 'obese' and 'nasty' the participant would hit the same response key for both words.

The first 2 experimental trials were considered the congruent trials because the concept and evaluative labels that people might be expected to normally associated with each other are paired together (i.e., Overweight/Bad and Slim/Good). Consequently, if people do associate 'Bad' with 'Overweight' and these 2 categories share the same response key, then the response latency between these items should be quite small. In the last 2 experimental trials the evaluative dimension labels were reversed so the new paired categories were 'Overweight/Good' and 'Slim/Bad' with each pair again sharing the same response key. If a person does not normally associate 'good' with 'overweight' (i.e., incongruent) then the response latency between words associated with these categories will be larger than those in the congruent trials. The size of the difference between the two sets of trials estimates the degree to which the participant's evaluations differ, whereby the larger the difference between the congruent and incongruent trials the more negative the association with 'Overweight'.

IAT data preparation and the calculation of the difference scores between the congruent and incongruent trials were conducted in accordance with the improved IAT scoring procedure as recommended by Greenwald et al. (2003). In brief, all subjects with greater than 10% response rates of less than 300 milliseconds and all trials with reaction times greater than 10,000 milliseconds were deleted. It should be noted that no error response penalty was applied for participants' incorrect response, since the time required to produce the second (i.e., correct) response was included in the response latency of the MediaLab program. To calculate the difference scores between the test

stages (Stages 3 and 6, and Stages 4 and 7) the mean difference scores were divided by their respective inclusive standard deviation, which resulted in an equal-weight average of the two resulting ratios (see Greenwald et al., 2003 for a complete description all IAT scoring procedures and recommendations).

Self-monitoring scale. For exploratory purposes the self-monitoring scale (Snyder, 1974) used in Study 1 (Cronbach $\alpha = .73$) was also included in the present study ($M = 10.10$, $SD = 3.50$, Range = 18.00, Skewness = $-.28$, Kurtosis = $-.30$).

The Balanced Inventory of Desirable Responding – Version 6 (BIDR). For exploratory purposes, a measure of impression management (IM) was also included in the study (Paulhus, 1991). Impression management is defined as a person's tendency to alter their behaviours or opinions in order to create a positive image on others. Because impression management can serve to conceal inappropriate behaviours and opinions one might postulate that those high in impression management may be particularly prone to socially desirable responding. The IM scale contained in the BIDR is comprised of 20 statements (an equal number of positively and negatively coded items) that require the respondent to rate their level of agreement for each statement using a 7 point rating scale (1 – True to 7 – Not True). Once the IM scale had been completed, negatively keyed items were reverse coded and 1 point was awarded for each response of 6 or 7. The participant's IM score was then summed, whereby higher scores reflected higher levels of impression management ($M = 5.98$, $SD = 3.07$, Range = 16.00, Skewness = 0.52, Kurtosis = $-.33$). The Cronbach alpha coefficient for the IM scale used in the present study was $\alpha = 0.63$.

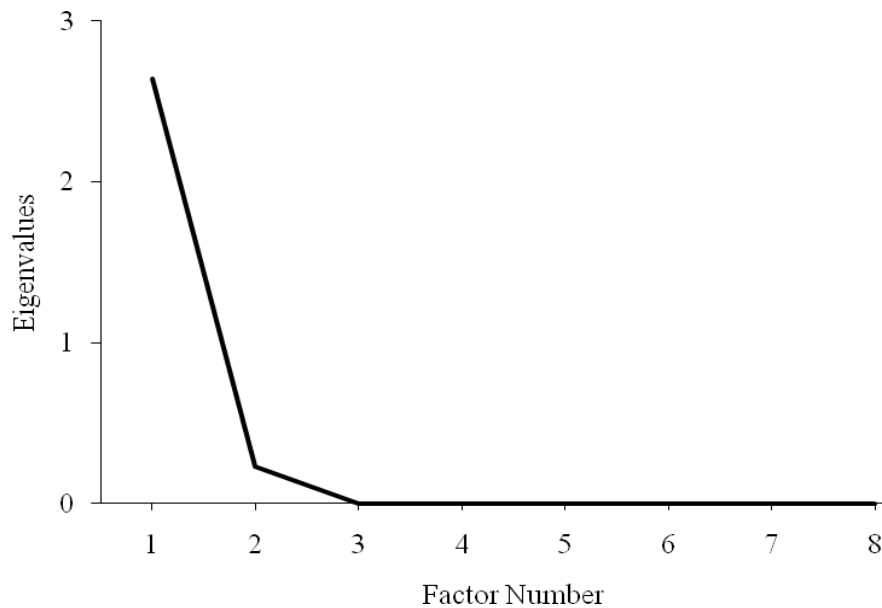
Results

Systematic Guess Pattern

Exploratory factor analysis. The initial purpose of this study was to replicate the findings of Study 1. Therefore, if participants' answers to the EC target items did reveal a systematic pattern of responding, then a factor analysis should reveal a single latent construct. Consistent with the exploratory factor analysis methodology used in Study 1, the same 3 separate statistical methods were used to identify the number of factors structure (i.e., Scree Plot, Parallel Analysis and Root Mean Square Error of Approximation).

The plot of the scree test of the eigenvalues from the reduced correlation matrix revealed that there was one dominant factor; however, without firm criteria for what constitutes a 'substantial' drop one cannot unequivocally rule out a possible 2 factor interpretation of the scree plot (see Figure 3). Although the drop for the second factor is

Figure 3. Scree Plot of Eigenvalues from the Reduced Correlation Matrix



fairly modest it does stand out from the remainder of the factors depicted in the graph.

Next, the parallel analysis suggests that a 1 factor solution best represents the data (see Table 7) because only the first eigenvalue from the actual data set is greater than those produced by random data.

Table 7. *Parallel Analysis – Eigenvalues from the Reduced Correlation Matrix.*

Factor number	Eigenvalues from actual data set	Eigenvalues from random data
1	2.634	.409
2	.243	.278
3	.037	.124
4	.020	.051
5	-.035	-.019
6	-.132	-.080
7	-.208	-.147
8	-.217	-.147

In the final approach Root Mean Square Error of Approximation (RMSEA) was used to examine model fit. As in Study 1, 3 separate factor models were examined (i.e., 1, 2, and 3) for the EC measure (see Table 8). After examining the RMSEA results

for the 1 factor model (see Table 8, row 1) it appears as though 1 factor model provides acceptable fit (.063). The RMSEA value for the 2 factor model; however, did provide a substantial improvement in model fit (.000), and the confidence interval indicated

Table 8. *RMSEA Goodness-of-Model Fit.*

Number of Factors	RMSEA	Confidence Interval (90%)
1	0.063	0.032-0.093
2	0.000	0.000-0.045
3	0.000	0.000-0.061

that even at the upper boundary it was below the .05 heuristic for good model. The 3 factor model provided no improvement in RMESA value and there was a notable increase in the confidence interval. Although the 1 factor model did provide acceptable fit, the 2 factor model provided a lower RMSEA point estimate and a corresponding reduction in the size of the confidence interval. This suggests that either the 1 or the 2 factor model might best represent the data.

The preponderance of evidence from these 3 statistical approaches suggests that either a 1 or 2 factor model may be appropriate for the data. In order to make a final determination in identifying the correct number of factors an exploratory factor analysis was carried out on the 8 EC target items using the same factor analytic procedure as the first study. When factor interpretability is examined there does appear to be additional

support for the 1 factor model. In the 1 factor model all 8 EC target items loaded on a single factor demonstrating moderately strong factor loadings (see Table 9).

Table 9. *Factor Matrix for the 1 Factor Model.*

EC Target Items	Factor Loadings
EC 4	.67
EC 7	.64
EC 5	.60
EC 6	.59
EC 2	.59
EC 3	.57
EC 1	.46
EC 8	.43

Note: Extraction Method: Maximum Likelihood.

The interpretation of the 2 factor model for the 8 EC target items was more difficult (see Table 10 for the factor pattern matrix for the 2 factor model). First, the 5 EC target items that loaded stronger in the 1 factor model remain substantial loadings in the 2 factor model and they were joined by one of the lower loading EC target items (1); however, factor 2 separated out 2 items (6 and 8). The structure and content of these EC target items did not seem conceptually different than the EC target item that made

up the first factor (e.g., percentage of people report being disgusted when seeing an obese person eating a fast-food restaurant, and percent of people repulsed by the idea of a romantic involvement with an obese person). In fact, there are a number of EC

Table 10. *Pattern Matrix for the 2 Factor Model.*

EC Target Items	Factor 1 Loadings	Factor 2 Loadings
EC 3	.74	.15
EC 2	.67	.06
EC 4	.56	-.15
EC 5	.49	-.16
EC 7	.43	-.27
EC 1	.38	-.12
EC 6	.01	-.80
EC 8	.05	-.50

Note: Extraction Method: Maximum Likelihood.
Rotation Method: Oblimin with Kaiser Normalization.

target items in factor 1 of the 2 factor model with similar wording, structure and affect (e.g., percentage of students who felt “repulsed” by overweight people). So, other than the primary theme of obesity it is difficult to determine what other underlying construct these items might share. Additionally, both factors were moderately correlated with

each other ($r = .58$) indicating that there was a large overlap in the underlying construct of both factors in this model. Given the more complex nature, the strong correlation between factors 1 and 2, and the difficulty interpreting the second factor contained in the 2 factor model, the 1 factor model was selected as the most parsimonious and interpretable.

Given that a 1 factor solution best represents the data, the next important question is what, if any, effect might limiting the amount of time available for participants to complete the EC measure have on the factor loadings. Hammond (1948) suggested that limiting the amount of time participants had to complete the EC questions should improve the performance of the EC measure. If this is the case, then it might be expected that there would be an increase in the factor loadings of the EC target items when participants were given a limited amount of time to complete the EC measure. However, when a factor analyses was carried out in the 'time' condition, the results revealed that the factor loadings of the EC target items in the 'restricted' time condition showed a slight decrease compared to the 'unrestricted' time condition (see Table 11). Additionally, there was a slight decrease in the reliability of the EC measure between the 2 conditions (Unrestricted Time, $\alpha = 0.82$; Restricted Time, $\alpha = 0.74$). Therefore, contrary to what Hammond had suggested, time restriction did not improve the performance of the EC measure. Indeed, it appears the correlations between the EC and direct measure in the 'restricted time' condition were trending slightly in the positive direction. It should be noted that the 'unrestricted time' condition in this study was a replication of the Study 1 procedure; consequently it provided further support of a systematic response pattern of guessing to the EC items and a single factor solution.

Table 11. *Factor Matrix for the Time Condition.*

EC Target Items	“Unrestricted” Time Loadings	“Restricted” Time Loadings
EC 7	.72	.64
EC 5	.66	.63
EC 6	.63	.60
EC 4	.62	.57
EC 3	.61	.57
EC 8	.58	.52
EC 1	.57	.44
EC 2	.45	.35

Interpreting the correlations between the measures.

Given that participants’ responses to the EC target items once again revealed a systematic guess pattern, the next question is to what extent were participants using their attitudes to guide their responses? If it is the case that participants’ scores across the EC target items are being influenced by their attitudes, then their EC scores should be correlated with the direct measure. In fact, the correlation between these measures was significant ($r(232) = .21, p < .03$). Although the correlation is low, it does suggest

that, at least in part, the response pattern has some evaluative content (see Table F for correlation table of all dependent variables).

One of the questions that arose from Study 1 regarded the correlational analysis and whether the correlations between the EC and direct measure of attitude varied between social desirability conditions. The correlations between the EC and direct measure in Study 2 for both social desirability conditions did not differ (low social desirability: $r(232) = .21, p = .05$; high social desirability: $r(232) = .19, p = .05$) $z = .21, p = .41$. As discussed in Study 1 this result is not necessarily problematic; if all the participants were adjusting their responses by the same amount when they responded this would not affect the correlation between the measures. As a result, the homogeneity in participant responses may have been a result of the strong social desirability manipulation, which was designed to create a situation where participants would not be highly variable in their responding.

In the 'time' condition, if restricting the amount of time improved performance of the EC measure, then the correlation between the EC and direct measure should be stronger. However, in the time 'restricted' condition the correlation between the two measures was no stronger ($r(232) = .17, p = .06$) than the 'unrestricted' condition ($r(232) = .21, p = .03$) $z = .43, p = .32$.

If the IAT is, at least partially, capturing the same evaluative content as the direct measure, then one might expect that it would be correlated with the direct measure of attitude. However, the analysis revealed that there was no significant correlation between the IAT and the direct measure ($r(232) = -.06, p = .34$). Additionally, there was no correlation between the EC measure and the IAT

($r(232) = .02, p = .75$). This suggests that while the EC and the direct measure are, to some degree, tapping into the same construct the IAT is measuring something else.

Evaluating the effect of socially desirable responding.

Direct measure of attitude. A 2 (social desirability: high social desirability versus low social desirability) x 2 (time: restricted versus unrestricted) x 2 (filler: with filler items versus without filler items) x 2 (order of indirect measures: IAT/EC versus EC/IAT) analysis of variance was conducted with the direct measure as the dependent variable. It was expected that there would be a main effect of social desirability, as found in Study 1. There were no other main or higher order effects expected in this analysis. As hypothesized, the analysis revealed that participant's scores on the direct measure of attitude were significantly different across social desirability conditions, $F(1,234) = 6.05, p < .01$, which indicated that participants' reported more favourable attitudes toward people who are overweight in the high social desirability condition (Low: $M = 26.97, SD = 6.09$; High: $M = 28.99, SD = 6.58$). This confirmed that the social desirability manipulation was successful⁶. Additionally, there were no significant 2 or 3-way interactions.

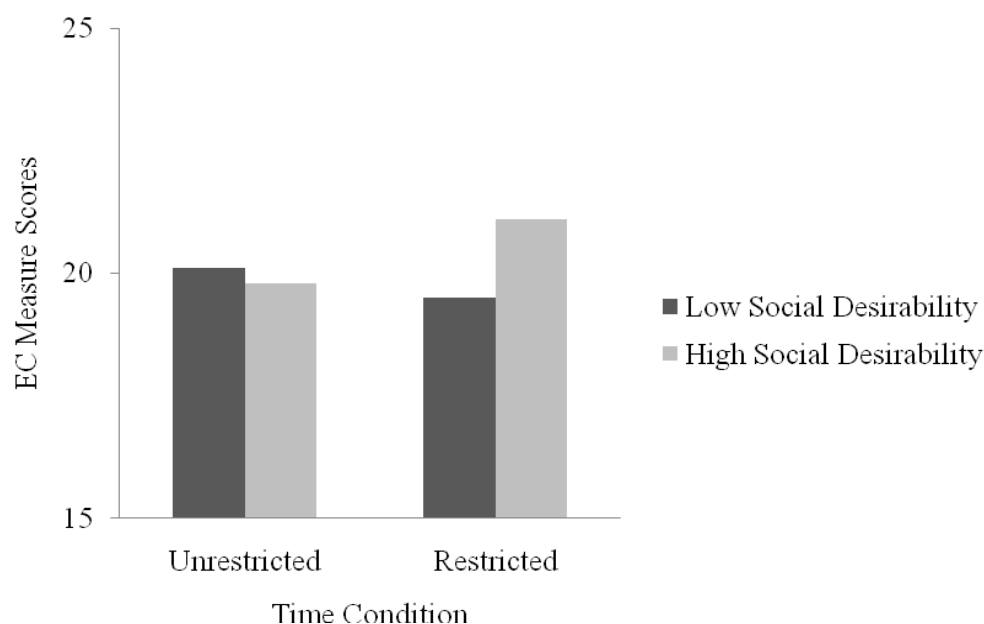
EC measure of attitude. The 2 (social desirability: high social desirability versus low social desirability) x 2 (time: restricted versus unrestricted) x 2 (order of indirect measures: IAT/EC versus EC/IAT) analysis of variance procedure was repeated with

⁶ As in Study 1, a mixed-model analysis of variance was conducted with the type of measure as the within-subjects factor with 2 levels (i.e., EC and the direct measure) and Social Desirability as the between-subjects factor. The results suggest that the effects of socially desirable responding were weaker on the EC measure than the direct measure of attitude ($F(1,231) = 3.12, p = .09$). Note, both the EC and direct measure were re-scaled to a common metric from 0 to 1.

the EC measure as the dependent variable. It was expected that the results of the analysis of variance would find no significant main effects or interactions between the variables. As hypothesized, the analysis revealed that there was no significant main effect of social desirability, $F(1, 234) = .08, p = .78$. This suggests that participant's scores did not reflect more favourable attitudes toward people who are overweight in the high social desirability condition compared to the low condition (Low: $M = 20.25, SD = 4.86$; High: $M = 20.41, SD = 4.10$). Although no main effect was expected for the EC/IAT order condition, it was significant ($F(1, 234) = 4.93, p = .03$; Low Social Desirability: $M = 20.82, SD = 5.01$; High Social Desirability: $M = 19.78, SD = 3.77$). There is no theoretical rationale for this main effect and without replication of this result no speculation regarding its interpretation will be made at this time.

