

The Order Dominance Scale:

Validity Analysis

Kyle Garrett Jones (206508190)

UKZN – PMB Campus

2012

Supervisor: Prof Lance Lachenicht

Submitted in partial fulfilment of the requirements of the degree of Master of Arts in the School
of Psychology in the University of KwaZulu-Natal, Pietermaritzburg

Abstract

The Order Dominance Scale (ODS) was constructed and proven reliable by Jones (2009). The scale is based on the combination of two theories; namely, Broken Windows Theory or BWT (Wilson & Kelling, 1982) and Reversal Theory (Apter, 1982). BWT provided the gap in the literature on which the personality trait of order dominance is based, while Reversal Theory provided the means with which to measure this trait. The purpose of this research was to take the completed scale constructed by Jones (2009) and test it for various types of psychometric validity. The research was able to establish predictive validity (although not in the way it would conventionally be done) and construct validity (which was established in conjuncture with convergent and discriminant validity). After going through testing for both reliability and validity, the ODS can now be considered a worthwhile scale by psychometric standards. The sample with which the data for this research was collect yielded a new set of norms for the ODS. The new set of norms gives it a mean of 31.38 with a standard deviation of 7.79, an absolute range of 0-50, an observed range of 8-48, and a reliability coefficient (alpha) of 0.7527.

DECLARATION

Submitted in partial fulfillment of the requirements for the degree of
Master of Arts in the Graduate Programme of Research Psychology, University
of KwaZulu-Natal, Pietermaritzburg, South Africa.

I declare that this dissertation is my own unaided work. All citations, references and
borrowed ideas have been duly acknowledged. It is being submitted for the degree of
Master of Arts in the College of Humanities, University of KwaZulu-
Natal, Pietermaritzburg, South Africa. None of the present work has been submitted
previously for any degree or examination in any other University.

Kyle Garrett Jones

Date

Professor Lance Lachenicht

Date

Contents

1. Introduction.....	8
2. Literature Review.....	10
2.1. Broken Windows Theory.....	10
2.2. Reversal Theory.....	17
2.3. Secondary Theories.....	24
2.3.1. Evolutionary Game Theory.....	24
2.3.2. Social Identity Theory.....	26
2.4. The Order Dominance Scale: Construction and Reliability Analysis.....	30
2.4.1. Construction.....	30
2.4.2. Sample.....	32
2.4.2.1. Pilot Sample.....	32
2.4.2.2. Main Sample.....	33
2.4.3. Results and Brief Discussion.....	34
3. Aims and Rationale.....	40
3.1. Aims.....	40
3.2. Needs for the Scale.....	44
4. Methodology.....	47
4.1. Research Design.....	47
4.1.1. Construct Validity.....	47
4.1.1.1. Designing the Vignettes.....	48
4.1.2. Predictive Validity.....	49
4.1.3. Convergent and Discriminant Validity.....	50

4.1.4.	Establishing Validity magnitude.....	50
4.1.5.	Re-Establishing Norms.....	51
4.2.	Sample.....	51
4.3.	Procedure.....	53
4.3.1.	Data Collection.....	53
4.3.2.	Data Coding.....	54
4.3.2.1.	Coding Vignettes.....	54
4.3.2.2.	Quantitative.....	58
4.4.	Ethical Considerations.....	59
5.	Results.....	61
5.1.	Demographic differences.....	61
5.1.1.	Age.....	61
5.1.2.	Gender.....	63
5.1.3.	Race.....	64
5.1.4.	Religion.....	66
5.2.	Construct Validity.....	68
5.3.	Predictive Validity.....	75
5.4.	Convergent and Discriminant Validity.....	75
5.5.	Establishing Validity magnitude.....	76
5.6.	Re-Establishing Norms.....	77
6.	Discussion.....	78
6.1.	Interpreting Results.....	78
6.1.1.	Demographic differences.....	78

6.1.1.1.	Age.....	79
6.1.1.2.	Gender.....	80
6.1.1.3.	Race.....	81
6.1.1.4.	Religion.....	83
6.1.2.	Construct Validity.....	83
6.1.3.	Predictive Validity.....	87
6.1.4.	Convergent and Discriminant Validity.....	88
6.1.5.	Establishing Validity magnitude.....	89
6.1.6.	Re-Establishing Norms.....	90
6.2.	Theoretical Implications.....	91
6.2.1.	Broken Windows Theory.....	91
6.2.2.	Reversal Theory.....	94
6.2.3.	Secondary Theories.....	96
6.2.3.1.	Evolutionary Game Theory.....	96
6.2.3.2.	Social Identity Theory.....	98
6.3.	Weaknesses of Research.....	99
6.4.	Future Research.....	100
7.	Conclusion.....	103
8.	References.....	106
9.	Appendices.....	109
9.1.	Consent Form.....	109
9.2.	The Order Dominance Scale and Scoring sheet.....	110
9.3.	The Negativism Dominance Scale and Scoring sheet.....	113

9.4. The Telic Dominance Scale and Scoring sheet.....	115
9.5. The Vignettes.....	119
9.5.1. Pair 1.....	119
9.5.2. Pair 2.....	120
9.5.3. Pair 3.....	121
9.5.4. Pair 4.....	122
9.5.5. Pair 5.....	123

1. Introduction

This research is a continuation of the research by Jones (2009) in the construction and reliability analysis of the Order Dominance Scale (ODS). The scale is based on two primary theories: Broken Windows Theory (BWT) (Wilson & Kelling, 1982), which revolves around the idea that disorder provokes petty crime which may provoke more and more serious criminal activity, and Reversal Theory (Apter, 1982), which revolves around the idea that personality is bistable and every person has the ability to conduct themselves in one of two oppositional states of each part of their personality.

The idea for order dominance to be treated as its own separate personality trait is new to both these theories. Previous research involving BWT has labelled order dominance as a phenomenon, leaving potential for it to be explored as an underlying personality trait. The trait was always a social phenomenon and the result of manipulating that trait was always measured in the effected change in the surrounding environments. In the case of the ODS, the trait becomes a personal attribute existing on different levels of each person, potentially dependant on a number of social and genetically inherent factors. This research is taking the effect seen in other studies and applying it to the individual, suggesting that the phenomenon is a result of a number of individuals acting according to their level of order dominance at any one time. The ODS suggests that environment and the social context still play an important role in each person's observed order dominant behaviour, but the behaviour is ultimately down to that individual, not the environment.

