Use of Modified Paulhus Deception Scales to Compare Socially Desirable Responding Between a Forensic and Non-Forensic Sample

by

Colleen Robb

A thesis submitted to the School of Community Services
In partial fulfillment of the requirements for the Honours Bachelor of Behavioural Psychology.

St. Lawrence College
Kingston, Ontario
Canada
March 15th, 2018
Dedication

To you Papa- your faith and humour has inspired me in inconceivable ways.
I will take your lessons with me throughout life.
Abstract
The use of self-report measures in psychological testing is advantageous for collecting data on discrete or unobservable behaviours; however, there is a possibility that individuals will provide false or exaggerated responses to appear more favorable. This behaviour is referred to as socially desirable responding (SDR) and will irrefutably distort the results of an assessment. This research study sought to determine if individuals required to complete a court-ordered parental capacity assessment (PCA; the forensic group) would engage in more SDR than a comparative sample (the non-forensic group). PCAs use self-report measures as part of the overall assessment to assist the court in reaching decisions regarding an individual’s ability to provide proper care for a child. This research study recruited participants (N = 201) for both a forensic and non-forensic group. The Paulhus Deception Scales (PDS) and the L²P/F v.54 form were used to measure SDR and a two-way quasi-independent ANOVA design was employed to compare the amount of SDR between the two groups. The PDS is a tool used to measure the level of SDR in self-report measures and the L²P/F v.54 form was developed specifically for this research study. Results found that the forensic group engaged in more SDR than the non-forensic group across most of the constructs. Using correlational matrices, it was also determined that the L²P/F v.54 form demonstrated convergent validity with the PDS, thus indicating that the modified scale is similar to the PDS in measuring SDR. The results were consistent with the current literature and the proposed hypothesis which stated that the forensic group would engage in more SDR overall. Recommendations for further research include gathering more participants for the forensic group and modifying some questions of the L²P/F v.54 form to further enhance the convergent validity with the PDS.
Acknowledgements

First and foremost, I want to thank my family for their unwavering support. Support demonstrated not only during my undergraduate career, but for everything before, and everything that is to come.

To my friends who have supported me throughout life, I would not be who I am or where I am without you all. You are home to me. Wherever we go in life, it is incredible to know that we will always be by each other’s side. To my BPsyc family, thank you for your endless love, support, and laughter. I can truly say you have become the sisters that I never had and our lives will forever be interconnected in the most remarkable way.

To my professor and supervisor Dr. Robert Rowe, thank you all you have done for me throughout the development of my thesis. Because of you I have developed a new outlook on so many things and gained invaluable knowledge and experience. Your support, kindness, and humour will never be forgotten.

Finally, thank you to the professors and supervisors that have guided me throughout my undergraduate career. You have all shaped me academically, professionally, and personally.
Table of Contents

Dedication ......................................................................................................................... ii
Abstract ............................................................................................................................... iii
Acknowledgements ............................................................................................................ iv
Chapter I: Introduction ...................................................................................................... 1
  Psychological Assessments ............................................................................................... 1
  Parental Capacity Assessments ....................................................................................... 1
  Socially Desirable Responding ....................................................................................... 1
  Paulhus Deception Scale ............................................................................................... 2
  Conclusion ....................................................................................................................... 2
Chapter II: Literature Review ............................................................................................ 3
  Psychological Testing and Self-Report Measures ............................................................ 3
  Socially Desirable Responding ....................................................................................... 4
  Evaluating Socially Desirable Responding ..................................................................... 6
  Parental Capacity Assessments ....................................................................................... 7
Chapter III: Method ............................................................................................................ 10
  Participants ..................................................................................................................... 10
  Design ............................................................................................................................. 11
  Setting/Apparatus .......................................................................................................... 11
  Measures ......................................................................................................................... 11
  Procedures ....................................................................................................................... 12
Chapter IV: Results ............................................................................................................ 13
Chapter V: Discussion ......................................................................................................... 18
  Summary of Findings ...................................................................................................... 18
  Contributions to the Field .............................................................................................. 18
  Strengths and Limitations .............................................................................................. 19
  Study Challenges and Ethical Concerns ........................................................................ 19
  Recommendations for Future Research ....................................................................... 20
References .......................................................................................................................... 21
List of Tables

Table 1. Average of Socially Desirable Responses Provided by Forensic and Non-Forensic Groups Using Raw Scores..................................................13
Table 2. Average of Socially Desirable Responses Provided by Forensic and Non-Forensic Groups Using Extreme Scores........................................13
Table 3. Analysis of Variance for Total Raw Scores Between the Forensic and Non-Forensic Groups..............................................................14
Table 4. Analysis of Variance for Extreme Scores Between the Forensic and Non-Forensic Groups..............................................................14
Table 5. Correlational Matrix of Constructs Measured Using Total Raw Scores.............15
Table 6. Correlational Matrix for Constructs Using Extreme Scores..........................16
Table 7. Correlational Matrix of Constructs on the PDS and the L²P/F...........................17
Chapter I: Introduction

Psychological Assessments

Assessments and testing are prominent tools used in modern society to measure the abilities, skills, and specific needs of individuals (Rust & Golombok, 2009). Occupational performance assessments, practical testing for driving skills, and knowledge testing for school admittance are all common uses of testing and assessments; however, psychological testing and assessments are one of the most prominent and historical forms of this tool (Rust & Golombok, 2009). In fact, Watkins et al. report that 91% of psychologists use tests as regular practice (as cited in Groth-Marnat, 2003). By definition, a test is a single measure used to describe a behaviour in terms of categories and/or scores, which are then compared to norms or standards (Gregory, 2016). Alternatively, assessments refer to the total process of gathering information about an individual and using the information to forecast future behaviour (Gregory, 2016). In essence, a test is only one part of an overall assessment. Tests may be issued for numerous reasons, but the majority are used to make a decision (Gregory, 2016). Psychological tests for example can assess personality traits, intelligence, strengths, abilities, etc. These tests may also be used in part to identify a diagnosis or to provide supplementary information to a larger assessment. A parental capacity assessment (PCA), for example, is a commonly employed psychological assessment.