Although none of the other 2-way interactions were significant, there was a significant interaction between social desirability and time (see Figure 4.). This result is unexpected given the results of Study 1 and the previously described analyses of the present study. That is to say, given that the non-significant results the EC measure had already demonstrated in the social desirability condition in Study 1, it was unclear how restricting time might improve upon that result. In fact, this interaction showed that in the 'unrestricted' time condition (this condition is a replication of study 1) there was no significant difference between the participant's EC scores regardless of the social desirability condition (Low: $M = 20.82, SD = 4.49$; High: $M = 19.78, SD = 3.77$). However, in the 'restricted' time condition participants in the 'high' social desirability condition reported more favourable attitudes towards people who are overweight (Low: $M = 19.68, SD = 4.26$; High: $M = 21.05, SD = 4.33$).

Figure 4. Effects of Time on Socially Desirable Responding



The IAT. The analysis of variance procedure was repeated with the IAT as the dependent variable. No significant main effects or interactions were expected from this analysis of variance. There were no significant main effects; however, there was one significant interaction between the time and EC/IAT order condition. When examining the means from these conditions it appears as though in the ‘restricted’ time condition when the EC measure is given first the IAT score is lower (EC/IAT: $M = 1.14$, $SD = .37$; IAT/EC: $M = 1.31$, $SD = .35$), while there is no such difference in the ‘unrestricted’ time condition (EC/IAT: $M = 1.31$, $SD = .38$; IAT/EC: $M = 1.27$, $SD = .39$). Since the IAT was not included in the time condition it is unclear why this interaction has taken place and there is no obvious theoretical explanation.

Self-monitoring scale. The analyses of the SM scale in Study 2 replicated those found in Study 1. As in the first study, this analysis revealed that there were no

significant correlations between the SM scale and direct measure of attitude in either the high ($r(68) = .11, p = .20$) or low ($r(112) = .14, p = .19$) social desirability condition.

As well, there were no significant correlations between the SM scale and the EC measure in either social desirability condition (high: $r(119) = .02, p = .85$; low: $r(112) = .02, p = .80$).

Additionally, the GLM analyses conducted in Study 1 were also repeated to examine whether self-monitoring may moderate the relationship between the social desirability manipulation and the EC and/or the direct measure of attitude. These analyses revealed that there was no significant interaction between the SM scale and the social desirability manipulation when either the direct measure ($F(1,234) = .04, p = .85$) or the EC ($F(1,234) = 1.71, p = .20$) were the dependent variables. These analyses suggest that the self-monitoring scale was not functioning as a moderator between social desirability manipulation and the attitude measures.

The Balanced Inventory of Desirable Responding – Version 6 (BIDR). As mentioned earlier, because impression management can serve to conceal inappropriate behaviours and opinions one might postulate that those high in impression management may be particularly prone to socially desirable responding. If participants high in impression management are more likely to engage in socially desirable responding than those low in impression management, then there would be positive correlations between their respective attitudes and impression management scores in the social desirability manipulation. The analysis revealed that there was a significant correlation between the direct measure of attitude and the IM scale in the low social desirability condition ($r(112) = .19, p = .05$), and a marginally significant correlation in the high social

desirability condition ($r(119) = .15, p < .10$). This result is not surprising, as one might expect that those who are high in impression management would inflate their scores on the direct attitude measure. It should be noted that these correlations were not significantly different from each other ($z = .31, p = .64$). There were, however, no significant correlations between the IM scale and EC measure in either the low ($r(112) = .11, p < .24$) or high ($r(119) = .06, p = .52$) social desirability conditions. The lack of correlation between the IM scale and EC measure in the social desirability provides some additional support for the EC measure's resistance to the social desirability manipulation.

Next, to examine whether the IM scale functioned as a moderator between the direct measure and the social desirability manipulation a GLM was specified with a dummy variable representing the social desirability manipulation, a continuous variable representing self-monitoring, and an interaction term (IM scale x social desirability) as predictor variables for the direct measure of attitude. Before performing these analyses participants' IM scores were centred. The GLM analysis revealed that there was no significant interaction ($F(1,234) = .01, p = .94$) between the IM scale and social desirability manipulation. The GLM analysis was repeated with EC as the dependent variable, and again there was no significant interaction ($F(1,234) = .01, p = .94$) between the IM scale or either condition of the social desirability manipulation. These analyses indicate that the impression management scale was not functioning as a moderator between the social desirability manipulation and the attitude measures.

Finally, to examine whether the social desirability manipulation had an effect on participants IM scores a GLM analysis was performed. The specified model was a 2

(social desirability: high versus low) x 2 (time: restricted versus unrestricted) x 2 (filler item ratio: 2:1 versus 0:1) with the IM scale as the dependent variable. There was no significant main effect of the social desirability manipulation ($F(1,234) = .56, p = .45$) nor any higher order effects. This suggests that participants did not report higher levels of impression management as a result of the social desirability manipulation. This result may be somewhat surprising, suggesting that participants did not show higher impression management scores when they were placed in the high social desirability condition. However, recall that the social desirability manipulation was very issue specific and the IM scale is a measure of general socially desirable responding.

Discussion.

The purpose of Study 2 was first to replicate of the systematic response pattern and the effects of social desirability responding on the EC measure and a direct self-report measure of attitude found in Study 1. Second, to study systematically examined ‘time restriction’ which was presumed to be an optimal condition for the functioning of an EC measure. Finally, compare the EC technique with the IAT.

Overall, this study provided some additional support for the EC technique. Using the same statistical methods as Study 1, the dimensionality of the systematic guess pattern was replicated. Furthermore, the correlation between the EC measure and the direct measure of attitude indicates that the systematic guess pattern was again, at least in part, attitudinal. This study also systematically examined ‘time restriction’, which was presumed to be an optimal condition for the functioning of an EC measure. This was not the case. Rather, this study found evidence that not only did restricting time not improve the performance of the EC measure; it may have actually increased its

susceptibility to socially desirable responding. Finally, this study compared the EC technique with the IAT. There was no significant main effect of the social desirability manipulation for the IAT; neither did this measure correlate with the EC or the direct attitude measure. The latter finding is consistent with the results from other IAT research (Greenwald, McGhee, & Schwartz, 1998; Nosek & Banaji, 2001, Olson & Fazio, 2004). These researchers have argued that the modest correlations between indirect and direct measures of attitude might be because they measure fundamentally different types of attitudes (i.e., implicit and explicit). That being said, it might still be expected that two indirect measures of attitude would correlate with each other since they would both be measuring the same type of attitude, however, this was not the case. It is unclear why the EC measure did not correlate with the IAT.

Based on the aforementioned results future EC research should re-examine the effects of time restriction on the EC measure. As well, the time restriction condition should be extended to include both the EC and direct measure of attitude. Finally, future EC research should examine how the EC measure performs compared with other contemporary indirect measures of attitude.

Chapter 4

Examining Item Ratio, an Optimal Feature of the Error Choice (EC) Technique and Comparing the EC with the Personalized Implicit Association Test.

This study was a replication of Study 2 (i.e., dimensionality of EC items, the effects of socially desirability responding and time on the EC measure and a direct self-report measure of attitude found) with three exceptions. First, the time condition in this study was extended to include both the EC and direct measures of attitude. In Study 2, only the EC measure was placed under time constraint. Next, this study examined another previously identified optimal condition for EC measures (i.e., target to filler item ratio). Following his initial study, Hammond (1948) recommended that to mask the purpose of the target items there should be at least twice as many filler items as target items in the EC measure. This recommendation, however, has never been empirically tested. Finally, this study compared the EC technique to another current indirect measure of attitude (i.e., the Personalized Implicit Association Test, PIAT).

The PIAT (Olson & Fazio, 2004) was selected because it has correlated more strongly with direct measures of attitude compared to other indirect measures of attitude (i.e., the IAT). In general, the PIAT procedure is quite similar to that of the IAT. The changes to the IAT procedure resulted from concerns raised by Olson and Fazio (2004) that the IAT could be contaminated by extrapersonal associations. Olson and Fazio described these extrapersonal associations as information (either from their culture or specific others) that a person has available in memory, but do not contribute to their evaluation of an attitude object. They argued that what are most relevant to a person's assessment of an attitude object are their own experiences with that object. More

specifically, Olson and Fazio describe 3 aspects of the IAT that they believed contributed to its contamination by extrapersonal associations; the normative implication of the category labels (i.e., “Pleasant” and “Unpleasant”), the category items themselves, which are portrayed as either normatively positive or negative, and the error feedback provided by the IAT (i.e., a red “x” appeared for incorrect response and the IAT would not continue until the correct response was made) that reinforced the normatively correct responses of each participant.

In order to overcome these concerns, Olson and Fazio (2004) made a number of recommendations. First, that the category labels displayed at the top of the computer screen should be selected such that it would be unlikely that they would have a normative construal (e.g., “I liked” and “I don’t like”). Olson and Fazio suggested that these types of category labels might decrease the normative influence associated with “Pleasant” and “Unpleasant” when categorising IAT items. Additionally, they recommended selecting category items that were not associated with a normative positive or negative valence. Rather, category items should be value-laden; however, there should be a lack of social consensus regarding their overall evaluation, enabling a more idiosyncratic assessment to be made by the participant. In order to accomplish this, Olson and Fazio developed a list of words that would be used in place of the items in both the “Pleasant” and “Unpleasant” categories (see appendix, Olson & Fazio, 2004). This list was comprised of words that would bring to mind an attitude (i.e., coffee, disco, jogging, opera, etc.); however, it was unlikely that these words would be normatively connected with a particular valence. Finally, Olson and Fazio recommended that participants should not receive any performance feedback during the

categorisation procedure, because it further reinforced the normative quality of the category items.

Method

Participants

A total of 354 undergraduate students were recruited from either introductory psychology or other university courses at Queen's University. All participants received either partial course credit or five dollars for their participation. A total of 5 participants were dropped from the study because they did not complete all of the attitude measures or the response latencies had an excessive error rate (i.e., 30% or greater as recommended in Greenwald, Nosek, & Banaji, 2003). As a result, the data from 349 participants were included in the analyses.

Procedure

The design of this study was a 2 (social desirability: high social desirability versus low social desirability) x 2 (filler to target item ratio: 2:1 versus 0:1) x 2 (time: restricted versus unrestricted) x 2 (order of scales: PIAT/EC versus EC/PIAT). The sessions were conducted in groups of up to 5 participants. Before each session, the groups were randomly assigned to one of the social desirability conditions. The item ratio, EC/IAT order and time restriction conditions were assigned at the individual level within sessions. The procedures for Study 3 were, for the most part, the same as Study 2. Following their consent to participate in the study participants were told by the research assistant that they would be completing several computer based tasks including a categorisation task and a general knowledge test about eating habits, weight and exercise (See Appendix G for all Study 3 measures). The optimal method for PIAT

delivery is computer based; as a result digital versions of all measures were used in the study. The social desirability, time and EC/PIAT order of presentation manipulations remained unchanged from Study 2.

EC filler to target item ratio. In the item ratio condition participants received either an EC measure that was comprised of at least twice as many filler items (8) as EC target items (23) or an EC measure in which all of the filler items had been removed (0:1) and consisted of only the 8 EC target items. The instructions for completing the short version of the EC measure remained unchanged for the time restricted condition. That is to say, participants were told that they would only have 3 minutes and 30 seconds to complete the EC measure regardless of whether it contained all 31 or only 8 EC questions. Since the participants did not know the number of items contained in any of the measures, the perception of time pressure to complete EC measures was maintained.

Measures

The Error-Choice (EC) Measure. The EC measure used in this study was the same as study 2 and consisted of 23 neutral (filler) and 8 target items. The computerised version of the EC measure did not allow participants to go back to previously answered questions and review or change their responses ($M = 20.62$, $SD = 4.68$, Range = 24.00, Skewness = $-.10$, Kurtosis = $-.35$).

Direct measure of attitude. Participants' attitudes were also assessed using the same 8 item semantic differential scale as Studies 1 and 2 ($M = 27.61$, $SD = 6.68$, Range = 44.00, Skewness = $-.15$, Kurtosis = 1.41). The Cronbach alpha coefficient for the direct measure used in the present study was $\alpha = 0.90$.

The Personalized Implicit Association Test (PIAT). The PIAT protocol used in this study was based on the procedure described in Olson and Fazio (2004), which required participants to complete 3 practice and 4 experimental trials. Following the PIAT procedure described in Olson and Fazio (2004) the category labels in this study were changed to “I like” and “I don’t like” rather than the “Good” and “Bad” labels used in study 2. Additionally, no performance feedback was provided to the participants during the PIAT procedure. The PIAT items for the “Slim” and “Overweight” categories were the same as those used in the IAT procedure in study 2. The category items for “I like” and “I don’t like” were selected from the list of idiosyncratic items listed in the appendix of Olson and Fazio (2004). These items were selected because they had no extrapersonal culturally associated valence. These items formed a common pool ($n = 24$) that were displayed with the category labels “I like” and “I don’t like”. In each of the experimental trials all of the words from the 3 pools of categories (i.e., Slim, Overweight, and I like and I don’t like) were presented randomly (without replacement). As in Olson and Fazio (2004) and study 2 the improved IAT scoring algorithm described in Greenwald et al. (2003) was used to calculate participants PIAT scores.

For the PIAT experimental trials, 2 pairs of category labels consisting of an attitudinal and evaluative dimension appeared at the top of the computer monitor. In the first 2 experimental trials the paired category labels were ‘Overweight/I don’t like’ and ‘Slim/I Like’. Participants were asked to correctly identify, as quickly and accurately as possible, which category the presented word belonged to by hitting the appropriate computer response key. Although there were 4 categories there were only 2 assigned

response keys for the participants to correctly categorise each of the words. That is to say, participants had to hit the same response key if the word belonged to either of the 2 category labels that constituted that pair of labels. For example, in the first set of experimental trials if the random word sequence was ‘obese’ and a word the participant associated with something they did not like then the participant would hit the same response key for both words.

The first 2 PIAT experimental trials were considered to be the congruent trials, because the concept and evaluative labels that people might be expected to normally associate with each other are paired together (i.e., Overweight/I don’t like and Slim/I Like). Consequently, if people do associate “I don’t like” with “Overweight” and these 2 categories share the same response key, then the response latency between these items should be quite small. In the last 2 experimental trials, the evaluative dimension labels were reversed so the new paired categories were “Overweight/ I like” and “Slim/ I don’t like” with each pair again sharing the same response key. If a person does not normally associate “I like” with “overweight” (i.e., incongruent) then the response latency between words associated with these categories will be larger than those in the congruent trials. The size of the difference between the two sets of trials estimates the degree to which the participant’s evaluations differ, whereby large positive differences between the congruent and incongruent trials reflect more negative associations with “Overweight”.