One of the reasons that the individual is the key component in the nature of their surrounding environment is because this trait is assumed to be at least bistable, meaning someone has the potential to be either order or chaos dominant. Dominance in one of these two states means that opposing environmental and social conditions need to be stronger than they would have to be in an individual who exhibits no signs of either form of dominance in order for switching to occur. This means external conditions are not sufficient predictors of behaviour for this trait.

The purpose of this research is to complete the validity analysis on the ODS created by Jones (2009) and complete all the psychometric requirements for the scale to be considered a worthwhile measure of the order dominance trait. The research was plagued by a few logistical problems and the scope for the research was narrowed slightly, aiming to measure three types of construct validity (including convergent and discriminant validity), and one type of predictive validity with the use of one sample. The data from this research will also be used to re-establish the norms for the scale originally given by Jones (2009) as those norms were calculated on a scale that was quite different (structurally) to the now 25-item long order dominance scale that exists.

2. Literature review

The order dominance scale (Appendix 9.2) (Jones, 2009) was constructed on the basis of two psychological theories. The idea for the ‘personality trait’ that the scale attempts to measure came from a theory introduced by James Wilson and George Kelling (1982) called Broken Windows Theory (BWT), while the concept for how the trait might present itself in different scenarios was taken from a theory called Reversal theory introduced by Michael Apter (1982). These two primary theories, along with other potential theories on how the trait may present itself will be discussed in detail. As the research is a continuation of a previously completed construction and reliability analysis of the order dominance scale (ODS), a section will be dedicated to include an explanation of these previous processes. Literature behind the scales based on the premises of Reversal theory, used in the reliability analysis and the validation procedure, will also be discussed in this section.

2.1. Broken Windows Theory (BWT)

The theory of broken windows is a simple one as far as social theories go. It states that a high level of disorder in an area can contribute to, or provoke, a higher rate of crime. In this respect, the theory claims that by removing the disorder in an area, through increasing the level of social control, one ultimately lowers the overall crime rate of the area (Sampson and Raudenbush, 1999). If one were to consider an area under high levels of social control, and thus (according to BWT) one with low levels of crime, and slowly begin to introduce more visible disorder, then

BWT would assume that not only the amount of crime, but the severity of the committed crimes would begin to increase. Thus, the theory does not claim to work in one direction, but that an intervention of both a favourable and unfavourable kind can affect the levels and severity of the observed crimes in any particular area.

BWT only becomes relevant in specific social circumstances. Under normal circumstances humans look to others for social cues regarding how they should behave in any given situation. However, BWT is relevant when the persons involved do not have the social cues they would normally get from others and they are forced to look for other clues as to the appropriate social behaviour. This is where the surrounding environment becomes so important, as this becomes the next best source for these social norms. The state the area is in becomes the role model to which the exposed individual benchmarks their own behaviour. In other words, an area which is clean and ordered suggests that the people in this area conform to the laws as laid down by their relevant authorities or that the area is well monitored and criminal behaviour is more difficult to get away with. In a similar vein, a derelict and unkempt area suggests that social norm is in conflict with the laws laid down and that criminal behaviour is tolerated and easier to get away with.

Wilson and Kelling's (1982) article gives the impression that the ideas the theory introduces can be applied in any setting and there will be some level of change in criminal activity. However, critics of the theory claim that the more poverty stricken a population is the less impact BWT would have on reducing the crime in the area. Harcourt and Ludwig (2006) have also found contradictory evidence in a re-housing project that moved inner city New York tenants to more

orderly neighbourhoods. The result of the move did not, as BWT would suggest, see a significant drop in crime rate, but instead the old rate of criminal activity moved with the relocated tenants.

Regardless of its potential problems the theory has been implemented in a number of different areas around the United States, including: Albuquerque, New Mexico; Lowell, Massachusetts; and most famously in New York City all with varying rates of success (Corman and Mocan, 2005). All these programs were similar in nature; effort was put into policing small crimes (without neglecting the more severe ones) to decrease the amount of visible disorder in an area with the intended outcome of eliminating the prevalence of the more severe crimes. All areas where the theory was implemented reported decreased criminal activity. It is also important to note that all areas partook in this type of program around the same time in the 1990s (Corman and Mocan, 2005).

Wilson and Kelling (1982) identify the Stanford psychologist Philip Zimbardo as making one of the earlier examples of this theory in 1969, a long time before the phenomenon was recognised by Wilson and Kelling (1982) as 'Broken Windows Theory'. They describe an experiment by Zimbardo where he leaves two cars with their bonnets open in two different areas. The one car is left in the Bronx in New York City, an area that by reputation alone is known as an unsafe area, and the other car is left in Palo Alto in California, an area considered much safer. The car in the Bronx was subjected to vandalism within minutes and was nothing more than a shell for children to play in after three days. The car left in Palo Alto was untouched for over a week. This cannot confirm BWT because it could simply be the case that Palo Alto is a safer place, while BWT would argue that it was the environmental surroundings that determined whether the car was

damaged or not. Zimbardo proved that this was indeed what was happening and that it was not only because Palo Alto is a safer place. He went on to damage the car in Palo Alto with a sledgehammer and soon afterwards, the car was being destroyed by passersby, eventually ending up on its roof. This suggests that the way the car was treated depended entirely on the moral code that was implicit in the environment. If the people felt that it was acceptable to vandalize the car, because it had already been damaged, then it did not matter what the surrounding environmental laws would have suggested but only what the immediate social law was dictating.

Over the years the theory has evolved and alternative explanations for the phenomenon have been given. Gault and Silver (2008) reviewed one of these possible alternatives. Wilson and Kelling's (1982) original theory stated that it was disorder that would lead to lower forms of informal social control and ultimately see a rise in the overall crime rate (Figure 1a). The alternative, originally offered up by Sampson and Raudenbush (1999), instead states that it is low informal social control that leads to both an increase in disorder and crime (Figure 1b). In this respect, an increase in the social control of an area, like the introduction of a foot patrol, will lead to a decrease in both criminal activity and disorder.