Parental Capacity Assessments

A PCA is a comprehensive evaluation of an individual’s parenting capacity, cognitive abilities, as well as risk and protective factors (Hotel Dieu Hospital, 2017). In Kington and surrounding areas, PCAs are ordered to the Family Court Clinic (FCC) through the Family Court system under section 54 of the Child and Family Services Act (Ontario Ministry of the Attorney General, 2016). Lead Psychologist, Dr. Robert Rowe, of the FCC conducts PCAs when the court requires supplementary information to reach a decision regarding child welfare cases (Hotel Dieu Hospital, 2017). PCAs consist of interviews with the parent(s), consultations with professionals involved in the matter, and a battery of psychological and cognitive tests (Hotel Dieu Hospital, 2017). The majority of the tests conducted for a PCA are self-report measures which are unfortunately susceptible to misrepresentation due to socially desirable responding (SDR; Tracey, 2016).

Socially Desirable Responding

SDR is defined as an individual’s propensity to highlight positive or desirable social traits while simultaneously abating negative or undesirable social traits (Watts et al., 2016). There are two forms of SDR: self-deceptive enhancement (SDE) and impression management (IM; Paulhus, 1998). SDE is the tendency to unconsciously provide enhanced descriptions of oneself (Paulhus, 1998). In contrast, IM is the tendency to deliberately provide exaggerated descriptions of oneself (Paulhus, 1998). Although SDR may be present for numerous reasons, IM has been shown to be more prominent when an individual has motive to appear favourably (Parmač Kovačić, Galić, & Jerneić, 2014). Stein et al., (2012) also suggest that a desire to be socially accepted and fear of repercussions are contributing factors to the use of SDR. As previously
noted, self-report measures are prone to erroneous information due to SDR; thus it is essential that a measure be employed to account for inaccurate responses (Tracey, 2016).

**Paulhus Deception Scale**

The Paulhus Deception Scale (PDS) is an established instrument used to quantify the level of SDR present in a self-report measure (Paulhus, 1998). The PDS has been found, through multiple studies, to be a valid and reliable measure of SDR (Paulhus, 1998). The PDS is a practical measure as it can be used with a variety of individuals—both clinical and non-clinical—and across multiple settings (Paulhus, 1998). Due to the possible incentive to appear favourable during a PCA, the FCC routinely provides measures to account for SDR. The L²P/F form is a modified version of the PDS that consists of additional questions that specifically target parental attitudes, attributes, and behaviours.

**Conclusion**

Due to the continuous use of psychological testing by professionals and the corresponding use of SDR in self-report measures, the study of SDR amongst individuals required to complete a PCA is considered to be a suitable topic of research. In conclusion, it is hypothesized that participants of the forensic group will engage in more SDR than participants of the non-forensic group.

This thesis consists of five chapters: introduction (see Chapter I), literature review (see Chapter II), method (see Chapter III), results (see Chapter IV), and discussion (see Chapter V). The first chapter, the introduction, provides an overview of PCAs, SDR, and the method for correctly measuring it. The literature review presents material from empirical studies, books, and other literature reviews to provide valid and reliable information to be used for the current study. The method chapter provides details on the participants and the research design used for the study, as well as information on the L²P/F form, which is the instrument used to measure SDR. This chapter also includes detailed information regarding the setting and apparatus, necessary materials, and procedures. The results chapter showcases the raw data collected from the study, a statistical analysis of the results, as well as tables and figures to provide a visual representation of the outcomes. Finally, the discussion chapter will provide an overview of conclusions related to the research question, as well as limitations, contributions to the field, and implications for future research.
Chapter II: Literature Review

Psychological Testing and Self-Report Measures

Psychological assessments can be used for a variety of reasons: classification, diagnosis, risk identification, evaluation, treatment planning, etc. (Gregory, 2016). However, the general purpose of these instruments is to generate a decision about an individual or situation (Gregory, 2016). According to Groth-Marnat (2003), it is estimated that practicing psychologists spend anywhere between 10% and 25% of their time conducting psychological assessments. Assessments generally consist of various tests and interviews from multiple sources to provide the clinician with well-rounded information regarding the individual and/or specific events (Gregory, 2016). A popular tool used amongst clinicians to gather this information is self-report measures. A self-report measure is any tool or measurement that relies on the participant to answer questions regarding their own behaviours, attributes, attitudes, etc. (Merriam-Webster’s Collegiate Dictionary, 2016). Self-report measures provide the unique advantage of gathering information that is generally discrete and not easily observable. Contrary to the strengths of this measurement, there are notable limitations; specifically, results may be distorted by the participant—either consciously or unconsciously—to enhance their chance of a favourable outcome (Parmač Kovačić, Galić, & Jerneić, 2014). Various studies have highlighted both the strengths and limitations of using self-report measures.

In a psychological assessment to evaluate men who were accused of domestic violence, researchers conducted a study to determine if results from self-report measures corresponded with results from semi-structured interviews and various scales (Helfritz, et al., 2006). These scales were specifically designed to measure impulsivity, aggression, and personality (Helfritz, et al., 2006). Eighty male participants who were involved with child custody cases completed complex psychological assessments to determine their risk of engaging in future domestic violence (Helfritz, et al., 2006). Results of the study showed that the majority of self-reported results did not match that of the interviews or scales (Helfritz, et al., 2006). These findings suggest that if self-report measures are used exclusively in assessments, the results may be misrepresented (Helfritz, et al., 2006).

In addition to their use in psychological assessments, self-report measures have been used across a variety of settings and have yielded various opinions due to conflicting results from multiple studies. For example, in a study conducted by Otten, Littenberg, and Harvey-Berino (2010), it was discovered that results from self-report measures were not congruent with the objective measures. Forty adult participants were enlisted for the study and asked to record the amount of time they spent watching television per day (Otten, et al., 2010). Devices were also placed inside the televisions of each participant to provide an objective measure of how often the television was being watched (Otten, et al., 2010). Upon completion of the study, Otten, et al. (2010) noted that 58% of the participants underreported their television viewing time by more than two hours. On average, the time spent watching television was underreported by 4.6 hours per week (Otten, et al., 2010). Participants of this study who had large discrepancies between the objective measure and self-report measure suggested that this was due to leaving the television on while completing other tasks and not actively watching (Otten, et al., 2010). In light of this possibility, the data was re-analyzed without the outliers and it was determined that the
average participant underestimated television viewing by 84 minutes per week (Otten, et al., 2010).