Self-monitoring scale. For exploratory purposes the self-monitoring scale (Snyder, 1974) used in Study 1 and 2 (Cronbach $\alpha = .76$) was also included in the present study ($M = 9.75$, $SD = 3.37$, Range = 18.00, Skewness = 0.01, Kurtosis = -.61).

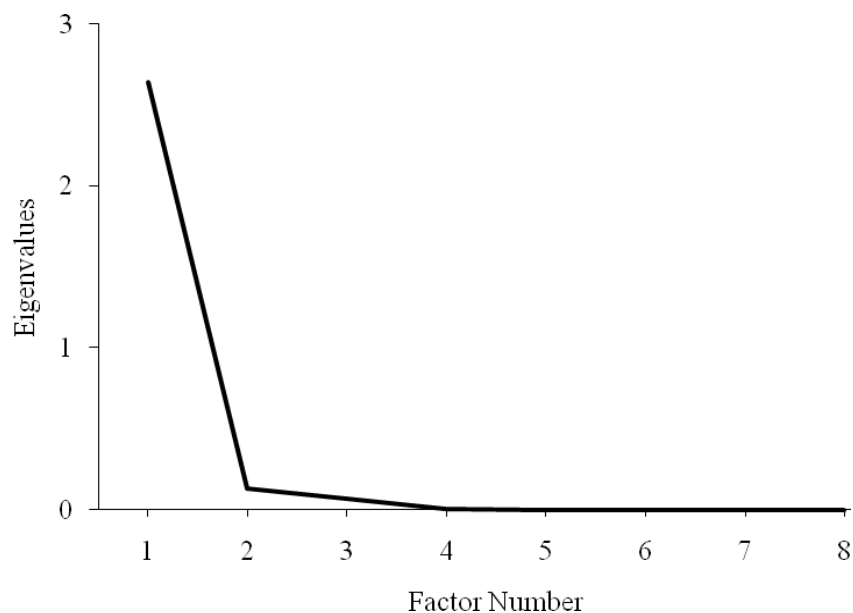
The Balanced Inventory of Desirable Responding – Version 6 (BIDR). For exploratory purposes the Impression Management (IM) scale of the BIDR (Paulhus, 1991) used in Study 2 (Cronbach $\alpha = .63$) was also included in the present study ($M = 5.86$, $SD = 3.10$, Range = 16.00, Skewness = 0.63, Kurtosis = 0.31).

Results

Systematic Guess Pattern

Exploratory factor analysis. The initial purpose of this study was to replicate participant's systematic guess pattern to the EC target items that was found in studies 1 and 2. If participants' answers to the EC target items does reveal a systematic pattern of responding, then the dimensionality of the factor analysis should reveal a single latent construct. Consistent with the exploratory factor analysis methodology used in studies 1 and 2, the same 3 statistical procedures were used to identify the number of factors. The plot of the eigenvalues from the reduced correlation matrix (see Figure 5)

Figure 5. Scree plot of Eigenvalues from the Reduced Correlation Matrix



revealed that there was one dominant factor; however, one cannot unequivocally rule out a 2 or 3 factor interpretation.

The parallel analysis suggested that a 1 factor solution best represented the data (see Table 12) because only the first eigenvalue from the actual data set was greater than those produced by random data.

Table 12. *Parallel Analysis – Eigenvalues from the Reduced Correlation Matrix.*

Factor number	Eigenvalues from actual data set	Eigenvalues from random data
1	2.640	.351
2	.129	.222
3	.067	.149
4	.006	.064
5	-.062	.003
6	-.106	-.037
7	-.171	-.083
8	-.203	-.140

The Root Mean Square Error of Approximation (RMSEA) was used to examine model fit. As in study 1, 3 separate factor models were examined for the EC measure

(see Table 13). After examining the RMSEA results for the 1 factor model (see Table 8, row 1) it appeared as though this model did provide good fit to the data (.044). The RMSEA value for the 2 factor model (.037) dropped only slightly from the one factor model and there was almost complete overlap in the confidence intervals of these models. Although the 3 factor model provided a small improvement in RMSEA value (a decrease of .017); however, there was a notable increase in the confidence interval, which completely overlapped the confidence intervals of both the 1 and 2 factor models. Given that the 1 factor model provided good fit and the confidence intervals were almost completely overlapping, these analyses favoured the more parsimonious 1 factor model, but nonetheless the 2 and 3 factor models were also considered further.

Table 13. *RMSEA Goodness-of-Model Fit.*

Number of Factors	RMSEA	Confidence Interval (90%)
1	0.044	0.013-0.069
2	0.037	0.000-0.069
3	0.020	0.000-0.071

The preponderance of evidence from these 3 statistical approaches suggests that the 1 factor model may be most appropriate for the data. However, in order to make a final determination in identifying the correct number of factors their interpretability was also examined. To accomplish this an exploratory factor analysis was carried out on the 8 EC target items using the same factor analytic procedure as studies 1 and 2. In the 1

factor model all 8 EC target items provided moderate to strong factor loadings on a single factor (see Table 14).

The interpretation of the 2 factor model was more difficult (see Table 15 for the factor pattern matrix). Five of the EC target items that loaded strongly in the 1 factor model remained substantial loadings in the 2 factor model and they were joined by one of the lower loading EC target items (1). The second factor was comprised of 3 items (5, 7 and 8). Item 7 loaded strongly on this factor (i.e., how many students

Table 14. *Factor Matrix for the 1 Factor Model.*

EC Target Items	Factor Loadings
EC 7	.73
EC 5	.73
EC 3	.63
EC 6	.61
EC 4	.53
EC 8	.49
EC 2	.48
EC 1	.37

reported being disgusted by obese people), while item 5 (i.e., what percentage of students said they were repulsed by overweight people) loaded on both factors and item

8 loaded only weakly on the second factor (i.e., what percentage of people are repulsed by the idea of romantic involvement with someone who is obese). The structure and content of these EC target items was not conceptually different than the EC target items that made up the first factor. Additionally, both factors were strongly correlated ($r = .74, p < .01$) which indicated that there was a large overlap in the underlying construct of both factors in this analysis.

Table 15. *Pattern Matrix for the 2 Factor Model.*

EC Target Items	Factor 1 Loadings	Factor 2 Loadings
EC 4	.60	-.02
EC 2	.52	.01
EC 1	.43	-.03
EC 3	.43	.25
EC 6	.38	.26
EC 7	-.07	.89
EC 5	.30	.47
EC 8	.21	.31

Note: Extraction Method: Maximum Likelihood.
Rotation Method: Oblimin with Kaiser Normalization.

The first factor in the 3 factor model was comprised of 6 EC target items, while the second and third factors were each made up of a single item. Although the factor loading for the item in each of these factors was quite strong, because they had of only a single item they did not contribute to a parsimonious interpretation of the data. As well, there was a moderate correlation between the factors (factors 1 and 2: $r = .42$; factor 1 and 3: $r = .59$; factor 2 and 3: $r = .34$), which suggested that there was some overlap in the underlying construct of the 3 factors. Given the more complex nature and difficulty interpreting the factors contained in the 2 and 3 factor models, the 1 factor solution provided the most parsimonious and interpretable solution for the data.

Table 16. Factor Pattern Matrix for the 3 Factor Model.

EC Target Items	Factor 1 Loadings	Factor 2 Loadings	Factor 3 Loadings
EC 4	.62	-.03	-.04
EC 2	.58	-.05	-.04
EC 3	.57	-.04	.13
EC 5	.50	.04	.24
EC 6	.47	.13	.09
EC 1	.37	.09	-.03
EC 8	.01	.99	.01
EC 7	.36	.02	.97

Note: Extraction Method: Maximum Likelihood.
Rotation Method: Oblimin with Kaiser Normalization.

Given that a 1 factor solution once again best represents the data, the next question is what, if any, effect might the optimal conditions have on the EC measure. This study examined two of the conditions; restricting the time available to complete an EC measure, and removing all of the filler items have on the factor loadings of the EC target items.

A factor analysis of the ‘time’ condition revealed that the factor loadings of the EC target items showed trivial changes between the time ‘restricted’ and ‘unrestricted’ conditions (see Table 17). Moreover, the factor loadings of the EC target items of both models remained above recommended cutoffs with the exception of item 1 in the time restricted condition.

Table 17. *Factor Matrix for the Time Condition.*

EC Target Items	Unrestricted Time Loadings	Restricted Time Loadings
EC 5	.76	.68
EC 7	.73	.72
EC 3	.66	.58
EC 6	.57	.64
EC 2	.51	.46
EC 4	.51	.54
EC 8	.47	.53
EC 1	.40	.33

Additionally, the reliability of the EC measure ($\alpha = 0.80$) between the 2 time conditions remained virtually unchanged (Unrestricted Time, $\alpha = 0.79$; Restricted Time, $\alpha = 0.80$). Therefore, restricting the amount of time participants had to complete the EC questions did not appear to improve the factor loadings or reliability of the EC measure.

The factor analyses of the ‘filler item’ conditions revealed that the factor loadings of the EC target items in the ‘filler item’ conditions showed minimal changes (see Table 18). The factor loadings remained above recommended cutoffs with the

Table 18. *Factor Matrix for the Filler Item Condition.*

EC Target Items	With Filler Items Loadings	Without Filler Items Loadings
EC 7	.74	.70
EC 5	.73	.71
EC 6	.63	.60
EC 3	.63	.60
EC 8	.57	.42
EC 4	.53	.53
EC 2	.51	.45
EC 1	.36	.36

exception of item 1 in both of the models. Additionally, the reliability of the EC measure between the 2 time conditions remained virtually unchanged (With Filler

Items, $\alpha = 0.81$; Without Filler Items, $\alpha = 0.78$). As a result, there is no evidence that having filler items improved the performance of the EC measure.

Interpreting the correlations between the measures.

If it is the case that participants' scores across the EC target items were being influenced by their attitudes, their EC scores should be correlated with the direct measure (see Appendix H for a correlation table of all dependent variables). In fact, the correlation between these 2 measures was significant, $r(347) = .29, p < .01$. Although the correlation is modest it does suggest that, at least in part, the response pattern has some evaluative content.

The correlations between the EC and direct measure in study 3 only changed slightly across the social desirability conditions (low social desirability: $r(347) = .35, p < .01$; high social desirability: $r(347) = .24, p < .01$). These findings replicate the weak tendency found in studies 1 and 2, whereby participants in the high social desirability condition responded in a desirable way to the direct measure, while the EC measure remained unaffected. However, as in Studies 1 and 2 the difference between the correlations between these conditions was not significant ($z = 1.21, p = .11$).

In the time conditions, if restricting the amount of time improves the EC measure's resistance to socially desirable responding then the correlation between the EC and direct measure should be stronger. In fact, the time 'restricted' condition correlation between the 2 measures was stronger ($r(171) = .37, p < .01$) than the 'unrestricted' condition ($r(176) = .16, p = .03$). The difference between the correlations in these conditions was significant ($z = 2.10, p = .02$). It appears as though restricting the amount of time participants had to complete the EC measure in this study did

increase the correlation between the EC and direct measure. This finding is contrary to what was found in Study 2, which did not reveal any difference between the EC and direct measure correlations in the time condition. However, the time condition in Studies 2 and 3 may not be directly comparable, because unlike Study 2, where just the EC measure was placed under time restriction, in this study both the EC and direct measure were placed in this condition.

In the ‘filler item’ conditions, if removing the filler items impaired the performance of the EC measure, as suggested by Hammond, then the correlation between the EC and direct measure should have been weaker. In the ‘without filler item’ condition; however, the correlation between the EC and direct measures ($r(176) = .28, p < .01$) was not appreciably different from the ‘with filler item’ condition ($r(171) = .26, p < .01$). Not surprisingly, the difference between these correlations was not significant ($z = .20, p = .42$).

The next question is; does the PIAT also capture the same construct as the direct measure of attitude? The analysis revealed that there was a significant correlation between the PIAT and the direct measure ($r(340) = .24, p < .01$). This suggests that the PIAT and the direct measure are, at least to some degree, tapping into the same construct. This correlation is somewhat weaker than Olson and Fazio (2004) reported between the PIAT and a semantic differential scale in their research ($r = .38$ and $r = .42$). It should be noted, however, that there are only 5 published PIAT articles, and the results from this study are more consistent with those of other PIAT researchers who found somewhat lower correlations between direct measures and the PIAT (Dambrun &

Villate, 2008; Nosek & Hansen, 2008). Interestingly, there was no significant correlation between the EC measure and the PIAT, $r(340) = .05, p = .38$.

Evaluating the effect of socially desirable responding.

Direct measure of attitude. A 2 (social desirability: high social desirability versus low social desirability) x 2 (time: restricted versus unrestricted) x 2 (filler: with filler items versus without filler items) x 2 (order of indirect measures: IAT/EC versus EC/IAT) analysis of variance was conducted with the direct measure as the dependent variable. It was expected that there would be a main effect of social desirability, as found in Studies 1 and 2. There were no other main or higher order effects expected in this analysis. As hypothesized the analysis revealed that participant's scores on the direct measure of attitude were significantly different across social desirability conditions, $F(1,349) = 10.22, p < .01$, which indicated that participants' reported more favourable attitudes towards people who are overweight in the high social desirability condition ($M = 29.32, SD = 6.30$) than the low social desirability condition ($M = 26.97, SD = 7.11$). This confirmed that the social desirability manipulation was successful⁷. As well, the analyses supported the hypotheses that participants' responses would be unaffected by whether they were placed under time pressure to complete the measure or that the order of indirect measures would have an effect on their responses. Additionally, there were no significant 2 or 3-way interactions.

⁷ The mixed-model analysis of variance conducted in Studies 1 and 2 was repeated for Study 3. The type of measure was the within-subjects factor (i.e., EC and the direct measure) and Social Desirability was the between-subjects factor. The results suggest that the effects of socially desirable responding were weaker on the EC measure than the direct measure of attitude ($F(1,327) = 3.20, p = .07$). Note, both the EC and direct measure were re-scaled to a common metric from 0 to 1.

EC measure of attitude. The analysis of variance procedure was repeated with the EC measure as the dependent variable. It was expected that the results of the analysis of variance would find no significant main effect of social desirability or interactions between the other variables. As hypothesized the analysis revealed that there was no significant main effect of social desirability, $F(1,349) = .75, p = .39$. This suggests that participants' scores did not reflect more favourable attitudes toward people are overweight in the high social desirability condition ($M = 20.88, SD = 5.19$) compared to the low social desirability condition ($M = 20.45, SD = 3.94$).

Although no main effects were expected, there was a main effect for the filler item condition, $F(1,349) = 5.24, p = .02$. On closer examination this analysis revealed that, when the filler items were present, participants reported more favourable attitudes towards people who were overweight ($M = 21.25, SD = 4.60$) than when the filler items were not present ($M = 20.07, SD = 4.67$). If filler items are necessary for the EC measure to function properly, then there should have been a significant 2-way interaction between the filler item and social desirability condition. However, the 2-way interaction between these 2 independent variables was not significant, $F(1,349) = .01, p = .93$. This suggests that regardless of whether the filler items were present or not participants did not report more favourable attitudes towards people who were overweight (see *Figure 6*).

The 2-way interaction between social desirability and time (see *Figure 7*) that was present in Study 2 did not replicate in Study 3, $F(1,349) = .03, p = .96$. This result is not problematic given that the non-significant results the EC measure had already

Figure 6. Effects of Filler Items on Socially Desirable Responding.