Figure 1



Figure 1a: Wilson and Kelling (1982) BWT model

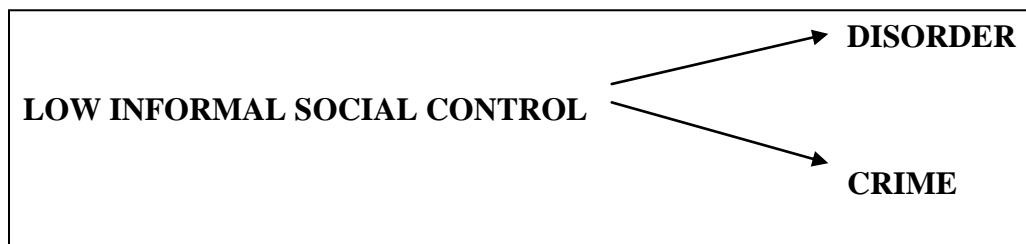


Figure 1b: Sampson and Raudenbush (1999) BWT model

The Order Dominance scale, however, has been constructed around the original model of BWT as this was the model that was adopted by Keizer, Lindenberg & Steg (2008) for their small-scale study in the Netherlands (the experiments in which the ODS was based on). The idea they were testing was that by varying the levels of disorder in an area they would be able to create an environment of low informal social control and provoke their participants into committing more of their own petty crimes. The difference in their experiment was that the social control was not always informal, but in some cases, clearly stated by law. The social control was depicted as either high or low in the different situations. A high level of social control is essentially the equivalent of a social norm that is in line with the general law; conversely, a low level of social control is the equivalent of a social norm, either directly or indirectly, in conflict with the general law.

These ideas were replayed in six different scenarios by adding or removing disorder and witnessing the effect it had on the people exposed to it. In one scenario, the walls of an alley full

of parked bicycles were given a sign prohibiting graffiti (general law). A flyer was then placed on every bicycle. One circumstance required that the alley be full of graffiti and the other required that it be clean of graffiti (social norm). In the circumstance where the social norm was in conflict with the general law 69% of the subjects littered by discarding their flyer. When there was no conflict, this figure decreased to 33% and proved to be significantly different (Keizer et al., 2008).

The major critics of the theory claim that the large-scale implementations of the theory's ideologies make the crucial error of assuming correlation implies causation. In other words, the results from the programs implemented in the various places in the United States were attributed to the program of BWT and did not consider the possibility of other social influences occurring at the same time. Although these areas all report drops in criminal activity (Corman and Mocan, 2005), Harcourt (2001) highlights that crime rates in the United States dropped in general during the 1990s when the Broken Windows Theory was famously put into effect in New York City and other areas. Possible reasons for this could have been the police reforms at the time, the 500,000 people that were moved into jobs by certain welfare programs or the housing vouchers that enabled poorer families to move into better neighbourhoods (Harcourt, 2001). Other theories have suggested that the decrease in a nationwide crack cocaine problem or the decrease in the number of high-risk males (aged 16-24) could explain the observed decrease (Harcourt, 2001).

While these alternative explanations are speculations, Harcourt (2001) at least considers that the effect of the BWT programs as potential speculation, i.e. in order for BWT to have been considered a success crime rates would have had to drop significantly more in the areas exposed

to the programs than the national average, otherwise they fall into the trap of confusing correlation with causation. This does suggest that BWT has problems working on a large scale and that the effect perceived is no more than an illusion of safety rather than an actual decrease in the crime rate. In this regard, large scale implementations of BWT appear to follow a model closer to that of Sampson and Raudenbush's (1999) model than the original model. The programs based on BWT effectively attempted to increase the levels of social control, i.e. lessen the occurrences of conflicting social norms. This was only able to decrease the levels of disorder, but failed to have any effect on the rate of crime (Harcourt and Ludwig, 2006). Thus, the introduction of a foot patrol would decrease the levels of disorder in an area and result in a perceived decrease in crime. Crime has not actually decreased, but because one of the two outcomes of higher informal social control has been attended to (disorder), the perception is that both have been dealt with to some degree. If this logic were applied to the example of New York City, then the act of controlling the disorder would not have an effect on the level of crime in the area, as the informal social control may still be low (in accordance with Sampson and Raudenbush's (1999) model).

However, when the studies have been involved with people on a smaller scale or more personal level, then, as with Keizer et al (2008), the model appears to adhere to the original idea of BWT and not the model suggested by Sampson and Raudenbush (1999), i.e. controlling disorder raises the level of social control and decreases the amount of crime. This would imply that the alternative model may not actually apply to any implementation of BWT, i.e. the various programs in the U.S.A, and that it may only exist as another speculated co-variable removing

from the actual effect of the BWT programs. It is for these reasons that the original model has been the model of choice for the remainder of the research and the original scale construction.

The idea that programs and experiments that have used BWT only create a perception of decreased crime (although the evidence from Keizer et al. (2008) suggests that this is far less perception than reality) may lend credence to the need for a personality test that is able to distinguish between those who are stopped by the illusion of order and those who still continue to commit the crimes. For example, if one assumed that the samples that Keizer et al. (2008) used in each circumstance were taken from the same population, then by the laws of random selection these samples would be as near as identical as possible. This means that the conflict between the general law and the social norm reached a tipping point for 36% of that population. This introduces the idea that there is a personality trait in every individual that has not yet been considered for psychological measurement and that this trait has two oppositional states that are changeable through the introduction of external forces.

2.2. Reversal Theory

The trait of order dominance that Keizer et al (2008) exposes presents itself differently in different situations. Again, if one assumes that the samples taken for each of their 6 studies is random and, therefore, effectively identical then the percentage of people who reached tipping point in each experiment is: 36% for study 1 (109.1% increase), 55% for study 2 (203.7% increase), 28% for study 3 (93.3% increase), 28% for study 4 (53.8% increase), 14% for study 5 (107.7% increase), and 12% for study 6 (92.3% increase). One can see that these tipping points

vary fairly randomly from situation to situation. This suggests that the trait of order dominance is not a stable one and that a person may present themselves as order dominant in some scenarios and chaos dominant in others.

Michael Apter's (1982) theory of reversals provides the perfect framework off which to build a scale that could measure this personality trait. Reversal theory rejects the commonly accepted idea of homeostasis in social sciences, an idea that states that 'the value of some variable can be held reasonably steady by interacting forces in a complex system' (Apter, 1982, p. 19).