In contrast to this study, Quick et al., (2015) found that the results from self-report measures were highly correlated with the results from objective measures. Quick et al., (2015) recruited 1,686 university students in the United States to participate in a study comparing self-report responses and objective measures of individuals’ body mass index. Quick et al. (2015) asked the participants to provide researchers with their height and weight. Following this, researchers then objectively measured the height and weight of each participant (Quick, et al., 2015). The body mass index was then calculated for both the self-report measures and the objective measures gathered by the researchers. Results of the study found that 93% of the self-report responses were accurate when compared to the objective measures (Quick, et al., 2015). This statistic provides support for the use of self-report measures as a means of collecting and utilizing concrete data (Quick et al., 2015).

Evidently, research has highlighted both the positive and negative features of using self-report measures for testing and assessments; however, self-report measures are still a useful tool for gaining information about discrete or unobservable behaviours. One of the most prominent hindrances of self-report measures is the use of SDR amongst respondents.

**Socially Desirable Responding**

SDR is used to describe an individual’s tendency to promote socially desirable traits and to conceal socially undesirable traits. (Watts et al., 2016). Paulhus (1998) reports that there are two variations of SDR. The first version is IM and it refers to a conscious and purposeful attempt to enhance one’s appearance on self-report measures; the second version is SDE and it refers to the unconscious process of providing honest, yet somewhat inflated characteristics about oneself (Paulhus, 1998). Although occurring under different circumstances, both IM and SDE can still lead to inaccurate results. SDR has been found to occur throughout numerous settings and various populations; however, as previously noted, it is more common when an individual is motivated to achieve a certain outcome (Parmač Kovačić, Galić, & Jerneić, 2014).

A study completed by McEwan, Davis, Mackenzie, and Mullen (2009) discovered high levels of SDR amongst a group of men from a forensic mental health program. One-hundred and fifty-nine male participants were recruited from a forensic mental health service agency; specifically, each were referred to the agency due to stalking behaviours (McEwan, et al., 2009). The majority of the participants reported lower internal and external expressions of anger as well as higher levels of anger control (McEwan, et al., 2009). The reported decrease in expressions of anger and reported increase of anger control were positively correlated with high levels of SDR (McEwan, et al., 2009). The results of the study suggest that individuals were altering responses to appear more favourable to the researchers. Davis, Doherty, and Moser (2014) completed a similar study amongst a group of institutionalized men receiving a new form of addiction treatment. The comparison of pre-test and post-test results depicted a decrease in positive attitudes toward substance use, increased self-esteem, and improved use of coping skills amongst the participants (Davis, et al., 2014). However, the rate of these traits had a positive correlation with the rate of SDR for the same individuals (Davis, et al., 2014). Specifically, the mean score for IM increased from 4.77 at pretest to 6.72 at post-test (Davis, et al., 2014). These results
suggest that the participants engaged in SDR to give the impression that they were successful with the treatment program; the authors of the study propose that this alteration was done in an attempt to receive privileges (Davis, et al., 2014).

A recent study of a dissimilar population revealed comparable results of the aforementioned studies. Student athletes have been found to engage in SDR to mask psychological symptoms that they feel may hinder their future success in their sport (Gross, Wolanin, Pess, & Hong, 2017). Of the 244 participants in the study, 25% received scores of ten or more on the deception scale which indicated that SDR was present (Gross, et al., 2017). The results displayed a high mean SDR score of 7.56; according to Gross, et al. (2017), this score is comparable to a forensic group that had a mean SDR score of 7.61.

Further, Bornstein et al., (2015), discovered that SDR is also present across multiple countries. The study conducted by Bornstein, et al. (2015), enlisted mothers and fathers from 1110 families across nine different countries: China, Columbia, Italy, Jordan, Kenya, Sweden, Thailand, United States, and the Philippines (Bornstein, et al., 2015). All participants were required to complete a social desirability scale and The Parental Acceptance-Rejection/Control Questionnaire (Bornstein, et al., 2015). Results of the study indicated that there were no significant differences in SDR between the countries (Bornstein, et al., 2015). The study also discovered that mean levels of SDR between mothers and fathers did not differ (Bornstein, et al., 2015). That being said, some cultural differences did emerge in levels of SDR. Specifically, results from China indicated higher SDR in comparison to the mean scores of the study and results from Sweden indicated low SDR in comparison to the mean scores of the study (Bornstein, et al., 2015).

In addition to the various populations and settings in which SDR is present, it is also consistent across different mediums of self-report measures. Dodou and de Winter (2014) completed a meta-analysis of 51 studies comparing the level of SDR between online surveys and pencil-and-paper surveys. Of the 51 studies and 16, 700 participants analyzed, Dodou and de Winter (2014) discovered that the effect size of the mode of administration was nearly zero. The results of this study indicate that there is no difference in level of SDR between the two mediums; therefore, SDR should be expected in self-report measures regardless of their format (Dodou & de Winter, 2014).

Contradictory to the aforementioned research, there are individuals who propose that SDR is actually proof of highly virtuous behaviour. Zettler, Hilbig, Moshagen, and de Vries (2014) conducted two studies regarding this topic. The first study recruited 673 participants through a test and questionnaire website (Zettler, et al., 2014). This was done to ensure anonymity amongst respondents and to generate a low demand situation to decrease the possible desire to provide overly-positive responses (Zettler, et al., 2014). Participants completed a personality inventory and The Balanced Inventory of Desirable Responding scale (later released as the PDS; Zettler, et al., 2014). Results of this study indicated that high IM scores were positively correlated with three personality factors: honesty-humility, agreeableness, and conscientiousness (Zettler, et al., 2014). The correlation between these traits and high IM scores suggest that under low demand situations, high IM scores are representative of virtuous characteristics especially honesty (Zettler, et al., 2014). A second study conducted by the same
authors sought to discover if low or high IM scores were associated with dishonest behaviour. Zettler, et al. (2014) recruited 134 participants to complete The Balanced Inventory of Desirable Responding Scale (hereinafter referred to as the PDS) as well as a contrived dishonesty test. Dishonest behaviour was assessed by instructing participants to choose a side of a coin — heads or tails — and to toss the coin twice and record the results (Zettler, et al., 2014). Participants were informed that if both tosses resulted in their chosen side both times, they would receive an additional monetary compensation of five dollars (Zettler, et al., 2014). Participants were also informed that there was no possible way for researchers to determine if they were truthful or not (Zettler, et al., 2014). As it was impossible for researchers to determine if participants were being honest about the results, dishonest behaviour was based on the statistical baseline that the chance of obtaining two successful results out of two tosses was 25% (Zettler, et al., 2014). Results of the second study indicated that high IM scores were consistent with the baseline probability of winning the two coin tosses (Zettler, et al., 2014). Contrary, low IM scores were associated with exceptionally high probability of winning the two coin tosses (Zettler, et al., 2014). This statistic suggests that high IM scores are related to virtuous characteristics, especially honesty, when administered in low demand situations (Zettler, et al., 2014).