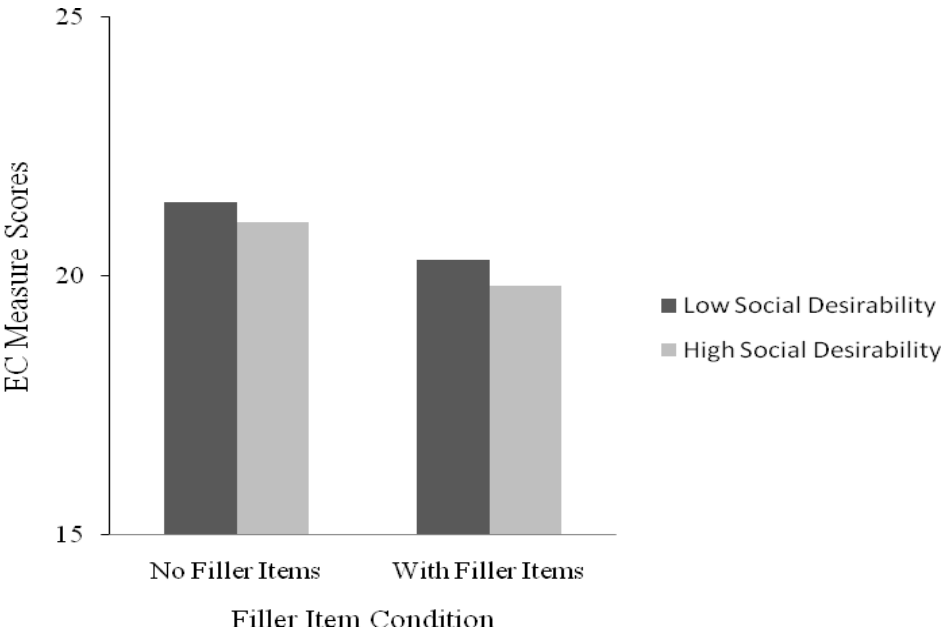
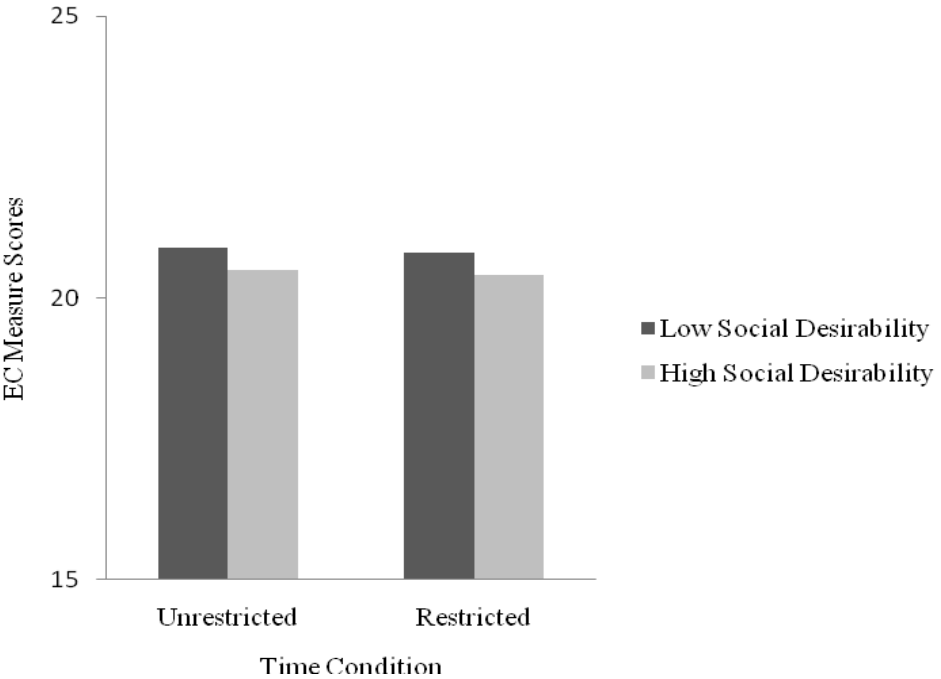


Figure 7. Effects of Time Restriction on Socially Desirable Responding.



demonstrated in the social desirability condition in studies 1 and 2, it was unclear how restricting time might have improved upon that result.

The PIAT. The analysis of variance procedure was repeated with the PIAT as the dependent variable. No significant main effects or interactions were expected from this analysis of variance. The ANOVA revealed that there was no significant main effect of social desirability ($F(1,340) = 2.33, p = .13$). This suggests that participant's performance on the PIAT was, for the most part, unaffected by the social desirability manipulation (Low: $M = .44, SD = .30$; High: $M = .49, SD = .32$). There were no significant interactions.

Self-monitoring scale. The analyses of the SM scale in Study 3 replicated those found in Study 1 and 2. As in the first and second studies, this analysis revealed that there were no significant correlations between the SM scale and direct measure of attitude in either the high, $r(174) = .02, p = .81$, or low, $r(173) = .08, p = .28$, social desirability conditions. As well, there were no significant correlations between the SM scale and the EC measure in either social desirability condition (high: $r(174) = .05, p = .50$; low: $r(173) = .00, p = .96$).

As well, the GLM analyses conducted in Study 1 were also repeated to examine whether self-monitoring may moderate the relationship between the social desirability manipulation and the EC and/or the direct measure of attitude. These analyses GLM analyses revealed that there was no significant interaction between the SM scale and the social desirability manipulation when either the direct measure ($F(1,346) = .00, p = .97$) or the EC were the dependent variable ($F(1,346) = 1.15, p = .29$). These GLM analyses

provide further evidence that the self-monitoring scale does not function as a moderator between the social desirability manipulation and the attitude measures.

The Balanced Inventory of Desirable Responding – Version 6 (BIDR). The analyses of the IM scale in Study 3 replicated those presented in Study 2. In these analyses, unlike Study 2, there were no significant correlations between the IM scale and the direct measure of attitude in the high ($r(174) = .07, p = .36$) or low ($r(173) = .06, p = .44$) social desirability condition. Additionally, there were no significant correlations between the IM scale and EC measure in either the high ($r(174) = .12, p = .11$) or low ($r(173) = .03, p = .70$) social desirability condition.

As in Study 2, the GLM analyses revealed that there was no significant interaction between the IM scale and the social desirability manipulation when either the direct measure ($F(1,346) = .00, p = .97$) or EC ($F(1,346) = 1.15, p = .29$) were the dependent variables. These analyses support the assertion that the IM scale does not function as a moderator between socially desirable responding and the attitude measures. Finally, I examined whether the social desirability manipulation had an effect on participants' IM scores a GLM analysis was performed. The specified model was a 2 (social desirability: high versus low) x 2 (time: restricted versus unrestricted) x 2 (filler item ratio: 2:1 versus 0:1) with the IM scale as the dependent variable. The analysis revealed that there was no significant main effect with the social desirability manipulation ($F(1,346) = .19, p = .67$) nor were there any higher order effects. This suggests that participants did not report higher levels of impression management as result of the social desirability manipulation, nor were their IM scores affected by any of the other experimental manipulations.

Discussion.

Using the same statistical methods as Studies 1 and 2 the systematic guess pattern was replicated. As well, the correlation between the EC measure and the direct measure of attitude indicated that the systematic guess pattern was again, at least in part, attitudinal. Additionally, this study systematically examined two of the recommended EC procedures thought to be optimal for the functioning of an EC measure (i.e., time restriction and filler item ratio). Although time restriction appeared to increase the EC measure's susceptibility to social desirability in study 2, this was not the case in study 3. It is unclear why placing participants under time pressure to complete the EC measure did not increase its susceptibility to socially desirable responding. The only difference in the time conditions between the 2 studies was that in study 3 'time restriction was placed on the both EC and direct measure, but because the direct measure was always given after the EC measure that cannot plausibly account for why the findings failed to replicate. The other EC procedural recommendation examined in this study was the number of filler items included in the EC measure. Again, contrary to what Hammond had expected removing the filler items from the EC measure did not negatively impact its performance. Rather, the EC measure seemed to perform well whether or not the filler items were present. This suggests that the purpose of the EC measure may be better concealed than even Hammond expected.

Finally, this study compared the PIAT with the direct measure of attitude and the EC technique. The analyses of the PIAT revealed that there was no significant main effect of social desirability, which indicated that participants' responses to the PIAT items were unaffected by the social desirability manipulation. As well, the PIAT did

correlate with the direct attitude measure, which suggests that it is at least partially measuring the same attitudinal construct as the direct measure. However, the PIAT did not correlate with the EC measure.

Chapter 5

General Discussion

Summary of findings and implications.

Insights into the functioning of the Error Choice technique. This programme of research has produced a number of important findings regarding indirect attitude measures, in particular the Error Choice (EC) technique. Hammond (1948) hypothesised that when responding to apparent knowledge based questions they did not know the answer to people would guess, but these guesses would not be random. Rather, their guesses would reveal a systematic response pattern that was consistent with that person's attitude. In each of the studies the exploratory factor analysis indicated that a single underlying dimension best accounted for the data. This suggested that if the participants responded positively (or negatively) to one of the EC target items then they tended to select positive (or negative) responses to the other EC target items. The presence of a systemic response pattern was also supported by the good reliability of the EC target items in Hammond's studies. These basic findings have been replicated by some EC researchers; however, the reliability of the items in these measures varied from low to moderate. Because many of these early EC studies did not provide any detail on how the EC items were developed it is impossible to accurately identify factors that may have contributed to these poor reliabilities. In the present group of studies, to ensure optimal item quality, the EC target items were selected based on prior item pretesting. As a result, the EC target items in the present measure consistently produced high reliability across all 3 studies.

Examining the reliability of the EC items was one of the few statistical approaches early researchers had available to examine the response pattern, however this is a necessary but not sufficient condition to determine dimensionality. This programme of research goes beyond these early EC studies by using an exploratory factor analysis to examine an EC measure for evidence of a one-dimensional systematic response pattern. If participants did respond systematically and the EC target items were homogenous, then a factor analysis should reveal a single latent construct. In fact, in all 3 of these studies there was compelling evidence supporting a one factor structure for the EC target items.

In previous EC studies researchers inferred whether the EC measure was attitudinal using the known groups paradigm. That is, early studies explored whether people's response patterns to the EC target items varied in ways that could be predicted based on their group membership. This programme of research has extended these findings by using direct correlations to provide additional support for the attitudinal nature of the EC measure. Surprisingly, directly correlating the EC measure with a direct measure of attitude to determine whether an EC measure was attitudinal had not been used by early EC researchers. The present research revealed a modest significant correlation between the two measures in each of the 3 studies. These findings, in combination with the results of the earlier EC research provides converging evidence that EC responses are, at least in part, attitudinal.

One point of interest regarding the correlations between the EC and direct measures that arose from the analyses of the three studies was whether the susceptibility of the direct measure versus the lack of susceptibility of the indirect measure to socially

desirable responding might affect the correlations between the two measures. Indeed, in each of the three studies there was a non-significant trend whereby the correlations between the direct and indirect measures were slightly higher in the low than in the high social desirability condition. Although it was not significant in any single study, a meta-analysis was conducted to determine if this trend might be significant when the correlations for the social desirability conditions were collapsed across the three studies. The results of the meta-analysis revealed that this trend was not significant ($z = 1.47$, $p = .14$).

One of the major criticisms of early indirect measures of attitude was that researchers had not attempted to demonstrate that they were any more resistant to socially desirable responding than direct measures. The lone exception to this criticism provided ambiguous results (Antonak & Livneh, 1995b). The present group of studies are the first attempt to examine this problem through the direct manipulation of social desirability. Across the studies the participants' EC scores did not change significantly between the high and low social desirability conditions, unlike the direct measure of attitude which showed that participants' scores increased significantly in the high social desirability condition (i.e., showed more favourable attitudes toward people who were overweight). The findings from these three studies provide very encouraging evidence that the EC measure is comparatively resistant to socially desirable responding.

Hammond (1948) and a number of EC researchers over the years have proposed guidelines for improving the efficacy of the EC technique. Although these guidelines have existed for many years EC researchers have failed to examine whether they were actually necessary. As a result, there has been considerable variability in EC measures

and methodology across past studies, possibly contributing to the inconsistent findings obtained in some of these EC studies. Consequently, it seems important that these guidelines be understood so that those that are critical to an EC measure's functioning can be included and those that do not improve its efficacy can be discarded. The present group of studies examined 2 of these guidelines (i.e., time restriction and filler items).

It has been argued that placing participants under time constraints would not allow them to over-think their responses to the EC questions, thereby providing a much accurate assessment of their attitude. In Studies 2 and 3 some of the hypothesised optimal conditions (i.e., time restriction and filler items) for the EC, as recommended by Hammond, were examined. In Study 2, contrary to what was expected, restricting the amount of time participants had to complete the EC measure made it more susceptible to socially desirable responding. Additionally, there was no evidence that restricting time increased the reliability or factor loadings of the EC target items. In Study 3 the time restriction condition was applied to both the EC and direct measure. Again, the factor loadings of the EC target items and the reliability of the EC measure remained unchanged between conditions. Unlike Study 2, however, the findings from this study revealed non-significant differences between the EC and direct measure scores across the time conditions. Given that the only difference for the time condition between the 2 studies was that the direct measure in Study 3 was also placed under time restriction and the direct measure was always administered after the EC measure, it is unclear why the results of these 2 studies are different.

Study 3 also examined EC filler to target item ratio. EC researchers have recommended that there should be at least 2:1 filler to target item ratio to help disguise the purpose of the EC measure. However, the results of Study 3 revealed that the EC measure performed equally well with or without filler items. It may be that the opaqueness of the EC target items, for some topics, may be sufficient to disguise the purpose of the EC measure. These findings suggest that the EC measures may be even more practical to use than early EC researchers initially suggested. That is to say, this would make the EC technique more efficient in situations where only a limited number of EC items could be included in the measure. Contrary to what was expected, removing the filler items from the EC measure did not negatively impact its performance. This suggests that in some cases, when the target items are sufficiently opaque, the purpose of the EC measure may be better disguised than Hammond had expected.

Finally, this programme of research compared the EC measure with more contemporary indirect measures of attitude. In Study 2 the IAT, probably the best known indirect measure, was compared with EC measure and the direct measure. The results revealed that the IAT did not correlate with the direct measure, nor did it correlate with the EC measure. In Study 3 the PIAT was compared to the EC and direct measure. The participants' PIAT and EC scores did show positive correlations with the direct measure, which suggested that they were, to some degree, measuring the same construct. The PIAT, however, did not correlate with the EC measure. The analyses in Studies 2 and 3 revealed that there was no significant main effect of social desirability

for either the IAT or PIAT, which indicated that participants' responses to these measures were unaffected by the social desirability manipulation.

General implications for early indirect measures of attitude. In addition to furthering research on the Error Choice technique, one of the corollaries of these findings is that they challenge criticisms made by Kidder and Campbell (1970) of indirect measures attitude measures. In their review of indirect measures of attitude the authors had concluded that these types of measures were inherently unreliable, did not correspond with direct measures and failed to demonstrate resistance to socially desirable responding. Consequently, Kidder and Campbell concluded that the use of indirect measures of attitude were simply not worth the risk. This present programme of research has examined the Error Choice technique, one of those early indirect measures of attitude, and provided evidence that this reproof of at least one of these early indirect measures of attitude may have been premature and unwarranted. Indeed, recent research by Vargas, von Hippel, and Petty (2004) found support for partially structured stimuli (i.e., a set of six short vignettes similar to a thematic appreciation test) another of the early indirect measures. These promising findings may suggest that perhaps the Error Choice technique, and other early indirect measures of attitude, may deserve further investigation.

Theoretical implications of the Error Choice technique. The analyses of the data from this programme of research have demonstrated that the Error Choice technique has some promise. Assuming that future research reinforces the validity of the Error Choice technique, what is the value of yet another indirect measure of attitude? A review of the indirect measurement of attitude research over the past 20

years reveals a couple of common characteristics of many these techniques. First, much of this research was predicated on the notion that when a person was presented an attitude object that had a particular evaluation it would facilitate the ease of response to a second attitude object that had the same evaluation, or conversely inhibit their response if the second attitude object had an inconsistent or incongruent evaluation. One of the most common methods used to examine the association between construct evaluations has been response latency (e.g., IAT, PIAT, Evaluative Priming, Go/No-Go Association Test, etc.). The underlying premise of this method has been that attitude objects that shared similar evaluations would activate one another faster than those that were contradictory, which would be reflected in the participants' response latencies (Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Greenwald, McGhee, & Schwartz, 1998; Nosek & Banaji, 2001; Olson & Fazio, 2004).