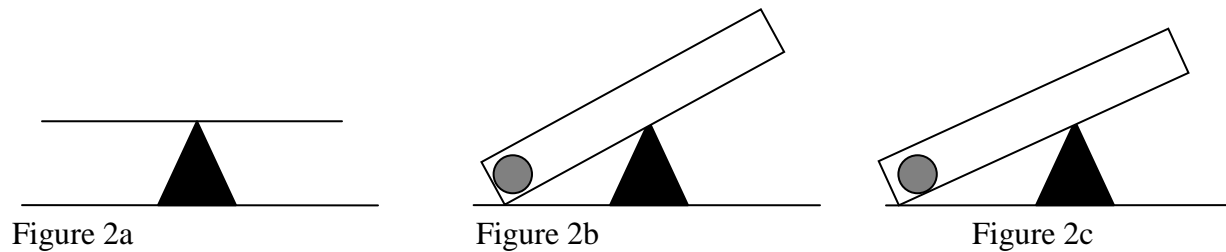
Therefore, if the interacting forces were those of chaos then it is most likely that the value of disorder will be held constant and promoted within this environment. It is clear though that this does not happen, as one can say with fair confidence that the most chaotic environments must play host to at least a few very ordered individuals and vice versa.

Apter (1982) introduces the idea of bistability instead. Bistability states that these variables would have two preferred states that would work in a discontinuous nature to one another. Apter (1988) also describes it as 'the conjunction of two self-correcting (homeostatic) mechanisms so as to form a single but more complex system (p.9). Lachenicht (1988) describes these as mutually exclusive pairs, each with their own internal stability. In essence then, these variables act like a switch; stable in the position of on and off, but with so little control in between the two states as to be negligible. In psychology then, bistability can be used to describe opposing mental or behavioural states occurring in an individual, such as order dominance. According to Apter (1982), switched states must be oppositional to each other and while the switch itself is always discontinuous, the adjustment to the new state is slow and more continuous (for this, imagine the

switch turning on an air-conditioner; the state of 'on' is immediate, but the effect of cooling the room down is delayed). It is also important to note that out of the two oppositional states, one state usually dominates over the other, and the external forces acting on the state must be extreme to force a change to the opposing state.

Apter (1982) uses the example of a fulcrum with a board balancing on top in such a way as to achieve a perfect horizontal balance. Any external force that acts on this board will cause oscillations, but the board will eventually return to its stable horizontal state. This is an example of how homeostasis might work (Fig. 2a). Apter (1982) introduces a slightly more complicated model for the explanation of bistability. The same fulcrum now balances a board on top of it that has Perspex attached all around the edges of the top surface. A marble is now inserted inside the hollow cavity that has been created by the surrounding Perspex (Fig. 2b). Now the board will behave more like a switch, resting on one side of the fulcrum until an external force shifts the board causing the ball to roll to the other side of the board where it will rest on the other side. When the board can be shifted from one side to the other with an equal force on both sides, there is absolutely no discrimination between the two states. However, if one were to shift the fulcrum from left to right the idea of dominance is now being introduced. An order-dominant or chaos-dominant individual would have a fulcrum shift slightly to the left or right (Fig. 2c).

Figure 2: Visual Representation of stabilities



This illustrates how the order-dominant personality trait can be seen to have bistability. Apter (1982) would argue that there must be a point, although not the same point, in every individual where it becomes beneficial for them to switch from contributing to the chaos and disorder to wanting to fix the problem and the same may be said for the reverse effect. Apter (1982) goes on to make a crucial distinction between two different types of bistability; value-determined bistability and externally-controlled bistability. Value-determined bistabilities rely on the amount or value of the variable to determine which of the two states is preferred, whilst externally-controlled bistabilities rely on an external force to change between states and the value of the variable is not important.

Apter (1982) illustrates the difference with the example of a playing card. If one were to place a playing card vertically on a flat surface, using a finger to hold it up, the exact angle that the card is at, should the finger be removed, will determine which side the card ultimately lands on. In this case it is the angle of the card that determines its final position (face-up or down), or it is the value attributed to 'angle' that identifies the final value. This is considered a Value-determined bistability. For this to be considered an externally-controlled bistability, the finger would have to play a more important role. Again, imagine the playing card being fixed vertically (i.e. exactly

90°; although this is not a pre-requisite for the experiment to work as the value of the variable is not important), but this time being held in place between two fingers. If the one finger is removed the other finger will push the card into one of its two stable states. Now, it is not the angle of the card that determines the card's final position, but which finger has been removed, i.e. the finger is external to the variable of the card's angle. The individual (either order or chaos-dominant) will favour one state over the other, as mentioned earlier, and, from the research on BWT, the choice to let one of those two states dominate is not made consciously by this individual, but rather forced upon them as their external environments change. The scale, then, attempts to exploit these externally-controlled bistabilities as items forcing the participants to choose between order and chaos-dominant states.

Both the idea of dominance and switching between states is dependent on a number of other factors. The external force may manifest itself in a number of different ways. Apter (1982) mentions how the environment is one major contributing factor, both to a switch and to a dominant state. An extreme change in an environment may force a change in states, just as a stable environment will cause one state to be preferred or dominant. Time also plays a major role in the change of states. A dominant state is likely to be one that an individual has been in for a long period. Converse to the idea of time, is the idea of frustration; if being in one state for a long period has not helped the individual move forward and achieve certain goals then a switch to the opposing state is likely.

These all give credence to the idea that reversal theory has its own form of phenomenology embedded into it. Although this is not in the strict definition of the word, reversal theory is

primarily concerned with subjective experience and meaning as it pertains to each person; potentially bringing the theory closer to one of experientialism, more than phenomenology (Apter, 1982). Lachenicht (1988) points out that much of the social and personality theories in psychology tend to assume that man tends to seek consistency if they are not already consistent in their personality. Reversal theory takes an entirely different stance from this, as the example of the fulcrum clearly points out. Most theories seek to test an individual's personality and the result of those tests will narrow a person down to one specific personality type. Apter (1982) contradicts this notion by introducing his theory of reversals, i.e. a personality assessment score may vary depending on some dominant state an individual may be in at the time.

Human behaviour can be lumped into two major categories; essential and inessential behaviour.

Apter (1982) lists a number of inessential behaviours that are also harmful like sadistic and masochistic behaviour, vandalism and hooliganism, dangerous sports, gambling, alcoholism, drug-taking, both celibacy and recreational sex, and suicide to name a few. Behaviour like this lends evidence to both the idea of opposing psychological states and the inconsistency of people. People are, therefore, inconsistent by nature (Lachenicht, 1988) and if they were to remain consistent in anything for too long in life, the result could be harmful. The ability to adapt is what gives a person the edge in different situations. A person who reacts to every situation in accordance to a stable personality trait is highly unlikely to survive at a social level. The ability of two individuals of similar personality types to experience and interpret a situation in two different ways is a testament to that fact. Clearly, experience is a crucial factor in determining the state an individual is in and, therefore, the way they interpret the situation at hand. This is the

internal factor of an external experience, and theories that suggest consistency would be unable to predict an individual's experience based on a stable model.