These studies have demonstrated how SDR has been observed across various populations and settings. Specifically, it has been noted that SDR is present globally in forensic settings, familial settings, and vocational settings to name only a few. This expansive display of SDR is noteworthy because it highlights the desire and ability to engage in deception regardless of one’s background, current situation, or demand conditions. Due to its prevalence, it is vital that an instrument be used to account for SDR so that assessments and tests can provide accurate results and interpretations.

**Evaluating Socially Desirable Responding**

The ability to account for SDR within self-report measures is paramount to properly interpreting results and gathering accurate information about an individual and/or a situation. Multiple attempts have been made to generate a test or scale that measures the level of SDR within self-report measures, however, a few have become staples within psychological assessments.

According to Paulhus (1998), the PDS has been recognized as both a valid and reliable measure of SDR. The PDS consists of 40 items equally split amongst the two subscales of SDE and IM (Paulhus, 1998). The decision to create a scale with two subscales came from a study conducted by Paulhus (1984). Prior to releasing the PDS, Paulhus (1984) employed an exploratory factor analysis and a confirmatory factor analysis and discovered that there was a marked distinction between SDE and IM. Braun, Jackson, and Wiley (2002) report that earlier research compliments these results through multiple studies that noted a distinct difference between SDE and IM. These results demonstrate support for utilizing both SDE and IM when analyzing SDR within self-report measures. Although the PDS may be considered the ‘gold standard’ for measuring SDR (Lambert, Arbuckle, and Holden, 2015), questions have been raised by researchers regarding if it preforms better than other conventional measures of SDR.

The Marlowe-Crowne Social Desirability Scale (MCSDS) and the PDS are both measures used regularly in psychological assessments to determine if individuals are providing
honest responses. Lambert, et al. (2015) recently conducted a study to determine if these scales were equally capable of identifying SDR. A total of 709 undergraduate students were selected to participate in one of three groups (Lambert, et al., 2015). All groups completed both scales but were provided with different instructions (Lambert, et al., 2015). The first group was instructed to respond as they normally would; the second group was instructed to answer in a way that would increase their chance of being hired –faking good—; finally, the third group was instructed to answer questions in a way that would decrease their chance of being hired –faking bad— (Lambert, et al., 2015). All participants who were instructed to alter their responses to achieve a desired outcome were informed that validity tests would be done and that they were to attempt to avoid providing clues that they were being deceptive (Lambert, et al., 2015). Results of the study discovered that overall, the MCSDS was more efficient than the PDS at detecting those who were faking good and faking bad (Lambert, et al., 2015). However, Lambert, et al. (2015) go on to state that although the MCSDS was more effective in detecting SDR amongst respondents, it is considered somewhat outdated; having been released more than 50 years ago, some of the items throughout the scale may not be applicable within present society.

A study conducted by Lanyon and Carle (2007) examined the latest revision of the PDS and supported the use of the scale for both a forensic and non-forensic population. The forensic group consisted of individuals who were involved with child court cases, criminal offenses, personal injury litigation, etc. (Lanyon & Carle, 2007). The non-forensic sample consisted of undergraduate students; the two groups resulted in a total of 519 participants (Lanyon & Carle, 2007). The participants completed a battery of psychological tests, including the PDS, to examine the relationship between the two subscales and to determine the validity of the scales (Lanyon & Carle, 2007). Specifically, Lanyon and Carle (2007) discovered that there was sufficient discriminant validity between SDE scores and IM scores. Both subscales were also considered to be valid measures of social desirability across both the forensic and non-forensic groups (Lanyon & Carle, 2007).

**Parental Capacity Assessments**

The term parental capacity is subjective and difficult to operationally define as the exact parameters are dependent on individual children. Overall, it refers to an individual’s ability to genuinely understand and prioritize the needs of their particular child (Donald & Jureidini, 2004). Furthermore, parents must be willing to recognize and alter their own inherent traits that may hinder further constructive development of their parental capacity (Donald & Jureidini, 2004). A study completed by Eve, Bryne, and Gagliardi (2014) provides a more detailed explanation of parental capacity based on both qualitative and quantitative data gathered from 50 professionals who regularly conduct PCAs. This study identified six main characteristics that are pertinent to ‘good’ parenting: insight, willingness and ability, day-to-day and complex needs, priority of child’s needs, fostering attachment, and consistency and flexibility (Eve, et al., 2014). Insight refers to in-depth knowledge of the child’s needs, strengths, and weaknesses, as well as the ability to recognize one’s own parenting limitations (Eve, et al., 2014). Willingness and ability refers to an individual’s willingness to be held accountable for and attempt to improve any problems that may arise with the child; this trait also pertains to an individual’s ability to parent by understanding and providing the basic needs to the child (Eve, et al., 2014). The study also suggests that parents should be able to meet both the day-to-day needs of the child and the long-term and more complex needs as well as the ability to prioritize these needs before their
own (Eve, et al., 2014). Finally, parents need to foster a secure attachment with their child and provide consistency in daily routines while exercising flexibility in terms of being receptive to the child’s ideas and interests (Eve, et al., 2014).

When there are concerns of an individual’s ability to meet the needs of the child, the Family Court system will intervene (Hotel Dieu Hospital, 2017). Under section 54 of the Child and Family Services Act, the courts may order a PCA to be completed by a professional in the field, for example, a clinical or forensic psychologist (Ontario Ministry of the Attorney General, 2016). The PCA consists of interviews with the family, consultations with other professionals, and a battery of both psychological and cognitive tests (Hotel Dieu Hospital, 2017). PCAs act as an unbiased assessment to aid the courts in the decision-making process regarding the parent’s ability to properly care for and provide basic needs to the child (Hotel Dieu Hospital, 2017). If the court reaches a decision that the individual is not capable of independently caring for a child, the child may become a crown ward, be placed in a foster home, or placed with another family member. Due to the potential consequences associated with the overall process, it is probable that individuals will attempt to present themselves positively to those conducting the assessments. A recent study conducted by Carr, Moretti, and Cue (2005) discovered that multiple tests (e.g. the Minnesota Multiphasic Personality Inventory) conducted during a PCA presented evidence of overly positive self-presentation. This evidence of individuals providing positively inflated characteristics of themselves was recognized on all four of the measures that were examined (Carr, et al., 2005).