The aforementioned methods do provide a reasonable approach to the indirect measurement of attitudes; however, they have suffered problems with reliability (Buchner & Wippich, 2000). Additionally, like any category of measurement these approaches are probably subject to certain biases (Brauer, Wasel, & Niedenthal, 2000; Cunningham, Preacher, & Banaji, 2001). Because the Error Choice technique is not predicated on the same methodology, and perhaps not even the exact same psychological mechanisms, it may not be susceptible to the same types of biases. This is not to suggest the Error Choice technique is free of biases, but it may be that they are different types of biases. Because the Error Choice technique is a different approach to the attitude measurement problem, it could provide a very useful comparison point in a multi-method approach to attitude measurement.

Practical implications of the Error Choice technique. The Error Choice technique offers a number of practical benefits over the other current indirect approaches to attitude measurement (e.g., IAT, PIAT, Go/No-Go Association Test and Evaluative Priming). Foremost among these benefits is that almost all these current methods are comparatively lengthy to administer. For example, it has been recommended that the IAT contain 7 stages, each comprised of 5 to 50 trials (Greenwald, et al., 1998; Lane, Banaji, Nosek, & Greenwald, 2007) while an EC measure may function very well with only 8 items. Additionally, the newer indirect techniques are reliant upon moderately more elaborate technologies than the EC technique. Consequently, there are some contexts where the use of these methods might prove to be too unwieldy and difficult. The EC technique, however, is not faced with these limitations. Rather, the EC technique has demonstrated itself to be no more difficult to administer than a direct measure of attitude. The simplicity of the EC technique might increase the opportunities to conduct indirect measurement of attitude research outside the laboratory.

Unresolved Issues in the Current Programme of Research.

Undoubtedly this programme of research has provided some additional insight into the EC technique. That being said, there is also no doubt that this research has raised several additional questions.

Modest correlations between the EC and direct measure of attitude. Although there were significant correlations between the EC and direct measure, in each of the 3 studies they were only modest. The obvious question that arises then is why were these correlations not higher? One explanation often given when measures do not correlate

well is that perhaps it is due to their reliability. However, the reliabilities of both the EC and direct measures across the 3 studies were very good. Indeed, when corrected for attenuation there was only a small increase (average increase, $r = .05$) in the overall correlations between the direct and EC measure in each of the 3 studies (correlations corrected for attenuation: study 1, $r(120) = .38$, $p < .01$; study 2, $r(232) = .26$, $p < .01$; study 3, $r(347) = .36$, $p < .01$). Therefore, if the reliability of the EC and direct measure are contributing to the low correlations between the 2 measures, this contribution must be quite minimal.

Another possible explanation, presuming participants are responding somewhat heterogeneously, is that socially desirable responding is attenuating the correlations. One way to determine if this is the case would be to directly manipulate socially desirable responding, as was done in the 3 studies in this programme of research. The analyses from these studies revealed that there was a non-significant trend in the correlations between the EC and direct measure to increase in the low social desirability (SD) condition (Study 1, high SD $r(120) = .20$, $p = .11$, low SD $r(120) = .41$, $p < .01$; Study 2, high SD $r(232) = .19$, $p = .05$, low SD $r(232) = .21$, $p = .05$; Study 3, high SD $r(347) = .24$, $p < .01$, low SD $r(347) = .35$, $p < .01$). What these comparisons suggest is that although there is a small amount of attenuation taking place due to social desirability, even in the low social desirability condition (where presumably we have eliminated socially desirable responding) one can still see that the correlations never reach a very high level. This suggests two possibilities: that the social desirability is not really much of what is driving this attenuation, or alternatively that there is still substantial socially desirable responding even in the low social desirability condition.

This second explanation seems unlikely given that a number of procedures were incorporated into the experimental designs that have been shown to make salient participant anonymity (e.g., anonymity restated by research assistant and in instructions, no identification marks on measures, completed measure place in a box by the participant), thereby minimising social desirability concerns. Furthermore, based on the results of the direct measure across the 3 studies it appears as though these procedures did reduce socially desirable responding in the low social desirability condition. That being said, it is possible that even in the low social desirability condition that at least some socially desirable responding was still present, thus moderately attenuating the relationship between the EC and direct attitude measure.

So, it appears as though neither reliability nor socially desirable responding can fully account for the low correlations between the EC and direct measure. However, it is possible to estimate the effects from this combination of distortions. This can be accomplished by examining the high social desirability condition correlation (i.e., uncorrected for attenuation in reliability and social desirability). This correlation is where both sources of attenuation should be at their maximum (Study 1, $r = .20$; Study 2, $r = .19$; Study 3, $r = .24$). These uncorrected correlations in the high social desirability condition can then be compared with those in the low social desirability condition, where there should be very little socially desirable responding and the correlations have been corrected for attenuation. These latter correlations thus provide the upper boundary to what these correlations should be if both sources of attenuation are removed (Study 1, $r = .49$; Study 2, $r = .26$; Study 3, $r = .43$). When the correlations between the EC and direct measure are corrected for the attenuation of both reliability

and social desirability the effects are not trivial (average increase of $r = .18$). This suggests that the construct being measured by these instruments are related; however, it is still not clear that they are the same construct. These correlations are certainly not as high as what have been typically observed between direct measures of attitude.

Some attitude researchers have argued that the modest correlations between indirect and direct measures of attitude might be because they measure fundamentally different types of attitudes (Greenwald, McGhee, & Schwartz, 1998; Nosek & Banaji, 2001). That is, indirect measures may tap into an implicit attitude, while direct measures reflect an explicit attitude; consequently, one might not necessarily expect them to be strongly correlated with each other. In fact, the correlation between the IAT and the direct measure in Study 2 was not significant.

The finding from Study 2 is not inconsistent with what proponents of implicit attitudes have argued regarding the relationship between implicit and explicit attitudes. That is, because these two types of measures are capturing different attitudes they may not necessarily be correlated with each other. So if the direct measure is capturing an explicit attitude then the question becomes; whether the indirect measures should be correlated? If the indirect measures of attitude are capturing an implicit attitude one might expect that they would share at least a moderate significant correlation; however, in point of fact the IAT and EC measure in Study 2, and the PIAT and the EC measure in Study 3 were uncorrelated. Although these findings were not entirely unexpected, as attitude researchers have found that indirect measures of attitude frequently fail to provide even modest significant correlations (Decoster, Banner, Smith, & Semin, 2006; Kidder & Campbell, 1970; Petty, Fazio, & Brinol, 2009; Olsen & Fazio, 2003) it does

beg a couple of additional questions. Are the EC, PIAT and IAT capturing different implicit attitudes?

That is to say, do people have multiple implicit attitudes? It may seem reasonable to suggest that someone may possess two implicit attitudes; however, since none of the indirect measures in these studies correlated with each other, at the very least, it would have to be assumed that participants had three implicit attitudes. Additionally, given the lack of correlations generally found between implicit measures, which would suggest that with each additional indirect measure there would be a corresponding increase in implicit attitudes. At this point, the logic of this argument begins to become implausible. Additionally, this conclusion is not particularly parsimonious, because it is not clear why and under what circumstances it would be postulated that a person would have multiple implicit attitudes. Alternatively, it could be that either the IAT, PIAT and/or EC measure is capturing an implicit attitude, while the others are measuring something entirely different from an implicit attitude? This conclusion is more parsimonious but it still leaves open the questions of what are these other two measures capturing and which of these measures is in fact capturing the "real" implicit attitude?

There is yet another possible explanation for the low correlations between the EC and direct measure: the EC measure is assessing the same attitude as the direct measure, but for some reason it is doing so imperfectly. Perhaps it is the case that there is something about the nature of the Error Choice questions themselves that do not fully capture the attitude.

In retrospect, upon examining the EC target items it became apparent that they possessed an interesting characteristic. All of the EC target items appeared to be questions that were based on opinion poll or census results (i.e., the questions ask respondents to infer how others would act in a given situation). The EC target items were chosen because it was believed that people's inferences about how others would react would be driven by their own attitudes. That is to say, one way that people could select their responses to these questions would be to ask themselves; how would I react in the is situation? The person would then use this information to infer how others would react. For example, one of the target items asked participants what percentage of Canadians they thought were embarrassed by the weight of an obese family member, and then provided them with varying percentage levels of agreement (i.e., 16%, 38%, 61%, or 74%). This logic makes perfect sense and indeed is somewhat similar to the well-established effects of false consensus (i.e., that attitude shapes an individual's perceptions of the likelihood of others holding the same attitudes or engaging in the same behaviours; Ross, Green, & House, 1977).

Indeed, the EC target items used in the current studies are not alone. Although full EC measures were seldom provided in earlier EC research, in almost all cases the researcher provided a sample of at least 1 EC target item from their measures. When comparing the EC target items from the various studies it was interesting to note that almost all of the EC items appeared to be measuring a specific type of judgemental bias. That is to say, when participants were answering the EC target items they were using their own attitude to infer the prevalence of certain behaviours by others. Although this bias is almost certainly in part attitudinal, it may be only imperfectly capturing the

person's attitude. If true this might account for why the correlations between the EC and direct measure were not higher. One potential implication of this finding is that if it was possible to develop EC target items that tapped into a broader range of judgemental biases it might be possible for the EC measure more effectively capture the attitude.

Other unresolved issues. There are a couple of other unresolved issues that bear mentioning. First, this programme of research has provided evidence that the Error Choice Technique is at least in part attitudinal and is resistant to socially desirable responding. However, these studies only examined a single issue (i.e., attitudes towards people who are overweight) which begs the question; would the Error Choice Technique also function effectively for other attitude objects. Although there is certainly no theoretical reason to expect that this technique would not function as well for other attitude objects, it is nonetheless a question that requires further investigation.

Additionally, EC research has not yet directly examined the relationship between the Error Choice technique and behavioural outcomes. Although there is no theoretical reason to expect that the Error Choice technique would not be related to behaviour, it is a logical extension of this programme of research. Interestingly, indirect measures of attitude have been shown to predict behaviours that are difficult to predict with direct measures. In these studies a participant's attitude was inferred from the overt behaviour, judgments and/or bodily responses. For example, Dovidio et al. (1997) used seating distance and several physical behaviours to assess attitude. Additionally, some researchers have suggested that indirect measures may predict different kinds of behaviours than direct measures. That is, a person's unconscious behaviour (e.g., negative non-verbal) correlated higher with their scores on an indirect measure of

attitude compared to a direct measure (McConnell & Leibold, 2001; Poehlman, Uhlmann, Greenwald, & Banaji, 2005).

Self monitoring (SM) and impression management (IM) scales. One final issue to mention in this section pertains to the results of the SM and IM scales. The fact that these measures showed no effects with EC is entirely sensible, since the evidence provided by the present group of studies has shown the EC is comparatively free of social desirability. However, what may be puzzling to some readers is that both the SM and IM scales showed such few effects with the direct measure, which evidence suggests was vulnerable to social desirability. Regarding the SM scale there are a couple of points that should be considered. First, as was mentioned earlier, the SM scale probably captures much more general social desirability tendencies, while a very issue specific form of social desirability was explored in this programme of research. Second, although it seems reasonable to believe that people who engage in self-monitoring might be more susceptible to socially desirable responding, it should be remembered that the SM was not designed to measure social desirability per se. In terms of the IM scale, which is a measure of socially desirable responding, there were small effects in Study 2 and in Study 3, although not significant, the correlations were still in a positive direction. Given that the IM scale captures more general forms of socially desirable responding, it may be that specific nature of the social desirability examined in the present studies resulted in the weak effect.

Directions for Future Error Choice Technique Research.

One exciting feature of a programme of research that suggests new unresolved questions is that this sets the foundation for interesting novel new directions for future

research. Obviously the principal unresolved question is why the correlation between the EC and direct measure were modest. There appears to be two obvious strategies for dealing with this issue, which correspond to the possible explanations mentioned in the previous section. One possible explanation is that a substantial amount of social desirability is present, even in the low social desirability condition. If that is true, one way to test this would be to identify attitude objects where social desirability is at a minimum (i.e., the issue themselves are free of social desirability) and develop an EC scale for one of those attitude objects. Studies could then explore the extent to which the EC measure correlates with a direct measure of attitude. If the correlation is substantially higher this would suggest that, in part, the attenuation was perhaps due to some residual social desirable responding, even under low social desirability conditions. Of course, conducting this type of study would be fairly involved, in that it would require identification of a low social desirability attitude object, and the development of new EC items for that issue.

The second explanation for the low correlations was that perhaps the nature of the EC items themselves was the contributing factor. One approach to determine whether this was that case could be to use the same attitude object as the present programme of research (i.e., people who are overweight) and include an additional set of EC target items. The additional EC target items would not require participants to respond to the EC target items based on their beliefs about the prevalence of the behaviours of others towards overweight people. For example, items like “When compared to the general population what percentage of overweight people reported lower levels of self-esteem?” or “According to research being conducted at McGill

University, ___% of obese people are likely to have fulfilling romantic relationships.” This type of question is different from the current EC target items in that they ask how overweight people view themselves or possible attributes of overweight people. This new type of EC target item might also be biased by attitudes, but these might not involve the same kinds of judgmental bias as those that result from asking participants to predict other people’s behaviour. Therefore, these new EC items would also capture attitudes, but they may involve a different kind of bias of judgment; as a result, including them would enhance the correlations between the EC and direct measure of attitude.

There are a couple of additional directions that EC research should also pursue. One would be to examine whether the Error Choice technique is effective with other attitude objects. To some degree this will be achieved when examining the low correlations between the EC and direct measure by using a low social desirability attitude object. However, it would be also be useful look at other high social desirability objects in order to demonstrate that the results from the current programme of research were not idiosyncratic to that one high social desirability topic. Additionally, it would also be sensible to examine several low social desirability attitude objects. Finally, a logical extension of the EC research would include examining the relationship between the EC measure and behaviour. As discussed in the previous section indirect measures have been shown to predict behaviours that are impractical or difficult to predict with direct measures the relationship and/or indirect measures may predict different kinds of behaviours (i.e., implicit) than direct measures. Consequently, an appealing direction for future research might be to examine where the

Error Choice technique would be situated between these positions regarding the relationship between indirect measures of attitude and behavioural outcomes.

Conclusion.

The Error Choice technique has been in existence for over 60 years and for the most part neglected; however, current research suggests that this technique holds some promise as an approach to attitude measurement and is well worth resurrecting. So, rather than being simply a historical footnote the Error Choice technique should be viewed as a methodology with a distinguished history that may in fact have a promising future. Furthermore, it is my hope that this current programme of research will serve as a catalyst for future Error Choice investigations.

References

- Antonak, R. F. (1994). Development and psychometric analysis of an indirect measure of attitudes toward individuals with mental retardation using the error-choice method. *Mental Retardation, 32*, 347-355.
- Antonak, R. F., & Livneh, H. (1995a). Direct and indirect methods to measure attitudes toward persons with disabilities, with an exegesis of the error-choice test method. *Rehabilitation Psychology, 40*, 3-24
- Antonak, R. F., & Livneh, H. (1995b). Development, psychometric analysis, and validation of an error-choice test to measure attitudes toward persons with epilepsy. *Rehabilitation Psychology, 40*, 25-38.
- Balma, M. J., Maloney, J. C., & Lawshe, C. H. (1958). The role of the foreman in modern industry: I. the development of a measure of management identification. *Personnel Psychology, 11*, 195-205.
- Bernberg, R. E. (1951). The direction of perception technique of attitude measurement. *International Journal of Opinion & Attitude Research, 5*, 397-406.
- Bishop, M., & Slevin, B. (2004). Teachers' attitudes toward students with epilepsy: Results of a survey of elementary and middle school teachers. *Epilepsy & Behavior, 5*, 308-315.
- Blumenfeld, W. S. (1966). A research note on the method of error-choice. *Educational and Psychological Measurement, 26*, 847-851.
- Brauer, M., Wasel, W., & Niedenthal, P. (2000). Implicit and explicit components of prejudice. *Review of General Psychology, 4*, 79-101.

- Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research, 24*, 445-455.
- Buttery, T. J. (1978). Pre-service teacher's attitude regarding gifted children. *College Student Journal, 12*, 288-289.
- Campbell, D. T. (1950). The indirect assessment of social attitudes. *Psychological Bulletin, 47*, 15-38.
- Clarke, N. E., & Crewe, N. M. (2000). Stakeholder attitudes towards ADA Title I: Development of an indirect measurement method. *Rehabilitation Counselling Bulletin, 43*, 58-65.
- Coffin, T. E. (1941). Some conditions of suggestion and suggestibility: a study of certain attitudinal and situational factors influencing the process of suggestion. *Psychological Monographs, 53*, 125-127.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and application. *Journal of Applied Psychology, 78*, 98-104.
- Crites, S. L., Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1995). Bioelectrical echoes from evaluative categorization: II. A late positive brain potential that varies as a function of attitude registration rather than attitude report. *Journal of Personality and Social Psychology, 68*, 997-1013.
- Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C. & Banaji, M. R. (2003). Component processes of social evaluation. *Journal of Personality and Social Psychology, 85*, 639-649.

- Dambrun, M., Villate, M., & Richetin, J. (2008). Implicit racial attitudes and their relationships with explicit personal and cultural beliefs: What personalized and traditional IATs measure. *Current Research in Social Psychology, 13*, 185-198.
- DeCoster, J., Banner, M. J., Smith, E. R., & Semin, G. R. (2006). On the inexplicability of the implicit: Differences in the information provided by implicit and explicit tests. *Social Cognition, 24*, 5-21.
- Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33*, 510-540.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.
- Fabrigar, L. R., Krosnick, J. A., & MacDougall, B. L. (2005). Attitude measurement: Techniques for measuring the unobservable. In T. C. Brock, & C. M. Green, (Eds.), *Persuasion: Psychological insights and perspectives*, (2nd ed.) (pp.17-40). Thousand Oaks, CA.: Sage Publications, Inc.
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology, 54*, 297-327.
- Fazio, R.H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229-238.

- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013-1027.
- Fromme, A. (1941). On the use of certain qualitative methods of attitude research: A study of opinion on the methods of preventing war. *Journal of Social Psychology, 13*, 429-439.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.
- Greenwald, A. G., McGee, D. E., & Schwarz, J. L. K. (1998). Measuring individual differences in implicit cognition: the Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464-1480.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003a). Understanding and using the implicit association test: I. an improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197-216.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003b). "Understanding and using the implicit association test: I. an improved scoring algorithm": Correction to greenwald et al. (2003). *Journal of Personality and Social Psychology, 85*, 481-481.
- Hammond, K. R. (1948). Measuring attitudes by error-choice: an indirect method. *Journal of Abnormal & Social Psychology, 43*, 38-48.
- Hepperlen, T. M., Clay, D. L., Henly, G. A., & Barké, C. R. (2002). Measuring teacher attitudes and expectations toward students with ADHD: Development of the test of knowledge about ADHD (KADD). *Journal of Attention Disorders, 5*, 133-142.

- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179-185.
- Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion; psychological studies of opinion change*. New Haven, CT, US: Yale University Press.
- Jacobs, J. C. (1972). Teacher attitude toward gifted children. *Gifted Child Quarterly*, *16*, 23-26.
- Jones, E. E. & Sigall, H. (1971). The bogus pipeline: A new paradigm for measuring affect and attitude. *Psychological Bulletin*, *76*, 349-364.
- John, O. P., & Benet-Martinez, V. (2000). Measurement: Reliability, construct, and scale construction. In H. T. Reis, & C. M. Judd, (Eds.), *Handbook of research methods in social and personality psychology*, (pp. 339-369). New York: Cambridge University Press.
- Kavanaugh, R. D. (1976). On the synonymy of more and less: Comments on a methodology. *Child Development*, *47*, 885-887.
- Kidder, L. H., & Campbell, D. T. (1970). The indirect testing of social attitudes. In G.F. Summers (Ed.), *Attitude Measurement* (pp. 333-385). Chicago: Rand McNally & Co.
- Kubany, A. J. (1953). A validation study of the error-choice technique using attitudes on national health insurance. *Educational and Psychological Measurement*, *13*, 157-163

- Lane, A. L., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the implicit association test: IV What we know (so far) about the method. In B. Wittenbrink, & N. Schwarz, (Eds.). *Implicit measures of attitudes* (pp. 59-102). New York: Guilford Press.
- Loebowitz-Lennard, H., & Riessman, F., Jr. (1946). A proposed projective attitude test. *Psychiatry: Journal for the Study of Interpersonal Processes, 19*, 67-68.
- Lemon, N. (1973). *Attitudes and their measurement*. Oxford, England: John Wiley & Sons.
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149.
- Mascaro, G. F. (1970). Interpersonal attraction and uncertainty reduction as functions of judgmental similarity. *Perceptual and Motor Skills, 30*, 71-75.
- McNemar, Q. (1946). Opinion-attitude methodology. *Psychological Bulletin, 43*, 289-374.
- Milgram, S., Mann, L., & Harter., S. (1965). The lost-letter technique: A tool of social research. *Public Opinion Quarterly, 29*, 436-437.
- Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.
- Neely, J. H. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes. *Memory & Cognition, 4*, 648-654.
- Newcomb, T. M. (1940). Labor unions as seen by their members: An attempt to measure attitudes. In G. W. Hartmann, & T. M. Newcomb, (Eds). *Industrial conflict*, (pp. 313-338). New York: Cordon.

- Newcomb, T. M. (1946). The influence of attitude climate upon some determinants of information. *The Journal of Abnormal and Social Psychology, 41*, 291-302
- Nosek, B. A., & Banaji, M. R. (2001). The Go/No-go association task. *Social Cognition, 19*, 625-666.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers, 32*, 396-402.
- Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extrapersonal associations on the implicit association test: Personalizing the IAT. *Journal of Personality and Social Psychology, 86*, 653-667.
- Paulhus, D. L. (1991). Measurement and control of response bias. In Robinson, J. P., Shaver, P. R., Wrightsman, L. S. (ed.). *Measures of personality and social psychological attitudes, Vol. 1* (pp. 17-51). San Diego, CA: Academic Press, Inc.
- Petty, R. E., Fazio, R. H., & Brinol, P. (2009). The new implicit measures: An overview. In Petty, R.E., Fazio, R.H., & Brinol P. (Ed.), *Attitudes: Insights from new implicit measures*. New York, NY: Psychology Press.
- Proshansky, H. M. (1943). A projective method for the study of attitudes. *The Journal of Abnormal and Social Psychology, 38*, 393-395.
- Proshansky, H., & Murphy, G. (1942). The effects of reward and punishment on perception. *Journal of Psychology, 13*, 295-305.
- Riddleberger, A. B., & Motz, A. B. (1957). Prejudice and perception. *American Journal of Sociology, 62*, 498-503.

- Roese, N. J., & Jamieson, D. W. (1993). Twenty years of bogus pipeline research: A critical review and meta-analysis. *Psychological Bulletin*, *114*, 363-375.
- Ross, L., Greene, D., & House, P. (1977). The “false consensus effect”: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*, 279-301
- Smith, G. H. (1947). Beliefs in statements labelled Fact and Rumor. *Journal of Abnormal & Social Psychology*, *42*, 80-90.
- Tucker, L. R., Koopman, R. F., & Linn, R. L. (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika*, *34*, 421-459.
- Vargas, P. T., von Hippel, W., & Petty, R. E. (2004). Using partially structured attitude measures to enhance the attitude-behavior relationship. *Personality and Social Psychology Bulletin*, *30*, 197-211.
- Vaughan, G. M., & Thompson, R. H. T. (1961). New Zealand children's attitudes towards Maoris. *The Journal of Abnormal and Social Psychology*, *63*, 701-704.
- Webb, E. J., Campbell, D. T., Schwartz, R.D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences*. Oxford, England: Rand McNally.
- Weschler, I. R. (1950a). An investigation of attitudes toward labour and management by means of the error-choice method. *Journal of Social Psychology*, *32*, 51-62.
- Weschler, I. R. (1950b). A follow-up study on the measurement of attitudes toward labour and management by means of the error-choice method. *Journal of Social Psychology*, *32*, 63-69.

- Weschler, I. R. (1951). Problems in the use of indirect methods of attitude measurements. *The Public Opinion Quarterly*, *15*, 133-138.
- Wilde, G.J.S., & Fortuin, S. (1969). Self-report and Error-Choice; an application of the error-choice principle to the construction of personality test items. *British Journal of Psychology*, *60*, 101-108.
- Wilde, G.J.S., & de Wit, O. (1970). Self-report and Error-Choice; Individual differences in the operation of the error-choice principle and their validity in personality questionnaire items. *British Journal of Psychology*, *61*, 219-228.
- Wittenbrink, B. (2007). Measuring attitudes through priming. In B. Wittenbrink, & N. Schwarz, (Eds). *Implicit measures of attitudes* (pp. 17-58). New York: Guilford Press.

Appendix A

ERROR CHOICE Study: Ethics Forms

Consent Form, Information Sheet and Debriefing Form

ERROR CHOICE

Consent Form

Name (please print clearly):

1. I have read the information sheet and have had any questions answered to my satisfaction.
2. I understand that I will be asked to complete two categorisation tasks, and to share what I know about weight, eating habits, and exercising. I have been informed that the latter task consists of responding to two (2) questionnaires about weight, eating habits, and exercising, as well as a questionnaire about different issue.
3. I understand that my participation in this study is voluntary and that I may choose not to answer any questions and to terminate my participation at any time. I understand that such withdrawal will have no consequences of any sort.
4. I understand that every effort will be made to maintain the confidentiality of the data now and in the future.
5. I understand that individual results will not be made available, but I am entitled to a debriefing at the completion of the study in which the purpose and the expected outcomes of the study will be made available to me.

In the event that you have any complaints, concerns, or questions about this research, please feel free to contact either Ron Porter (533-6000, e75677, 3rdp@queensu.ca), Dr. Leandre R. Fabrigar (533-6492, fabrigar@queensu.ca), the Head of the Department of Psychology (533-2492), or the General Research Ethics Board (533-6081, chair.GREB@queensu.ca) at Queen's University.

I have read the above statements and freely consent to participate in this research:

Participant's Signature: _____ **Date:** _____

ERROR CHOICE

Information Sheet

This project is being conducted by Ron Porter, a graduate student in Psychology at Queen's University, under the supervision of Dr. Leandre R. Fabrigar, a professor in the Department of Psychology at Queen's University.

This study will entail finding out university students' opinions and reactions to a number of university projects. During this study you will be asked to complete several questionnaires. It is anticipated that this study, in conjunction with this session, will take approximately 50-60 minutes to complete. There are no known physical, psychological, economic, or social risks associated with this study. You will receive no additional payment for your participation in this study.

Your participation is completely voluntary and you are free to withdraw at any time with no effect. You are not obligated to answer any questions that you find objectionable or that make you feel uncomfortable.

Please also be assured that the data will be kept in a secure location, and that all of your responses will be coded to conceal your identity. Your responses will remain confidential and anonymous; only authorised researchers will have access to the data. The information provided will be used only for this project. Confidentiality will be maintained now and in the future, with no identification of individuals in publications resulting from this study. The raw data provided in this study will be kept in a locked room for which only experimenter will have access. When the data is no longer needed it will be destroyed.

Although individual results will not be published, and will not be made available to you, publications will consist of the collected data and you are entitled to a copy of the group findings and results.

In the event that you have any complaints, concerns or questions about this research, please feel free to contact either Ron Porter (533-6000, e75677, 3rdp@queensu.ca), Dr. Leandre R. Fabrigar (533-6492, fabrigar@queensu.ca), the Head of the Department of Psychology (533-2492), or the General Research Ethics Board (533-6081, chair.GREB@queensu.ca) at Queen's University.

ERROR CHOICE

Debriefing Sheet

A great deal of research has gone into studying the properties of peoples' attitudes and opinions. Of particular interest is how to measure these attitudes and opinions when the issues that are being assessed are of a sensitive nature. There are certain topics that many people are not forthcoming about when asked direct (explicit) questions (e.g., obesity, gender roles, racial stereotypes, etc.). Questionnaires that ask direct questions about these topics are called explicit measures. In these questionnaires, people can easily identify what they are being asked. Some researchers attempt to get information about people opinions or attitudes in a less direct way. When trying to assess a person's opinions on sensitive topics, it can be valuable to ask questions that are not obviously assessing these opinions. In fact, some past research has found that indirect (implicit) measures of attitude can provide a more accurate assessment of someone's attitude when the topic is of a sensitive nature.

One such implicit measure was developed by Hammond (1948), called the "error-choice technique". The basic assumption made by Hammond was that under a certain set of conditions, attitudes might bias how a person responds to a particular line of questioning. He proposed that an attitude could be assessed by presenting a person with an objective information test composed of a series of questions, each with a number of answer choices. The questions are designed such that the participant will be unlikely to know the correct answers, but their selected answers will imply positivity or negativity toward the target attribute. Thus, a forced answer would itself reflect a positive or negative attitude towards a topic. Further, because the person responding to the questions is unaware of the target attribute, error-choice should be less susceptible to social desirability than direct attitude measures, however this issue has not previously been addressed in research, nor has the issue of whether error-choice measures implicit (attitudes that cannot be consciously accessed) or explicit (attitudes that can be consciously accessed) attitudes.

When you arrived for this session, you were randomly assigned to a couple of experimental conditions. First, you were assigned to either an anonymous or public condition as part of a social desirability manipulation. In the anonymous condition, it was made clear to you that the researchers would in no way be able to identify who completed the measures. As a result, we hoped to decrease the likelihood that you would distort your responses. Conversely, in the public condition, a number of steps were taken to ensure that you felt completely identifiable in your response to the measures (although in fact no such identifications actually took place), thus increasing the likelihood of obscured responding. The second experimental condition related to social desirability (SD). In the low SD condition no attempt was made by the RA to draw your attention to the topic of interest, however in the high SD condition the RA provided a fabricated narrative about having had a weight control problem in the past. Although the RA has not had a weight problem, this deception was necessary to ensure that the topic of interest was made salient before completing the experimental measures.

You were then asked to complete four separate experimental measures. One of the first two measures (both of which you were told were tests of how people categorize), the test of categorizations was actually an implicit attitude measure (the IAT or PIAT). The other test, the test of general knowledge about overweight individuals, was an error-choice test. This deception was necessary in order to control for issues of social desirability. The second two tests consisted of an explicit attitude measure (the semantic differential test) and either an inventory of desirable responding (BIDR) or self-monitoring.

The purpose of this study is to examine the effectiveness of an implicit (i.e., indirect) measure of attitudes in assessing beliefs or opinions about obesity. It is expected that there will be a difference in responding for the explicit attitude measure based on whether a participant was in the anonymous or public condition. On the other hand, it is expected that reports on the error-choice test won't differ as a function of social desirability and that responses to this test will be uncorrelated with the social desirability condition.

Your participation in this study is greatly appreciated. In the event that you have any complaints, concerns, or questions about this research, please feel free to contact either Ron Porter (533-6000, e75677, 3rdp@queensu.ca), Dr. Leandre R. Fabrigar (533-6492, fabrigar@queensu.ca), Dr. Vernon Quinsey the Head of the Department of Psychology (533-2492), or the General Research Ethics Board (533-6081, chair.GREB@queensu.ca) at Queen's University. If you have any objections concerning your treatment, the methods or materials used, or the conduct of the experimenters during the course of this study, please inform the experimenter or anyone listed above. The data obtained from you will be destroyed, and you will still receive full credit for your participation.