What this can all be summarised to mean is that although the external forces acting on an individual may play a role in forcing them into one of two opposing states, each person is still ultimately entitled to their own internal experience of those external forces. In this regard, reversal theory can be thought of as mechanistic one, although not in the sense of 'behaviouristic' or 'stimulus-response' that would usually be attributed to it as the individual is relatively autonomous and not 'pushed around' by their external influences (Apter, 1982). If, as it is being assumed, the trait of order dominance is indeed bistable and adheres to the ideas of reversal theory it highlights the need for a measurement to identify the level of the order dominance trait that any individual may possess at any one time. This is because this combination of external and internal forces, including that of time, could lead two different types of people to experience the same event in exactly the same way. A scale would, be able to separate these two types of people through exposure to a series of different scenarios and accurately say whether they are in fact any different. However, if the ideologies of reversal theory presented here are accurate then the resulting scale may only be context specific and give measurements of the trait dependant on the environment and the time of administration.

2.3. Secondary theories

2.3.1. Evolutionary Game Theory

BWT and its adaptations seem to suggest that every individual, under the right conditions, has the capacity to tip into a form of chaos contribution. While this may be true, it is important to note that without the ability to tip back into some form of order contribution the world would slowly, but surely, slide into complete chaos and discord. Apter's (1982) theory of reversals offers up one plausible reason for why this does not happen on an individual level. However, evolutionary game theory may offer up a plausible explanation for why this happens on a social level through the idea of evolutionary stable strategies, introduced by John Maynard-Smith (1964). An evolutionary stable strategy (ESS) is defined as a strategy that cannot be bettered, as long as a sufficient amount of the group members adopts the strategy (Maynard-Smith, 1964). The most famous example of an ESS is given by evolutionary game theory, which is an adaptation of the Nash equilibrium (Cressman et al., 1998). This example is known as the Hawk-Dove game and it sets up a conflict between two strategies that work very differently within the same population. In this game, an individual's success in making choices depends on the strategy to which they subscribe as well as the choices of others, subscribing to their own or another strategy (Dawkins, 2006).

The Hawk-Dove game, also known as a strategic situation, is one of many examples of game theory in practice and specifically adopts the idea of aggression (fairly closely related to the idea

of order and chaos). In a hypothetical species only two types of aggression exist; the hawk strategy and the dove strategy. The hawk will always fight as hard and unrestrained as possible and only ever retreat when seriously injured, while the dove does nothing more than threaten without ever injuring anyone. In this respect, a fight between two hawks will end in serious injury or death for one of them, a fight between two doves will end in one eventually tiring and just giving up and a fight between a hawk and a dove will result in the dove retreating almost immediately. Therefore, the hawk will always win, even at great cost to itself. However, the ESS will be the one of the two strategies that dominates over the other and evolves. Contrary to what one might first think, the hawk is not an ESS, nor is the dove. In a population full of doves it would only take one hawk to change the dynamic of the population and as a result the hawk strategy will dominate. But a population full of hawks means each hawk pays a huge price every time they lose a fight; their risks outweigh their benefits. This makes the unpopular dove strategy more viable to the species once again. The strategies are unlikely to oscillate so violently and if one was to assign values to the risks and benefits that each strategy incur, one could work out the ratio of hawks to doves that would reach equilibrium and, therefore, be an ESS, i.e. stability refers to population proportions of these strategies and not the dominance of either one.

(Dawkins, 2006)

This translates well into a strategy for an Order-Chaos game that can apply to BWT. The desire to lean towards chaos may be strong in a world full of order strategies as the order strategies would sort out the problems that the chaos strategies cause. The world would slowly become dominated by chaos strategies (as it is a desirable strategy) until the level of chaos would make it hard for the chaos strategies to survive. This would lead to a return of order strategies and a

balance would be struck somewhere between the two. This means that if one were to measure this trait in a sample of people there would be both order-dominant and chaos-dominant individuals. However, unlike the Hawk-Dove game there would likely be far more than just two strategies in balance. There could be many different levels of order-dominant and chaos-dominant individuals ranging between completely order-dominant to completely chaos-dominant. This means that there are a large number of these strategies balanced within our different environments and developing a scale to measure an individual's level of order-dominance may help in identifying the tipping point of any individual, albeit potentially affected by the specific context. In other words, how chaotic does the environment need to be before an order-dominant individual adds to the chaos and how ordered does the environment need to be before a chaos-dominant individual refrains from chaotic behaviour. However, a scale would be able to give a measurement that could help to assess this tipping point in not only the extremes, but on every different level of order dominance. Although this simulation is not undertaken in this research, it may be something of interest for future research. As mentioned earlier, Keizer et al. (2008) were able to induce a relatively similar population to different tipping points over a series of studies, suggesting that this tipping point is indeed relative to the individual.

2.3.2. Social Identity theory

Social identity theory, like many psychological theories stretches far deeper and is far more intricate than its simple ideologies. However, it is these simple ideologies that apply most appropriately to the behaviours exposed by the studies of Keizer et al (2008). Social identity

theory was formally introduced by Henri Tajfel and John Turner (1979) and tries to explain a person's sense of who they are through their group membership.

Social identity theory, like reversal theory, also has the capacity to explain how the personality trait of order dominance might work. If one were to accept this as an alternate explanation to reversal theory, it would not have a serious effect on the way that the scale is developed, but it would offer an entirely different interpretation of the scores from each participant. Reversal theory sticks to the idea that order dominance is an individual experience, while if one were to apply social ideas to the concept it would be interpreted on an intergroup level. Social identity is constructed between two or more people on a linear scale that runs from the one extreme of interaction based only on the interpersonal level (between individuals) to the other extreme of interaction only based on the intergroup level (between different groups) (Tajfel and Turner, 1979). Both are unlikely to occur in their purest forms, but Tajfel and Turner (1979) gives the example of the relationship between a husband and wife as the closest to a pure interpersonal relationship and the relationship between soldiers on opposing sides as the closest to a pure intergroup relationship.