Milner and Crouch (1997) provide further evidence regarding response distortion in child protection cases. Milner and Crouch (1997) gathered two groups of participants: general population parents and ‘at risk’ parents; The ‘at risk’ parents were those who were attending services for child abuse either through court-orders or referrals from professionals. Each group was instructed to provide either honest, fake good, fake bad, or random answers to each test depending on the condition. The Adult/Adolescent Parenting Inventory (AAPI), the Child Abuse Potential (CAP) Inventory, and the Parenting Stress Index (PSI) were used to evaluate each instructional condition. Although participation was voluntary and every individual was ensured that results would be confidential, the ‘at risk’ parent group still provided more SDR that that of the general population group during the honest condition (Milner & Crouch, 1997). In fact, 33% of the ‘at risk’ parenting group were identified as faking good for the measures, compared to 8% of the general population group (Milner & Crouch, 1997). These results demonstrate that although there was no apparent incentive to fake good on these instruments, those who were involved with child well-being cases still provided false and overly positive attributes, thus invalidating the results of the measures (Milner & Crouch, 1997).

PCAs are a prime example of the use of self-report measures in psychological assessments. Self-report measures are prominent in PCAs because many qualities and traits that are characteristic of suitable parents are not easily observed. In general, PCAs may elicit anxiety and nervousness from participants due to the possibility of not gaining access or custody of their child. Due to the high-demand situation created by PCAs, participants may be more prone to accentuate socially desirable characteristics, thus distorting the results (Parmač Kovačić, Galić, & Jernejić, 2014). Further, as noted by Milner and Crouch (1997), individuals involved with child welfare cases have shown to provide more SDR than a comparative norm group. This amplified
probability of distorted results highlights the necessity to have measurements that are able to account for SDR within self-report measures. As the PDS has demonstrated success in identifying social desirability amongst both forensic and non-forensic populations (Lanyon & Carle, 2007), it is considered a suitable tool to measure SDR within PCAs.
Chapter III: Method

Participants

Two groups of participants were recruited for the present study: a forensic group and a comparative non-forensic group. Two participants—one from each group—were removed due to incompletion of multiple questions on the L2P/F v.54 form (Appendix A); this resulted in a total of 201 participants. The forensic group accounted for 38.3% (N=77) of the total number of participants and the non-forensic group accounted for 61.7% (N=124) of the total number of participants. Females represented 50.6% and 78.2% of the forensic and non-forensic groups respectfully. Additionally, males represented 49.4% and 21.8% of the forensic and non-forensic groups respectfully. The average age for participants of the forensic group was 33 years and ranged in age between 16 and 75 while the participants of the non-forensic group had a mean age of 43 years and ranged in age between 26 and 66. During selection procedures, neither group was restricted by age or gender.

The database that was generated from study indicated that the two groups were closely related regarding the average number of biological children per parent; participants of the forensic group had an average of 2.53 biological children per individual while the non-forensic group had an average of 2.26 biological children per individual. The database further indicated that 81.8% of the forensic group and 28.2% of the non-forensic group had previously had contact with the Children’s Aid Society regarding at least one biological child.

The 77 forensic group participants were selected based on their court-ordered requirement to complete a PCA through the FCC with Dr. Rowe, during which they completed the L2P/F v.54 form and demographic form (Appendix B). Prior to beginning the PCA, participants were informed of the purpose of the assessment, the possible uses of the information gathered, and the limits of confidentiality. The participants then signed a consent form (Appendix C). The 124 non-forensic group participants were randomly selected from the Child and Adolescent Mental Health department waiting room at a hospital in Southeastern Ontario. Participants of this group were included if they were the parent of at least one biological child, but were excluded if they were currently or had previously been employed by the Children’s Aid Society. This exclusion criteria was employed in order to circumvent possible bias when answering questions pertaining to the Children’s Aid Society. Participants were approached by the researcher, provided information about the study, and asked if they would voluntarily complete a consent form (Appendix D), demographic sheet, and the L2P/F v.54 form for research purposes. The participants were informed that the information would be confidential, that there would be no impact on the services they were receiving, and that there were no foreseen consequences associated with completing the forms.

The L2P/F v.54 form, the demographic sheet, and the information/consent form were approved by The Queen’s University Health Sciences & Affiliated Teaching Hospitals Research Ethics Board (HSREB) on June 9, 2016. This approval granted ethical clearance to conduct research on the study until June 9, 2018.
Design

The researcher employed a two-way quasi-independent ANOVA design to compare the amount of SDR (dependent variable) on the L²P/F v.54 from between the forensic group and non-forensic group (independent variables). SDR was defined as an individual’s propensity to provide exaggerated positive traits about themselves while simultaneously minimizing negative traits (Watts, et al., 2016). The forensic group was defined as individuals who completed the L²P/F v.54 form as part of their court-ordered PCA. The non-forensic group was defined as individuals who were randomly selected from the Child and Adolescent Mental Health department waiting room at a hospital in Southeastern Ontario for means of comparison.

Descriptive statistics were used to summarize demographic information about both groups of participants (e.g. age, sex, number of children, etc.). Further, the L²P/F v.54 form was analyzed by the researcher in two different ways. First, the data were analyzed simply by summing the total raw scores for each construct. The researcher then scored the form using the extreme scoring technique. Similar to the original PDS, the level of SDR on the L²P/F v.54 form was calculated based on ‘extreme’ scores. Thus, certain questions were given a score of one if a response of ‘6’ or ‘7’ was provided and other questions were given a score of one if a response of ‘1’ or ‘2’ was provided; all other responses were scored as zero. The researcher also used SPSS to create multiple correlational matrixes of the data to determine the relationship between all constructs on the L²P/F v.54 form for both the total raw scores and the extreme scores. A third correlational matrix was also constructed to determine the relationship of the L²P/F v.54 constructs and the two constructs of the original PDS. These forms of analysis were chosen due to the large sample size and due to the number of items and constructs being measured. Visual representation of the results from the ANOVA and correlational matrixes are provided through both figures and tables.