Finally, it would be appreciated if you do not discuss this experiment with anyone else who may be eligible to participate (e.g., other introductory psychology students), as such prior knowledge might lead to unwanted effects.

If you are interested in learning more about the error-choice technique, you may wish to consult the following reference:

Hammond, Kenneth R. (1948). Measuring attitudes by error-choice: an indirect method. *Journal of Abnormal & Social Psychology*. Vol 43(1), Jan 1948, pp. 38-48

Appendix B

Social Desirability Scripts

*High and Low Social Desirability Scripts for Study 1, and
High and Low Social Desirability Scripts for Studies 2 and 3*

ERROR CHOICE –SCRIPTS for Study 1

SCRIPT – PRIVATE (Low Social Desirability Condition)

(Note: Participants are seated as far apart as possible to decrease social desirability)

Hi. The purpose of this study is to investigate how much people know about different categories. You will be asked to complete four separate measures about categorizations. Please fill out the inventories in the order you receive them. As the researcher, I don't want be able to identify who completed the measures so please do not put any identifying marks on your inventories. At the end of the study, please drop off your completed questionnaires in the box by the door as you leave, again so I don't know whose is whose.

SCRIPT – PUBLIC (High Social Desirability Condition)

Hi. The purpose of this study is to investigate how much people know about different categories. You will be asked to complete four separate measures about categorizations. Please fill out the inventories in the order you receive them. Also, please write your name, address, and phone number on the cover sheet of the questionnaire package so that I can check through your answers later to make sure that you answered the questions properly. When you're done, please hand me your completed inventories directly so I can recognize you later if I have any further questions.

ERROR CHOICE –SCRIPTS for Studies 2 and 3

SCRIPT – PRIVATE (Low Social Desirability Condition)

(Participants are seated as far apart as possible to decrease social desirability)

Hi everyone, my name is _____ and I'm working with Ron Porter on this study. Basically, for this part of the session today, you will be completing a computerized categorization task and a survey about eating habits, weight, and exercising.

Now, feel free to move the mouse to the continue button located in the bottom right of the computer monitor and begin the survey/categorization task. Please note that the participant number entered on the first screen has been randomly assigned and we cannot match your survey to your name or any other identifiable information. Also, when you have completed the survey you will see the word "Finished" written across the screen, you may then inform me that you have completed the survey. I will then explain the second task/survey you will be completing for the rest of the session. Again, when you have completed the survey you will see the word "Finished" written across the screen, you may then inform me that you have completed the survey.

SCRIPT – PUBLIC (High Social Desirability Condition)

Hi everyone, my name is _____ and I'm working with Ron Porter on this study. Basically, for the first half of the session today, we're looking at the issue of weight and obesity. As someone who has struggled with weight and who has been overweight in the past, I am very familiar with the stigmas overweight and obese people encounter. So this issue is very important to me.

Now, feel free to move the mouse to the continue button located in the bottom right of the computer monitor and begin the survey/categorization task. When you have completed the survey you will see a comment box for to enter your contact information (name, email address and/or phone number). Once this information has been entered please let me know so that if we need to contact you in the future we can match your name to your face. I will then enter a password and explain the second task/survey you will be completing for the rest of the session. When you have completed this task/survey you will see a comment box for to enter your name (in order to ensure that data files match) you may then inform me that you have completed the survey.

Appendix C

ERROR CHOICE Study 1: Experimental Materials

Error Choice Measure, Semantic Differential Scale and Self-Monitoring Measure

Study: CHOICE

Researcher: Ron Porter

In this session, you will be presented with a General Test of Knowledge about Health Related Behaviours and Nutrition. We are interested in how much the general population knows about this subject. Although there is no time limit you should try and complete the test as quickly as possible. Please wait until the research assistant tells you to begin.

Participant Identification Code: _____

General Test of Knowledge about Eating Behaviours

Please respond to the following set of questions by circling the answer choice you believe to be true. Please answer all questions to the best of your ability.

- 1) According to current scientific research, on average, what percent of overweight adult individuals are university educated?
a) 20% b) 45% c) 75% d) 90%
- 2) Based on findings by Health Canada, which province has the largest population of overweight individuals?
a) Nova Scotia b) Ontario c) Alberta d) Manitoba
- 3) Based on research by the Canadian Institute of Health, out of 100 overweight people, how many selected low fat food alternatives when provided the choice?
a) 10 b) 47 c) 72 d) 93
- 4) According to a recent survey, what percent of children under the age of 12 are overweight?
a) 15% b) 5% c) 50% d) 65%
- 5) According to a recent study conducted by Cornell University in which the performance of overweight managers was compared to the performance of managers who were not overweight, what was the difference observed?
a) no difference b) overweight managers performed poorer
c) overweight managers performed better
- 6) According to research conducted by the health institution at the University of Toronto, what percent of the Canadian population experiences health problems due to being overweight?
a) 8% b) 18% c) 36% d) 46%
- 7) What percent of overweight individuals would technically be classified as obese?
a) 15% b) 40% c) 65% d) 80%
- 8) Based on research by the American Health Institute, how many overweight people out of 100 exercise on average 3 or more times per week?
a) 12 b) 28 c) 41 d) 65
- 9) According to a study conducted by the government of Ontario, which city had the highest number of overweight individuals?
a) Ottawa b) Toronto c) Waterloo d) Kingston

- 10) According to recent research, what percent of body fat is genetically determined?
a) 10% b) 35% c) 52% d) 70%
- 11) According to current research, to what extent are overweight individuals actively engaging in diet and exercise to control their behaviour?
a) 20% b) 40% c) 60% d) 80%
- 12) Research conducted at the University of British Columbia suggested which of the following characteristics to be *most* prominent among a population of overweight adults?
a) lack of motivation b) poor leadership c) industrious d) friendly
- 13) According to a recent report on eating disordered behaviour, what percent of overweight individuals engage in *binge* eating (eating larger amounts of food at one time than is considered normal)?
a) 10% b) 22% c) 68% d) 84%
- 14) According to a recent report on eating disordered behaviour, what percent of overweight individuals do not exert control over when and the amount of food that they eat?
a) 21% b) 35% c) 57% d) 71%
- 15) According to Health Canada, what percent of overweight individuals work minimum wage jobs or are unemployed?
a) 8% b) 29% c) 53% d) 76%
- 16) Recent studies would indicate what percent of overweight people quit a new diet within 2 weeks?
a) 16% b) 37% c) 69% d) 88%
- 17) What proportion of overweight people would select taking an elevator above stairs?
a) 38% b) 52% c) 78% d) 89%
- 18) Based on research by the Canadian Institute of Health, what percent of overweight individuals have one or more family members who are also overweight?
a) 10% b) 40% c) 75% d) 90%
- 19) Research by the health institute at the University of Michigan found what percent of the population would classify a person weighing 300lbs as obese?
a) 50% b) 68% c) 89% d) 100%

- 20) According to a recent report on eating habits, what percent of overweight people eat more than 3 meals per day?
 a) 16% b) 29% c) 41% d) 77%
- 21) Based on research by the Health Institute of Ontario, which job has the highest rate of overweight employees?
 a) doctor b) retail clerk c) custodian d) lawyer
- 22) According to current research, what percent of overweight people are likely to choose a *regular* product rather than a *low fat* alternative?
 a) 22% b) 36% c) 49% d) 66%
- 23) Scientific evidence suggest that what percent of body fat would a person have to have to be considered obese?
 a) 25% b) 33% c) 42% d) 54%
- 24) What percent of Canadian youth grow up eating chocolate bars according to Health Canada?
 a) 9% b) 25% c) 45% d) 80%
- 25) Given current research on effective means to weight loss, to what extent can overweight people lose weight successfully and keep the weight off for an extended period of time?
 a) 3% b) 8% c) 15% d) 22%
- 26) On average, how many times does an overweight adult try to start a new diet each year?
 a) 0 b) 1 c) 2 d) 3
- 27) According to a recent study, what percent of overweight people eat fast food on a regular basis?
 a) 28% b) 40% c) 65% d) 80%
- 28) On average, how many servings does an overweight person have per meal?
 a) 1 b) 2 c) 3 d) 4
- 29) On average, how long does it take an overweight individual to eat a meal?
 a) 5min b) 13min c) 18min d) 25min
- 30) Based on research by Health Canada, how many servings of fruits and vegetables, on average, does an overweight individual eat per week?
 a) 0 b) 3 c) 8 d) 15
- 31) What percent of overweight individuals drink diet soft drinks?
 a) 8% b) 14% c) 25% d) 36%

- 32) According to the U.S. Department of Health and Human Welfare, how many high fat snack foods does the average overweight individual eat per week?
a) 2 b) 4 c) 7 d) 10
- 33) According to research by the medical school at Queen's University, what percent of overweight people will become diabetic in their lifetime?
a) 35% b) 55% c) 75% d) 88%
- 34) According to research by the medical school at Queen's University, What percent of overweight individuals will have a heart attack in their lifetime?
a) 42% b) 55% c) 68% d) 77%

CHOICE - Semantic Differential Scale

Please rate how you feel about **overweight** individuals on each of the scales below.

bad	-3	-2	-1	0	1	2	3	good
unfavourable	-3	-2	-1	0	1	2	3	favorable
unpleasant	-3	-2	-1	0	1	2	3	pleasant
negative	-3	-2	-1	0	1	2	3	positive
dislike	-3	-2	-1	0	1	2	3	like
disgust	-3	-2	-1	0	1	2	3	delight
repulsion	-3	-2	-1	0	1	2	3	attraction
disrespect	-3	-2	-1	0	1	2	3	respect

Self-Monitoring Scale (Snyder, 1974)

For each statement, answer by circling (T)rue or (F)alse.

- | | | |
|------------------------------------------------------------------------------------------------------------|---|---|
| 1. I find it hard to imitate the behavior of other people | T | F |
| 2. At parties and social gatherings, I do not attempt to do or say things that others will like. | T | F |
| 3. I can only argue for ideas which I already believe. | T | F |
| 4. I can make impromptu speeches even on topics about which I have almost no information. | T | F |
| 5. I guess I put on a show to impress or entertain others. | T | F |
| 6. I would probably make a good actor. | T | F |
| 7. In a group of people I am rarely the center of attention. | T | F |
| 8. In different situations and with different people, I often act like very different persons. | T | F |
| 9. I am not particularly good at making other people like me. | T | F |
| 10. I'm not always the person I appear to be. | T | F |
| 11. I would not change my opinions (or the way I do things) in order to please someone or win their favor. | T | F |
| 12. I have considered being an entertainer. | T | F |
| 13. I have never been good at games like charades or improvisational acting. | T | F |
| 14. I have trouble changing my behavior to suit different people and different situations. | T | F |
| 15. At a party I let others keep the jokes and stories going. | T | F |
| 16. I feel a bit awkward in company and do not show up quite as I should. | T | F |
| 17. I can look anyone in the eye and tell a lie with a straight face (if for a right end). | T | F |
| 18. I may deceive people by being friendly when I really dislike them. | T | F |

Appendix D

ERROR CHOICE Study 1: Correlation Tables

*Error Choice Measure (EC), Semantic Differential Scale (Direct) and
Self-Monitoring Measure(S-M Scale)*

Study 1.

Bivariate Correlations Between Dependent Variables

Measure	1	2	3
1. EC	--	.30**	.05
2. Direct		--	.03
3. S-M Scale			--

Note. **p < .01 (2-tailed).

Appendix E

ERROR CHOICE Study 2: Experimental Materials

Error Choice Measure, Semantic Differential Scale, Implicit Association Test, Balanced Inventory of Desirable Responding (Version 6) and Self-Monitoring Measure

The experimental materials for study 2 were presented to participants on a computer using the Medialab© program, and did not appear in this format, which has been altered to adhere with the School of Graduate Studies and Research guidelines for these submissions.

General Test of Knowledge about Health Related Behaviors and Nutrition

- 1) An individual is considered **overweight** when their Body Mass Index (BMI) is:
 - a. Below 18.5
 - b. 18.5-24.9
 - c. 25.0 –29.9
 - d. 30 and above

- 2) Based on findings by Health Canada, which province has the largest population of overweight individuals?
 - a. Nova Scotia
 - b. Ontario
 - c. Alberta
 - d. Manitoba

- 3) According to current research what percentage of overweight people are likely to choose a *regular* product rather than a *low fat* alternative?
 - a. 22%
 - b. 36%
 - c. 49%
 - d. 66%

- 4) What percentage of Canadians with an obese family member report being embarrassed by this person's weight?
 - a. 16%
 - b. 38%
 - c. 61%
 - d. 74%

- 5) According to the Canadian Food Guide, how many servings of meat and alternatives should an adult have each day?
 - a. 1-2
 - b. 2-3
 - c. 3-4
 - d. 4-5

- 6) What is the leading cause of obesity in North America?
 - a. Fad dieting
 - b. Overeating
 - c. Lack of physical activity
 - d. Genetics

- 7) From the following, which activity burns the most calories per hour?
- Walking
 - Running
 - Biking
 - Swimming
- 8) The average Canadian eats fast food ____ times a month.
- 2
 - 4
 - 8
 - 12
- 9) According to a poll conducted by the National Association of Broadcasters (NAB), how many people reported that they would be inclined to change the channel if they saw images of overweight individuals on the station they were watching?
- 11%
 - 32%
 - 56%
 - 81%
- 10) Based on research by the World Health Organization, how much money do Americans annually spend on weight-loss products and services?
- \$5 billion
 - \$15 billion
 - \$40 billion
 - \$60 billion
- 11) According to the World Health Survey (2002), what percent of individuals said that they would be uncomfortable having dinner and socializing with someone who was overweight?
- 72%
 - 58%
 - 35%
 - 17%
- 12) In 2000, which age group had the highest rate of obesity in Canada?
- 30-39
 - 40-49
 - 50-59
 - 60-69
- 13) Based on research done at the University of Toronto, what percentage of students said that they would be uncomfortable being at a beach with people who are overweight?
- 16%
 - 33%
 - 50%
 - 74%

- 14) Research by the Health Institute at the University of Michigan found what percentage of the population would classify a person weighting 300lbs as obese?
- a. 50%
 - b. 68%
 - c. 89%
 - d. 100%
- 15) According to a recent study conducted by the government of Ontario, which city had the highest number of overweight individuals?
- a. Ottawa
 - b. Toronto
 - c. Waterloo
 - d. Kingston
- 16) The National Institutes of Health spends ____% of its annual budget on obesity research.
- a. 0.5%
 - b. 1%
 - c. 2.5%
 - d. 5%
- 17) Based on a recent survey, what percentage of students said that overweight people “repulsed” them?
- a. 14%
 - b. 38%
 - c. 53%
 - d. 77%
- 18) According to the Health Institute of Ontario, what percentage of body fat would a person have to have to be considered obese?
- a. 23%
 - b. 27%
 - c. 32%
 - d. 36%
- 19) According to the U.S. Department of Health and Human Welfare, how many high fat snack foods does the average overweight individual eat per week?
- a. 10
 - b. 7
 - c. 4
 - d. 2
- 20) Which of the following desserts is the most fattening?
- a. Chocolate
 - b. Frozen yogurt
 - c. Oatmeal cookie
 - d. Fruit cup

- 21) According to the American Obesity Association, approximately _____ deaths are caused by obesity each year.
- 150,000
 - 225,000
 - 300,000
 - 375,000
- 22) According to a recent study, what is the ratio of overweight men to overweight women in Canada?
- 1 to 2
 - 1 to 3
 - 2 to 1
 - equal
- 23) Based on research by Health Canada, how many servings of fruits and vegetables, on average, does an overweight individual eat per week?
- 0
 - 3
 - 8
 - 15
- 24) According to a poll done by the International Association for the Study of Obesity in 2006, what percentage of people reported disgust when they watched an obese person eating at a fast-food restaurant?
- 32%
 - 47%
 - 63%
 - 81%
- 25) What percent of Canadian youth grow up eating chocolate bars according to Health Canada?
- 9%
 - 25%
 - 45%
 - 80%
- 26) According to research by the Health Department at McGill University in 2001, how many students reported being “disgusted” by obese people?
- 78%
 - 53%
 - 32%
 - 18%

- 27) Based on research by the Health Institute of Ontario, which job has the highest rate of overweight employees.
- Custodian
 - Retail clerk
 - Lawyer
 - Doctor
- 28) According to the Canadian Institute of Disordered Eating (CIDE), how many Canadians suffer from disordered eating?
- 0.5%
 - 2.5%
 - 4.0%
 - 8.0%
- 29) Recent research conducted by the Psychology department of the University of Calgary found that ___ percent of people are repulsed by the idea of a romantic involvement with someone who is obese.
- 10%
 - 26%
 - 42%
 - 69%
- 30) Children who are heavy-set are _____ than children who are not heavy-set to become obese in their teenage years
- Less likely
 - No more likely
 - Two times more likely
 - Three times more likely
- 31) How many calories does the average adult male need per day?
- 900-2000 kcal
 - 1250-3000 kcal
 - 1500-4000 kcal
 - 2200-5000 kcal

Please quietly inform the researcher that you have completed the timed portion of this study and continue on to the next section.