Social identity is also created through the idea of in and out-group identity and there was no better illustration for the creation of these identities than with the "robber's cave" experiments Sherif (1966) conducted with a camp for school children designed to create intense competition between two sets of children. The result of the competition meant that each group of children became fiercely opposed to the idea of the other group, even if some bond had been formed between the children in opposite groups before they were split. The other group of children had

become the out group, while the relative in-group had found reason to promote themselves as an in-group and strengthen their identity. Tajfel and Turner (1979) hypothesized that creating in and out-groups would not require as much effort as Sherif's (1966) experiment. He created a scenario where in and out-groups were decided based on a minimal in-group affiliation (the preference for one of two artists) and participants were required to divide money between their group and the out-group (the group who favoured the other artist). It was discovered that participants went for maximum difference between themselves and the other group rather than going for either the highest amount of money for their group or the highest amount of money all together. In these experiments it appeared to be necessary for the participants to internalise their group membership as an aspect of their self-concept (Tajfel and Turner, 1979).

Taking the idea of interpersonal and intergroup relationships and the idea of in and out-groups and applying them to order dominance is not a huge leap. Order dominance presents itself as an almost purely intergroup relation as the interaction does not occur between people, but between a person and their immediate environment. It is an indirect relationship, as the environment is manipulated by its inhabitants and other people interacting with the environment get their cues from it and not the inhabitants (although this does not have to be the case). This is similar to the way a soldier gets cues about his or her duty in a battle scenario. They do not actually interact on any personal level with their opposition, but have a fully functioning social relationship with the opposition soldiers.

If this is true, then it is true to say that the individuals interacting with the order or chaos dominant scenario also have an indirect relationship with the inhabitants of the relevant

environment. The individual unfamiliar with the environment takes the cues given to them and uses them to create in and out-groups of their own. It is important to note that these in and out-groups are created by the individual themselves based on their previous social values and the strength of these values. If an individual interacts with an environment that is primarily chaos dominant and the social values they hold from either the people they might be with at the time or the people they normally surround themselves with are primarily order dominant, then it is likely the order dominant values will be held as in-group values and the external environment and its inhabitants will be named the out-group and vice versa.

Of course if the values of the environment and the individual are in line then there is no immediate conflict and the individual will go on as they normally do. However, as was mentioned earlier, order dominance and chaos dominance are not the only two states that exist and the ODS will at least be able to expose a number of the different levels that do exist. This means that an individual will rarely encounter a foreign environment that aligns perfectly with the way they interact with their own and so some form of in and out-group will be created in most scenarios. The other important factor is the strength of the individual's social values. If they are weak values then they are likely to identify the in-group as their immediate environment and values that are not in this environment as out-group values. This links back to the idea of reversal theory and presents the idea that the individual may not switch according to individual preferences but that the act itself might be by a more social one.

2.4. The Order Dominance scale – Construction and reliability analysis

The scale is based primarily on the premises of BWT and reversal theory and the initial items for the scale were constructed with both in mind.

2.4.1. Construction

The initial research by Jones (2009) started by constructing a pilot scale with around a hundred items that could set up short and easy to understand situations where a conflict between social norms and general law was present. A period of brainstorming, between the principal investigator and two other psychology honours students at UKZN, was required in order to acquire as many of these situations as possible that could be adapted for use in this pilot scale (Loewenthal, 2004). The ideas for the items revolved around the sorts of situations that were present in Keizer's et al. (2008) study, being careful to avoid offense. Order-dominance can manifest itself on a number of different levels and these levels needed to be addressed in the questions. Examples of these are: Ownership of the property in question (Does the reaction to the situation change depending on who owns the property), the level of disorder present in the question (Does the reaction change when there is considerably more chaos present or required in the scenario), the level of social influence (Does the reaction change when the individual is alone, in company with friends or members of authority) and responsibility (Does the reaction change when the individual has a social or personal responsibility). These different levels were assumed to have no significant effects on the scale at first and a number of each level was included in the original pilot.

The pilot also included 8 items from the Negativism Dominance scale (McDermott & Apter, 1988) and 7 from the Social Desirability scale (Strahan & Gerbasi, 1972). The Marlowe-Crowne Social Desirability scale (SDS) was used to remove socially desirable items as well as remove people who answered items in a purely socially desirable way.

The Negativism Dominance scale (NDS), also known as the rebelliousness dominance scale, measures the trait of negativistic dominance and conformist dominance. Behaving in a negativist or rebellious way is defined as ‘wanting, or feeling compelled, to do something contrary to that required by some external agency’ (Apter, 1982, p. 198). The NDS is an 18-item scale that has two 7-item subscales measuring reactive and proactive negativism and four ‘filler’ items (McDermott, 1988). The NDS is a very similar scale to the proposed order dominance scale. It deals with an individual’s tendency to rebel against the social norm or the general law. The crucial difference between the NDS and the ODS has to do with the social norm and the general law. Whilst the NDS is concerned with rebellion against either of these two situations the ODS will attempt to identify an individual’s reaction to a conflict between the social norm and the general law. The NDS provided the closest possible match to an existing scale measuring a similar trait and was used both for the item format it provided (as it was a scale constructed using reversal theory as a basis) and to ensure that the ODS items were sufficiently different from the idea of Negativism Dominance, without being on the entirely wrong track.

To avoid socially desirable answers the items needed to involve moderate conflicts, as an extreme conflict would likely yield a socially desirable answer every time. It was important that

the items did not display patterns that were easy to identify for the participant. However, a few repetitive items were added into the pilot scale in order to identify which of the slightly different situations produced the better item. A total of 75 original ODS items were designed using these and the guidelines of the primary theories, bringing the total number of items to 90. The final scale intended on having around 25 items and, according to DeVellis (2003), to ensure that this many good items were in the final scale would require that around three times the amount of intended items were in the pilot scale. Each item had three possible answers that reflect an order-dominant reaction, a chaos-dominant reaction and an undecided reaction to the scenario. Participants were told that the answers to each item could not possibly cover all the possible reactions to the scenario and that they should choose the answer that would most likely cover their reaction. They were also told to avoid the undecided or 'not sure' option as often as possible as this was a last resort option and could lead to a skew in the results.

2.4.2. Sample

2.4.2.1. Pilot Sample

The pilot study was administered to ten people (all family and friends of the researcher) between the ages of 18 and 74, all white, including 3 females and 7 males, and 3 Christians and 7 Atheists or Agnostics. They were simply asked to identify items that were hard to understand, ambiguous, contained spelling errors or were offensive. Each participant took somewhere between 10 and 20 minutes to complete the pilot. A few items were found to be repetitive and ambiguous and one or two had spelling errors. Items that could not be fixed were removed from

the test and a few were reworded to clear up any ambiguities. All-in-all 10 items were removed from the test and the main study would be conducted with 80 items (65 order-dominant, 8 NDS and & 7 SDS).