Setting/Apparatus

The forensic group completed the forms in a private office of the FCC at the hospital. The office contained a table and three to four chairs. The non-forensic group completed the forms in the Child and Adolescent Mental Health department waiting room. This room contained approximately 12 chairs, various toys and books, as well as a large television which played movies throughout the day. A demographic sheet, the L²P/F v.54 form, and writing utensils were the materials necessary for completing the study. Other required materials were Tim Horton’s gift cards with a value of two dollars each – equivalent to a large coffee – that were used as compensation for the non-forensic group upon completion of the forms.

Analysis was completed in a private office within the FCC and a laptop equipped with SPSS was used to record and analyze the data from both groups. The researcher also used archival data for both the forensic and non-forensic groups. All information collected was contained in a locked cabinet and within a locked office at the FCC.

Measures

Both archival data and collected data were used to measure the dependent variable – the amount of SDR – between the forensic group and the non-forensic group. The demographic form
provided basic information about the participants and their biological children, and the \(L^2P/F\) v.54 form was the specific tool used to measure SDR amongst the participants. The \(L^2P/F\) v.54 form is a modified version of the PDS and is equipped with questions to specifically target parent behaviour, attitudes, and attributes. The PDS has shown to be useful in evaluating SDR amongst both forensic and non-forensic populations (Lanyon & Carle, 2007), thus a slightly adapted version of the instrument was employed for this study. The \(L^2P/F\) v.54 form uses a seven-point Likert scale which ranges from 1 \textit{(not true)} to 7 \textit{(very true)}.

**Procedures**

Data for the forensic group was collected from archival data and from individuals who completed the necessary forms while completing the PCA. Although some archival data was also available for the non-forensic group, a larger sample was needed. To gain more data, the researcher approached individuals in the waiting room of the Child and Adolescent Mental Health department and introduced herself as a researcher within the hospital. The researcher described the study and explained that any information provided would remain confidential and that it would have no bearing on the services they were receiving. The participants were then asked to sign the confidentiality sheet and complete a demographic form as well as the \(L^2P/F\) v.54 form; on average, this took each participant approximately ten minutes to complete. Upon completion of the forms, each participant was provided with a two-dollar Tim Hortons gift card (equivalent to the price of a large coffee). The completed forms were brought back to the office and all data were entered into the SPSS database.

Once a sufficient number of participants had completed the forms, the two databases were merged into one. It was determined that 14 of the questions did not closely correlate with any others and therefore were deleted from the analysis. The subsequent 30 questions fell into three new constructs: Impression management (IM), Criminal Impression Management (CIM), and Hostility Toward the Children’s Aid Society (HCAS). The researcher then ran frequency scans to ensure that variables from each database was coded in the same manner and to ensure no anomalies were present in the merged database. The researcher then adjusted the coding of certain demographic variables accordingly.

Next, the researcher scored both the total raw scores and the extreme scores for the three constructs. The total score was calculated by adding all raw scores within each construct (IM = questions 1-19; CIM= questions 20, 22, 24, 26, 28, & 30; HCAS = questions 21, 23, 25, 27, & 29). The extreme scores were calculated in the same fashion, however only scores that were considered ‘extreme’ were assigned a score of one and all others were scored as zero. The researcher then ran an ANOVA for the total extreme scores. Finally, the researcher created multiple correlational matrixes to analyze the constructs in two different ways. These correlational matrixes display the relationship between the three constructs of the \(L^2P/F\) v.54 for both the total raw scores and the extreme scores. Another correlational matrix displays the relationship between the three constructs of the \(L^2P/F\) v.54 and the two constructs of the original PDS.
Chapter IV: Results

The mean levels of SDR for both the total scores and the extreme scores indicate that the forensic group engaged in more SDR than the non-forensic group across a majority of the constructs. In particular, the forensic group had a mean total score of 78.87, 22.36, and 19.48 in the IM, CIM, and HCAS constructs respectively. In comparison, the non-forensic group had a mean total score of 74.75, 19.75, and 15.17 in the IM, CIM, and HCAS constructs respectively. The mean extreme SDR scores of both groups depicted a similar pattern. The forensic group had a mean extreme score of 10.51 and 1.69 on the IM and HCAS constructs respectively, while the non-forensic group had a mean extreme score of 7.39 an 0.54 on the same constructs. In contrast to these results, the non-forensic group had a higher mean extreme score (3.84) on the CIM construct than the forensic group (3.17). The average SDR for both groups using the total raw scores and the extreme scores can be seen in Table 1 and Table 2 respectfully.

Table 1

*Average of Socially Desirable Responses Provided by Forensic and Non-forensic Groups Using Raw Scores*

<table>
<thead>
<tr>
<th>Construct</th>
<th>Forensic Group</th>
<th>Non-forensic Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$M$</td>
</tr>
<tr>
<td>Total IM</td>
<td>77</td>
<td>78.87</td>
</tr>
<tr>
<td>Total CIM</td>
<td>77</td>
<td>22.36</td>
</tr>
<tr>
<td>Total HCAS</td>
<td>77</td>
<td>19.48</td>
</tr>
</tbody>
</table>

Table 2

*Average of Socially Desirable Responses Provided by Forensic and Non-forensic Groups Using Extreme Scores*

<table>
<thead>
<tr>
<th>Construct</th>
<th>Forensic Group</th>
<th>Non-forensic Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$M$</td>
</tr>
<tr>
<td>Extreme IM</td>
<td>77</td>
<td>10.51</td>
</tr>
<tr>
<td>Extreme CIM</td>
<td>77</td>
<td>3.17</td>
</tr>
<tr>
<td>Extreme HCAS</td>
<td>77</td>
<td>1.69</td>
</tr>
</tbody>
</table>
14

Table 3

Analysis of Variance for Total Raw Scores Between the Forensic and Non-Forensic Groups

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>TotalIM</td>
<td>Between Groups</td>
<td>2122.625</td>
<td>1</td>
<td>2122.625</td>
<td>17.719</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>23839.435</td>
<td>199</td>
<td>119.796</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>25962.060</td>
<td>200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TotalCIM</td>
<td>Between Groups</td>
<td>851.809</td>
<td>1</td>
<td>851.809</td>
<td>28.683</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>5909.754</td>
<td>199</td>
<td>29.697</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>6761.562</td>
<td>200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TotalHCAS</td>
<td>Between Groups</td>
<td>2314.685</td>
<td>1</td>
<td>2314.685</td>
<td>37.297</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>12350.221</td>
<td>199</td>
<td>62.061</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>14664.905</td>
<td>200</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results of the ANOVA for the three constructs of the L²P/F v.54 form that were calculated using the total raw scores are displayed in Table 3. Further, results of the ANOVA using the extreme scores are displayed in Table 4. The majority of the constructs were statistically significant between groups (p < 0.001). The only exception was the construct of CIM in the extreme scoring method which had a significance of 0.006 between groups. Although this slight difference was present, all constructs were still considered to be statistically significant. These results suggest that the relationship between these constructs was caused by a factor other than random chance; more specifically, the relationships between the constructs had a less than 0.01% chance of occurring arbitrarily.