CHOICE - Semantic Differential Scale

Please rate how you feel about **overweight** individuals on each of the scales below.

bad	-3	-2	-1	0	1	2	3	good
unfavourable	-3	-2	-1	0	1	2	3	favorable
unpleasant	-3	-2	-1	0	1	2	3	pleasant
negative	-3	-2	-1	0	1	2	3	positive
dislike	-3	-2	-1	0	1	2	3	like
disgust	-3	-2	-1	0	1	2	3	delight
repulsion	-3	-2	-1	0	1	2	3	attraction
disrespect	-3	-2	-1	0	1	2	3	respect

Implicit Association Test Material

Initial IAT Instructions

In the next task you will be asked to make a series of judgments as quickly as you can. The judgments themselves are easy, but the difficult part is making them quickly. You will be pressing the “A” key and the “6” key (on the number pad) in order to make your judgments. In order to help you make these judgments as fast as you can, please keep your index fingers on the “A” and “6” keys throughout the task. This will help you respond more quickly, as you won’t have to move your hand once you have made your decision.

Please practice by pressing the “A” key now.

Key Assignment and Incorrect Entry Instructions

Be sure to notice the concept names at the top of the screen.

Press ‘A’ FAST for words described by the left concept name.
Press ‘6’ FAST for words described by the right concept name.

A red “X” will appear when you press the wrong key.
Whenever “X” appears press the other key immediately.

Some mistakes are OK.

Press the spacebar to proceed. GO FAST.

Final Instructions

Remember, that once you start making your judgments, it is important to go as fast as you can. You may make a few errors, but that’s OK. This particular test only works if you are going quickly enough that you might make a mistake.

If you have any questions, please see the RA.
Otherwise, press the spacebar to continue.

Implicit Association Test Category Stimulus

Slim

lean
thin
slim
skinny
slender
trim
small
average
in-shape
fit
light
buff

Overweight

chubby
hefty
heavy
thick
obese
pudgy
large
heavy-set
huge
overweight
plumb
fat

Good

favourable
sunshine
pleasant
rainbow
fortunate
positive
desirable
lovely
bright
heart
happy
good

Bad

harmful
ill
rotten
nasty
spoiled
awful
evil
sinful
unpleasant
offensive
negative
death

Self-Monitoring Scale (Snyder, 1974)

For each statement, answer by circling (T) rue or (F) alse.

- | | | |
|------------------------------------------------------------------------------------------------------------|---|---|
| 1. I find it hard to imitate the behavior of other people | T | F |
| 2. At parties and social gatherings, I do not attempt to do or say things that others will like. | T | F |
| 3. I can only argue for ideas which I already believe. | T | F |
| 4. I can make impromptu speeches even on topics about which I have almost no information. | T | F |
| 5. I guess I put on a show to impress or entertain others. | T | F |
| 6. I would probably make a good actor. | T | F |
| 7. In a group of people I am rarely the center of attention. | T | F |
| 8. In different situations and with different people, I often act like very different persons. | T | F |
| 9. I am not particularly good at making other people like me. | T | F |
| 10. I'm not always the person I appear to be. | T | F |
| 11. I would not change my opinions (or the way I do things) in order to please someone or win their favor. | T | F |
| 12. I have considered being an entertainer. | T | F |
| 13. I have never been good at games like charades or improvisational acting. | T | F |
| 14. I have trouble changing my behavior to suit different people and different situations. | T | F |
| 15. At a party I let others keep the jokes and stories going. | T | F |
| 16. I feel a bit awkward in company and do not show up quite as I should. | T | F |
| 17. I can look anyone in the eye and tell a lie with a straight face (if for a right end). | T | F |
| 18. I may deceive people by being friendly when I really dislike them. | T | F |

Appendix F

ERROR CHOICE Study 2: Correlation Tables

*Error Choice Measure (EC), Semantic Differential Scale (Direct),
Implicit Association Test (IAT), Balanced Inventory of Desirable Responding (IM Scale)
and Self-Monitoring Measure (S-M Scale)*

Study 2

Correlations Between Dependent Variables

Measure	1	2	3	4	5
1. EC	--	.19**	.01	.04	.02
2. Direct		--	.18**	.14*	.06
3. S-M Scale			--	.26**	.02
4. BIDR				--	.08
5. IAT					--

Note. * $p < .05$. (2-tailed), ** $p < .01$ (2-tailed).

Appendix G

ERROR CHOICE Study 3 Experimental Materials

Error Choice Measure, Semantic Differential Scale, Personalized Implicit Association Test, Balanced Inventory of Desirable Responding (Version 6) and Self-Monitoring Measure

The experimental materials for study 3 were presented to participants on a computer using the Medialab© program, and did not appear in this format, which has been altered to adhere with the School of Graduate Studies and Research guidelines for these submissions.

General Test of Knowledge about Health Related Behaviors and Nutrition

- 1) An individual is considered **overweight** when their Body Mass Index (BMI) is:
 - a) Below 18.5
 - b) 18.5-24.9
 - c) 25.0 –29.9
 - d) 30 and above

- 2) Based on findings by Health Canada, which province has the largest population of overweight individuals?
 - a) Nova Scotia
 - b) Ontario
 - c) Alberta
 - d) Manitoba

- 3) According to current research what percentage of overweight people are likely to choose a *regular* product rather than a *low fat* alternative?
 - a) 22%
 - b) 36%
 - c) 49%
 - d) 66%

- 4) What percentage of Canadians with an obese family member report being embarrassed by this person's weight?
 - a) 16%
 - b) 38%
 - c) 61%
 - d) 74%

- 5) According to the Canadian Food Guide, how many servings of meat and alternatives should an adult have each day?
 - a) 1-2
 - b) 2-3
 - c) 3-4
 - d) 4-5

- 6) What is the leading cause of obesity in North America?
 - a) Fad dieting
 - b) Overeating
 - c) Lack of physical activity
 - d) Genetics

- 7) From the following, which activity burns the most calories per hour?
- a) Walking
 - b) Running
 - c) Biking
 - d) Swimming
- 8) The average Canadian eats fast food ____ times a month.
- a) 2
 - b) 4
 - c) 8
 - d) 12
- 9) According to a poll conducted by the National Association of Broadcasters (NAB), how many people reported that they would be inclined to change the channel if they saw images of overweight individuals on the station they were watching?
- a) 11%
 - b) 32%
 - c) 56%
 - d) 81%
- 10) Based on research by the World Health Organization, how much money do Americans annually spend on weight-loss products and services?
- a) \$5 billion
 - b) \$15 billion
 - c) \$40 billion
 - d) \$60 billion
- 11) According to the World Health Survey (2002), what percent of individuals said that they would be uncomfortable having dinner and socializing with someone who was overweight?
- a) 72%
 - b) 58%
 - c) 35%
 - d) 17%
- 12) In 2000, which age group had the highest rate of obesity in Canada?
- a) 30-39
 - b) 40-49
 - c) 50-59
 - d) 60-69
- 13) Based on research done at the University of Toronto, what percentage of students said that they would be uncomfortable being at a beach with people who are overweight?
- a) 16%
 - b) 33%
 - c) 50%
 - d) 74%

- 14) Research by the Health Institute at the University of Michigan found what percentage of the population would classify a person weighting 300lbs as obese?
- a) 50%
 - b) 68%
 - c) 89%
 - d) 100%
- 15) According to a recent study conducted by the government of Ontario, which city had the highest number of overweight individuals?
- a) Ottawa
 - b) Toronto
 - c) Waterloo
 - d) Kingston
- 16) The National Institutes of Health spends ___% of its annual budget on obesity research.
- a) 0.5%
 - b) 1%
 - c) 2.5%
 - d) 5%
- 17) Based on a recent survey, what percentage of students said that overweight people “repulsed” them?
- a) 14%
 - b) 38%
 - c) 53%
 - d) 77%
- 18) According to the Health Institute of Ontario, what percentage of body fat would a person have to have to be considered obese?
- a) 23%
 - b) 27%
 - c) 32%
 - d) 36%
- 19) According to the U.S. Department of Health and Human Welfare, how many high fat snack foods does the average overweight individual eat per week?
- a) 10
 - b) 7
 - c) 4
 - d) 2
- 20) Which of the following desserts is the most fattening?
- a) Chocolate
 - b) Frozen yogurt
 - c) Oatmeal cookie
 - d) Fruit cup

- 21) According to the American Obesity Association, approximately _____ deaths are caused by obesity each year.
- a) 150,000
 - b) 225,000
 - c) 300,000
 - d) 375,000
- 22) According to a recent study, what is the ratio of overweight men to overweight women in Canada?
- a) 1 to 2
 - b) 1 to 3
 - c) 2 to 1
 - d) equal
- 23) Based on research by Health Canada, how many servings of fruits and vegetables, on average, does an overweight individual eat per week?
- a) 0
 - b) 3
 - c) 8
 - d) 15
- 24) According to a poll done by the International Association for the Study of Obesity in 2006, what percentage of people reported disgust when they watched an obese person eating at a fast-food restaurant?
- a) 32%
 - b) 47%
 - c) 63%
 - d) 81%
- 25) What percent of Canadian youth grow up eating chocolate bars according to Health Canada?
- a) 9%
 - b) 25%
 - c) 45%
 - d) 80%
- 26) According to research by the Health Department at McGill University in 2001, how many students reported being “disgusted” by obese people?
- a) 78%
 - b) 53%
 - c) 32%
 - d) 18%

- 27) Based on research by the Health Institute of Ontario, which job has the highest rate of overweight employees.
- a) Custodian
 - b) Retail clerk
 - c) Lawyer
 - d) Doctor
- 28) According to the Canadian Institute of Disordered Eating (CIDE), how many Canadians suffer from disordered eating?
- a) 0.5%
 - b) 2.5%
 - c) 4.0%
 - d) 8.0%
- 29) Recent research conducted by the Psychology department of the University of Calgary found that ___ percent of people are repulsed by the idea of a romantic involvement with someone who is obese.
- a) 10%
 - b) 26%
 - c) 42%
 - d) 69%
- 30) Children who are heavy-set are _____ than children who are not heavy-set to become obese in their teenage years
- a) Less likely
 - b) No more likely
 - c) Two times more likely
 - d) Three times more likely
- 31) How many calories does the average adult male need per day?
- a) 900-2000 kcal
 - b) 1250-3000 kcal
 - c) 1500-4000 kcal
 - d) 2200-5000 kcal

Please quietly inform the researcher that you have completed the timed portion of this study and continue on to the next section.

CHOICE - Semantic Differential Scale

Please rate how you feel about **overweight** individuals on each of the scales below.

bad	-3	-2	-1	0	1	2	3	good
unfavourable	-3	-2	-1	0	1	2	3	favorable
unpleasant	-3	-2	-1	0	1	2	3	pleasant
negative	-3	-2	-1	0	1	2	3	positive
dislike	-3	-2	-1	0	1	2	3	like
disgust	-3	-2	-1	0	1	2	3	delight
repulsion	-3	-2	-1	0	1	2	3	attraction
disrespect	-3	-2	-1	0	1	2	3	respect

Personalized Implicit Association Test Material

Initial PIAT Instructions

In the next task you will be asked to make a series of judgments. The judgments themselves are quite easy, however the challenging part is making them quickly.

You will be pressing the “A” key and the “6” key (on the number pad) in order to make your judgments.

In order to help you make these judgments as quickly as you can, please keep your index fingers on the “A” and “6” keys throughout the task. This will help you respond more quickly, as you will not have to move your hand once you have made your decision.

Key Assignment and Incorrect Entry Instructions

Be sure to notice the concept names that will appear at the top of the screen.

Press “A”, as fast as you can, for words that are associated with the concept on the **left** of the screen.

and

Press “6”, as fast as you can, for words that are associated with the concept on the **right** of the screen.

Final Instructions

Remember, once you start making your judgments it is important to go as fast as you can. You may make a few errors, but that’s okay. This particular test only works if you are going fast enough that you might make an occasional mistake.

If you have any questions, please ask the RA. Otherwise, press the space bar to continue.

Personalized Implicit Association Test Category Stimulus

Slim

lean
thin
slim
skinny
slender
trim
small
average
in-shape
fit
light
buff

Overweight

chubby
hefty
heavy
thick
obese
pudgy
large
heavy-set
huge
overweight
plumb
fat

I like / I don't like

Coffee	Disco
Spinach	Television
Tequila	Jogging
Cleaning House	Opera
Romance Novels	Garlic
Country Music	Motorcycles
Storms	Beer
Airplanes	Football
Caves	Shopping
Baseball	Cats
Science	Camping

Self-Monitoring Scale (Snyder, 1974)

For each statement, answer by circling (T) rue or (F) alse.

- | | | |
|------------------------------------------------------------------------------------------------------------|---|---|
| 1. I find it hard to imitate the behavior of other people | T | F |
| 2. At parties and social gatherings, I do not attempt to do or say things that others will like. | T | F |
| 3. I can only argue for ideas which I already believe. | T | F |
| 4. I can make impromptu speeches even on topics about which I have almost no information. | T | F |
| 5. I guess I put on a show to impress or entertain others. | T | F |
| 6. I would probably make a good actor. | T | F |
| 7. In a group of people I am rarely the center of attention. | T | F |
| 8. In different situations and with different people, I often act like very different persons. | T | F |
| 9. I am not particularly good at making other people like me. | T | F |
| 10. I'm not always the person I appear to be. | T | F |
| 11. I would not change my opinions (or the way I do things) in order to please someone or win their favor. | T | F |
| 12. I have considered being an entertainer. | T | F |
| 13. I have never been good at games like charades or improvisational acting. | T | F |
| 14. I have trouble changing my behavior to suit different people and different situations. | T | F |
| 15. At a party I let others keep the jokes and stories going. | T | F |
| 16. I feel a bit awkward in company and do not show up quite as I should. | T | F |
| 17. I can look anyone in the eye and tell a lie with a straight face (if for a right end). | T | F |
| 18. I may deceive people by being friendly when I really dislike them. | T | F |

Appendix H

ERROR CHOICE Study 3: Correlation Tables

*Error Choice Measure (EC), Semantic Differential Scale (Direct),
Personalized Implicit Association Test (PIAT), Balanced Inventory of Desirable
Responding (IM Scale) and Self-Monitoring Measure (S-M Scale)*

Study 3

Correlations Between Dependent Variables

Measure	1	2	3	4	5
1. EC	--	.30**	.04	.09	.08
2. Direct		--	.05	.07	.06
3. S-M Scale			--	.16**	.08
4. BIDR				--	.04
5. IAT					--

Note. *p < .05. (2-tailed), **p < .01 (2-tailed).