2.4.2.2. Main Sample

200 copies of the test were circulated in and around the university and most of these were given out in a 2008 Psychology 201 tutorial on statistics, however, a portion came from other associations affiliated with UKZN. The 200 copies elicited 177 full replies and the remaining 23 copies were either discarded by the participant or were handed back with too little input for analysis. The response rate for the study was then 88.5%, which is well above a rate that would bring potential responder bias into the study. While all 177 participants completed the entire test, 21 of these neglected to fill in the demographic information on the answer sheet. However, of the valid demographic cases, age ranged from 18 to 38, with a mean of 21.23 and a median of 20. 53.8% of the sample were either age 19 or 20. As is concurrent with university norms and psychology norms, a large portion of the sample was female, with 78.8% of the valid sample consisting of females. In terms of race, the sample was 44.9% Black, 32.1% White, 15.4% Indian and 6.4% Coloured. Religious beliefs were also gathered and 75.6% of the valid sample was Christian with the next biggest group of 'Ambivalent' contributing 14.7%. The 'Ambivalent' group included any person that did not subscribe to a form of organized religion.

2.4.3. Results and Brief Discussion

The ten items removed from the very first pilot were removed based on a number of factors. A few of the items were very close to repeats of early or later items in the scale and needed to be removed to avoid giving the scale a repetitive nature. A few of the items were culturally dependant and while culture is a potential mediating variable in the results, culture specific items were deemed undesirable, as they could be offensive or irrelevant to people with different cultural subscriptions. The scale as a whole had too many items that were concerned with driving and this is not a reality for a large portion of the target populations, so a few items on that subject were removed. Lastly, a few items were deemed too extreme for the scale or potentially silly and were unlikely to show any form of discrimination even in a large sample.

The data from the main study was coded according to the same format as the NDS. Order-dominant answers were given a value of 2, 'Not sure' answers were given a value of 1, and Chaos-dominant answers were given a value of 0. The answers from the NDS items were coded identically to how they were coded in the original test, which means that rebellious answers should ideally correlate negatively with the order-dominant answers. The answers from SDS items were coded in the same way as the Order-dominant items with socially desirable answers receiving a score of 2. The original social desirability scale only offered the option between 'true' or 'false' so for the purposes of this research some of the social desirability items were turned into question with yes/no answers and all were given third option of 'not sure'. The purpose of this was to make these items indistinguishable from the rest. The advantage of coding

the answers in this manner, as opposed to a -1, 0, 1 format, is that order-dominance cannot be cancelled out in the total score and an accurate total will be reached. The major disadvantage of this method is that discrimination calculations will reflect lower levels of discrimination when they do in fact discriminate well (DeVellis, 2003). All items were coded according to this format and no items received any form of weighting.

Each participant's score was totalled and then both the NDS answers and the SDS answers were removed from these totals and added separately. The demographics of the sample were analysed as best as they could, keeping in mind that there were 21 missing cases for the sample. An independent samples t-test was run on the order-dominance scores for the category of gender and gave a p-value of 0.024 (mean for male was 65.33 and for female 71.29 out of a possible 130). A one-way ANOVA procedure was run for both the categories of religion and race. The means for religion were as follows: Christian (70.83), Hindu (68.91), Muslim (59.67) and Ambivalent (67.74) out of a possible 130. The p-value for the category of religion was 0.404 and Tukey's HSD only identified one subset. The means for race were as follows: Black (68.84), White (72.86), Indian (68.67) and Coloured (65). The p-value for the category of race was 0.192 and Tukey's HSD again identified only one subset. Both ANOVA procedures did not violate the assumption of homogeneity of variance. It is, however, important to note that these procedures (including the t-test) were run using significantly different sample sizes for each group. The fact that these differences were not too large was important for the study as the focus was on the items not the participants. Having identified an extreme population would have been interesting, however, as the scale was only in early stages of construction, it would have been difficult to identify accurately the segment of the sample that was extreme.

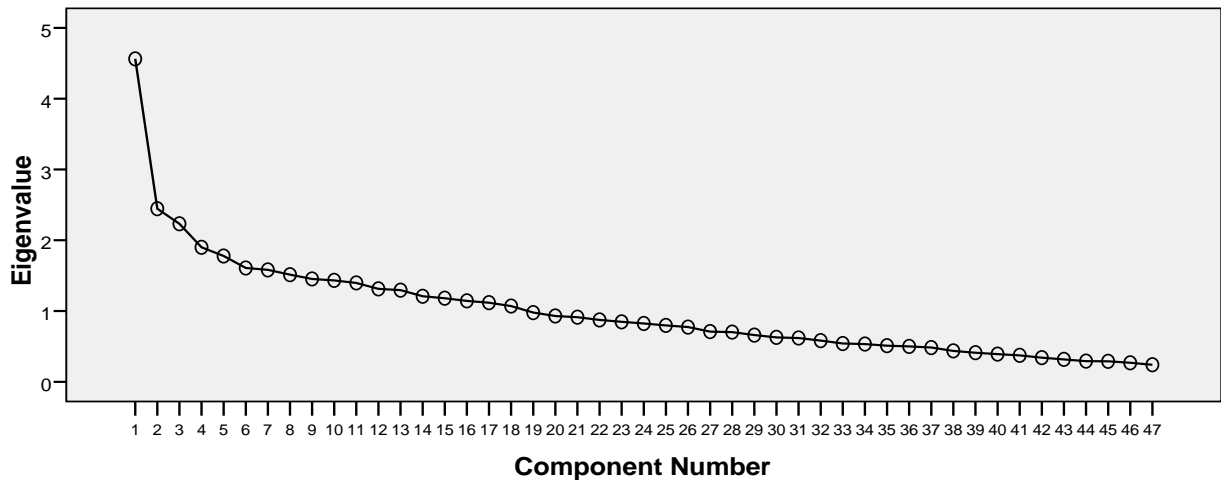
The mean order-dominant score for the sample was 69.62, with a standard deviation of 13.807, out of a possible score of 130. The scores ranged from a minimum of 32 to a maximum of 107. NDS answers correlated -0.361 with the order-dominant answers, accounting for a potential 13% of the variance, and SDS answers correlated 0.234 with the order-dominant answers, accounting for only 5.48% of the variance. The scores obtained by the NDS items were no longer used in the analysis and those scores got by the SDS were only used again to remove socially desirable items. At that stage, however, the correlation values given by Pearson's R suggested that the NDS was sufficiently related to the ODS, without there being any too much overlap, while the low correlation of SDS could be small enough that it could almost be attributed to chance.