Table 4

Analysis of Variance for Extreme Scores Between the Forensic and Non-Forensic Groups

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ExtremeIM</td>
<td>Between Groups</td>
<td>462.229</td>
<td>1</td>
<td>462.229</td>
<td>30.881</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>2978.666</td>
<td>199</td>
<td>14.968</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>3440.896</td>
<td>200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ExtremeCIM</td>
<td>Between Groups</td>
<td>21.316</td>
<td>1</td>
<td>21.316</td>
<td>7.804</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>543.579</td>
<td>199</td>
<td>2.732</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>564.896</td>
<td>200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ExtremeHCAS</td>
<td>Between Groups</td>
<td>62.603</td>
<td>1</td>
<td>62.603</td>
<td>36.287</td>
</tr>
<tr>
<td></td>
<td>Within Groups</td>
<td>343.318</td>
<td>199</td>
<td>1.725</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>405.920</td>
<td>200</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A correlational matrix was used to determine the relationship between constructs using both the total raw scores and the extreme scores; results of which are displayed in Table 5 and Table 6 respectively. Results of the correlational matrix indicate that most constructs were related at the 0.01 level (2-tailed), which suggests that these constructs were measuring similar material. Specifically, the total scores and the extreme scores produced correlations of 0.347 and 0.406 between the IM and CIM constructs respectively. Furthermore, the CIM and HCAS constructs had a correlation of 0.233 for the total scores and a correlation of -0.221 for the extreme scores. Of the three constructs, the only correlation that was not considered significant at the 0.01 level was the IM and HCAS correlation. The IM and HCAS constructs produced a correlation of 0.109 for the total scores and a correlation of 0.002 for the extreme scores.

Table 5

<table>
<thead>
<tr>
<th>Correlational Matrix of Constructs Measured Using Total Raw Scores</th>
<th>TotalIM</th>
<th>TotalCIM</th>
<th>TotalHCAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TotalIM</td>
<td>Pearson Correlation</td>
<td>1</td>
<td><strong>.347</strong></td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td></td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>201</td>
<td>201</td>
</tr>
<tr>
<td>TotalCIM</td>
<td>Pearson Correlation</td>
<td><strong>.347</strong></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>201</td>
<td>201</td>
</tr>
<tr>
<td>TotalHCAS</td>
<td>Pearson Correlation</td>
<td>.109</td>
<td><strong>.233</strong></td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.124</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>201</td>
<td>201</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
A correlational matrix was also used to determine the relationship between the three constructs of the \( L^2 P/F \) v.54 form and the two constructs of the PDS. As the PDS is always scored using extreme scores, only the extreme scores of the \( L^2 P/F \) v.54 form were used during this analysis. Results of this correlational matrix indicate that the majority of the constructs were highly related and were therefore measuring similar material. The results show that the IM construct of the \( L^2 P/F \) v.54 form was significantly correlated to both the IM and SDE construct of the PDS at the 0.01 level. Specifically, the IM construct of the \( L^2 P/F \) v.54 form had a correlation of 0.583 with the IM construct of the PDS and a correlation of 0.473 with the SDE construct of the PDS. The CIM construct of the \( L^2 P/F \) v.54 form had a significant correlation with the IM construct of the PDS (0.401), but not with the SDE construct (0.215). Lastly, the HCAS construct of the \( L^2 P/F \) v.54 form did not display a significant correlation with either PDS construct; the HCAS construct had a correlation of -0.156 with the PDS IM construct and a correlation of -0.030 with the PDS SDE construct. These results are displayed in Table 7.

Overall, the results of the correlational matrices indicate that all of the constructs within the \( L^2 P/F \) v.54 form –aside from one set –were closely related to one another. Due to the high correlation, this can be stated with a 99% confidence rate. Further, it can be stated with a 99% confidence rate that IM construct of the \( L^2 P/F \) v.54 form was closely related to both PDS constructs. Moreover, the CIM construct of the \( L^2 P/F \) v.54 form was highly related to the IM construct of the PDS, but not with the SDE construct. Finally, the HCAS construct of the \( L^2 P/F \) v.54 form was not closely related to either construct of the PDS.

**Table 6**

*Correlational Matrix for Constructs Using Extreme Scores*

<table>
<thead>
<tr>
<th></th>
<th>ExtremeIM</th>
<th>ExtremeCIM</th>
<th>ExtremeHCAS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pearson Correlation</strong></td>
<td>.406</td>
<td>.000</td>
<td>-.002</td>
</tr>
<tr>
<td><strong>Sig. (2-tailed)</strong></td>
<td>.000</td>
<td>.975</td>
<td>.002</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>201</td>
<td>201</td>
<td>201</td>
</tr>
<tr>
<td><strong>Pearson Correlation</strong></td>
<td>-.002</td>
<td><strong>-.221</strong></td>
<td>1</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
Table 7

**Correlational Matrix of Constructs on the PDS and the L²P/F**

<table>
<thead>
<tr>
<th></th>
<th>Extreme IM</th>
<th>Extreme CIM</th>
<th>Extreme HCAS</th>
<th>Paulhus SDE</th>
<th>Paulhus IM</th>
</tr>
</thead>
<tbody>
<tr>
<td>ExtremeIM</td>
<td>Pearson Correlation</td>
<td>1</td>
<td>-.406</td>
<td>-.02</td>
<td>.473</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.975</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>201</td>
<td>201</td>
<td>201</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Pearson Correlation</td>
<td><strong>.406</strong></td>
<td>1</td>
<td><strong>-.221</strong></td>
<td>.215</td>
</tr>
<tr>
<td>ExtremeCIM</td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.002</td>
<td>.085</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>201</td>
<td>201</td>
<td>201</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Pearson Correlation</td>
<td>-.002</td>
<td><strong>-.221</strong></td>
<td>1</td>
<td>-.030</td>
</tr>
<tr>
<td>ExtremeHCAS</td>
<td>Sig. (2-tailed)</td>
<td>.975</td>
<td>.002</td>
<td>.815</td>
<td>.215</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>201</td>
<td>201</td>
<td>201</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Pearson Correlation</td>
<td><strong>.473</strong></td>
<td>.215</td>
<td>-.030</td>
<td>1</td>
</tr>
<tr>
<td>PaulhusSDE</td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.085</td>
<td>.815</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Pearson Correlation</td>
<td><strong>.583</strong></td>
<td><strong>.401</strong></td>
<td><strong>-.156</strong></td>
<td><strong>.379</strong></td>
</tr>
<tr>
<td>PaulhusIM</td>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.001</td>
<td>.215</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>65</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
Chapter V: Discussion

Summary of Findings

The present study focused on the use of SDR within self-report assessments. Specifically, the aim of the study was to determine if individuals who were required to complete self-report measures for the purpose of a court-ordered PCA would engage in more SDR than a randomly selected comparative sample. The results of the study show that the forensic group displayed higher levels of SDR than the non-forensic group in all constructs except for one. When the extreme scoring method was used, the non-forensic group demonstrated higher levels of SDR on the CIM construct when compared to the forensic group. Although these higher levels of SDR amongst the non-forensic group may be indicative of SDR, this outcome could also be the result of highly virtuous behaviour; a theory which has been supported in two separate studies by Zettler, Hilbig, Moshagen, and de Vries (2014). Similarly, this theory could account for the higher levels of SDR amongst the forensic group; however, SDR is more likely present because this group understood that their responses would be related to their final report. Comparatively, the non-forensic group had no incentive to provide false or exaggerated answers. These results were consistent with the findings from Parmeč Kovačić, Galić, and Jerneić (2014) who claim that SDR occurs more often when an individual is motivated to achieve a specific outcome. The higher levels of SDR amongst the forensic group were also consistent with the findings of McEwan, Davis, Mackenzie, and Mullen (2009) and of Davis, Doherty, and Moser (2014); both studies examined forensic populations and noted that high levels of self-reported positive traits correlated with high levels of SDR, suggesting that these individuals were providing false or embellished answers to appear more favorable to the researchers. Overall, the results of this study support the hypothesis that the forensic group would engage in more SDR than the non-forensic group.

The results of this study were further evaluated through the use of ANOVAs and correlational matrices. The ANOVAs created for the L²P/F v.54 form produced statistically significant results across all constructs, indicating there was a marked difference between the forensic and non-forensic groups and that the relationship between constructs was not due to a random external factor. The correlational matrices demonstrated high correlations between all constructs of the L²P/F v.54 form except for one set (IM and HCAS constructs), which indicates that these constructs were measuring similar material and that participants were answering consistently. The L²P/F v.54 form also demonstrated convergent validity through positive correlation with the PDS constructs. Specifically, the IM construct of the L²P/F v.54 form was highly correlated with both constructs of the PDS. The CIM construct of the L²P/F v.54 form was also highly correlated with the IM construct of the PDS. These results indicate that the IM and CIM constructs of the L²P/F v.54 form identify SDR in the same fashion as the PDS. The HCAS construct of the L²P/F v.54 form had no significant correlation with the constructs of the PDS which suggests that these questions did not measure similar material; this result was anticipated as the PDS does not contain questions regarding attitudes toward the CAS.

Contributions to the Field

This study contributed to the field of behavioural psychology by providing new and evidence-based research on the topic of SDR engagement in self-report measures. Self-report measures are a prominent feature in psychological assessments, which are regularly used in this
field; therefore, the study of SDR and the study of the validity of these measures is essential in order to ensure that accurate and representative information is collected during assessments.

**Strengths and Limitations**

A marked strength of this study was the availability of archival data for both the forensic and non-forensic groups. Archival data provided the study with a large sample size and with a variety of participants. Due to the wide range of participants in the study, it is likely that these results can be generalized to a larger population. The study was also conducted within a large organization and was adopted from previous researchers, which allowed for access to valuable information and to monetary compensation for participants.

Despite these strengths, some limitations were noted within the study. Although the results of the study may be generalizable due to the large sample size, it may not account for societal or cultural differences across a larger area and therefore may be limited to this, or similar, communities. Further, the non-forensic group accounted for a majority of the overall number of participants. The smaller number of participants in the forensic group may have limited the true representation of this group.

**Study Challenges and Ethical Concerns**

The present study was completed with relatively minor challenges, however, there were some challenges noted that occurred within both groups and across multiple levels of service.

**Client level.** Many forms from both groups had to be deleted from analysis due to missing information. The non-forensic group often left a number of questions blank – most likely due to lack of motivation to fully participate – and the forensic group often missed or cancelled appointments. Excessive amounts of missing information from any form resulted in the form being deleted from the analysis.

**Program level.** Many participants of the non-forensic group were often not motivated to fully participate in the process of completing all the forms; many individuals skipped questions or completed the forms at a fast rate which may suggest that they were not completely engaged in the process.

**Organizational level.** The study faced challenges at the organization level in that the clients of the non-forensic group were attending the hospital for other services. Although each participant was informed that their participation had no connection to any other service being received, many individuals seemed suspicious of this fact and inquired if there was an underlying purpose or if it was somehow connected to their visit. It is possible that this noted distrust had some form of impact on the responses provided by this group.

**Societal level.** Societal views of individuals involved with the CAS seemed to have an effect on individuals in both groups, but more specifically within the non-forensic group. The majority of these participants did not have previous contact with the CAS, but those who did were determined to explain why the CAS had been involved. Although this may not have had a
direct impact on the L²P/F v.54 form, it does suggest that these participants were aware of possible social implications connected with completing the form.

**Recommendations for Future Research**

Recommendations for future research include gathering data from a larger number of participants in the forensic group in order to provide a more representative sample. Further, conducting this research within other communities may increase the generalizability of the results. Finally, it may prove beneficial to add or remove certain questions for the L²P/F v.54 form in order to establish a stronger correlation with the SDE construct of the PDS, as only one of the current constructs of the L²P/F v.54 form has a strong correlation with it. With the modification of certain questions, the L²P/F v.54 form may have stronger convergent validity with the PDS.
References


Eve, P. M., Byrne, M. K., & Gagliardi, C. R. (2014). What is good parenting? The perspectives of different professionals. *Family Court Review, 52*(1), 114-127. doi:10.1111/fcre.12074