The first part of the analysis procedure involved getting the original alpha of 0.744 closer to the more acceptable figure of 0.8 (DeVellis, 2003). The original alpha was calculated using all 65 order-dominant items in the scale. Reliability item analysis was used and removed any item that lowered the level of alpha (i.e. raised alpha when it was not included in the analysis), the original 65 order-dominant items were scaled down to 51 and the overall alpha level was raised to 0.783. The removed items were checked for high discrimination indices, as removing well-discriminating items may be detrimental to the final scale, but the process of removing items that lowered alpha appeared to have removed a number of items with low discrimination indices.

The remaining 51 items were then checked for high correlations with SDS and low discrimination. Items remained in regardless of their correlation with SDS if their discrimination index was over 0.3. 0.3 was chosen as an appropriate figure because the scoring system that was

mentioned above leads to slightly lower discrimination figures. These were calculated in the conventional way using the top and bottom 27% of the sample scores without NDS or SDS. For reference, the top and bottom 27% was calculated with and without NDS and SDS scores and the relevant t-test performed. The sig value comparing the top 27% with and without the NDS and SDS scores was 0.888 and the bottom 27% sig value was 0.935. Both of these are not significant and meant analysis could continue without the NDS and SDS questions. Items that had both discrimination below 0.3 and showed significant correlations with SDS were discarded and the remaining 47 items were now entered into a factor analysis procedure.

Figure 3: Factor Analysis Scree Plot



The factor analysis was left unrotated and factor loadings below 0.3 were suppressed. The factor analysis identified a large number of small factors. These 18 factors contributed to 64.329% of the total variance, with most factors failing to contribute more than 3 or 4%. The scree plot (Fig. 3) was able to identify the point at which these factors became negligible and the top 5 factors were carried through for interpretation contributing a total of 27.487% of the variance for all 47 items. The remaining factors all had no more than 5 loadings out of 47 over 0.3 and could not be

seriously considered for interpretation. The first factor had 25 loadings over 0.3 and all were positive. It was clear upon a closer inspection of the original questions that, barring a few of these, they were all related to the mental battle between the general law and the social norm. The second factor had 10 loadings over 0.3 and showed a bipolar nature. An interpretation of the second factor showed that it was concerned with social ideas of responsibility and culpability. The third factor had 9 loadings over 0.3 and also showed a bipolar nature. The nature of the questions concerned with this factor all revolved around the notion of a personal idea of responsibility, where one cannot be held culpable. Both the fourth and fifth factor only had 5 loadings over 0.3 and a closer inspection could not reveal any important links between the questions. Thus, the fourth and fifth factors were discarded and the resulting factor analysis identified 3 relevant factors that could be used to scale the test down to a more manageable 25 items.

Out of the 25 questions that loaded onto the first factor, 19 were chosen to be transferred to the final scale. The questions not included (questions 2, 6, 12, 32, 44, 75) were deemed to be slightly repetitive or not actually have anything to do with the factor of the general law vs. the social norm. Out of the 10 questions that loaded significantly onto factor 2, only 3 were chosen to be in the final scale. This was simply done by ignoring the questions that loaded negatively on the factor to avoid the bipolarity. Out of the 9 significant loadings on factor 3, 3 were also chosen for the final scale. The final 25 questions were then assigned numbers between 1 and 25. Using MS Excel's random number generator the questions were then randomly ordered according to the order in which the numbers were generated. The alpha level was recalculated with the final 25 items at 0.728 with absolutely no item that contributed to a decrease in alpha.

As can be seen the construction and reliability analysis of the ODS was comprehensive and followed many of the conventions of classical psychometrics. The scale required validation before it stood any chance of being published as a usable scale. The purpose of this research is to subject the scale to a rigorous validation procedure and establish it as a tool that could be used in future research.

3. Aims and Rationale

3.1. Aims

The order dominance scale has proven to be a reliable measure, but has not been shown to be a valid one. The present research has been designed exclusively to gauge a few levels of validity and establish whether the order dominance scale is a worthwhile measurement for use in future research. The present research then consists of two studies, which test the different types of validity required to validate any psychological measurement. As this scale does not have immediately similar scales, it was difficult to measure all types of validity, but the research has still been designed to identify the scale's level of predictive validity and construct validity. These two types of measurement validity outlined by Terre Blanche, Durrheim & Painter (2006) will ideally help in determining to what degree the Order Dominance scale does what it is intended to do. Up until this point, the scale is shown to reliably measure some personality construct. The current state of affairs assumes that this personality construct is order dominance, but cannot be shown to be true without validation.

Criterion-related validity is defined as the degree to which a measure is related to some other standard or criterion that is known to indicate the construct accurately. There are two main types of criterion-related validity; predictive validity (the degree to which the measure predicts future events logically related to the construct) and concurrent validity (the degree to which the measure is related to pre-existing measures of the construct) (Terre Blanche, Durrheim &

Painter, 2006). Predictive validity is something that was well within the scope of the research, but again due to the lack of similar scales, concurrent validity, was not. The second major form of validity, construct validity, with its opposites of convergent and discriminant validity, try to discover a relationship or lack of relationship between the measure in question and different theoretically associated or unrelated constructs respectively (Terre Blanche, Durrheim & Painter, 2006). So while similar scales do not exist, theoretically associated ones do, thereby bring construct validity within the scope of the research. Measuring these two types of validity was the purpose of the two studies carried out for this thesis.

The aim of the present research is to test these two forms of validity measures mentioned above on the Order Dominance scale itself. As already mentioned, the major problem in achieving these forms of validity is that the Order Dominance scale does not fall into a currently existing battery of personality tests. Literature on the trait is not explicit and the trait is only first identified as a distinct personality trait by Jones (2009). Previous literature has not distinguished the trait and has only played with it in experimental and observational scenarios, leaving the present research with practical applications off which to test the aspect of predictive validity, but no leg to stand on in regards to concurrent validity. For this reason, the validation procedure did not include concurrent validity as one of the validating mechanisms.

Predictive validity will not be measured in the conventional way, where a score on the ODS should be used to predict an observed behaviour at a later date. Due to the time constraints of the research and the ethics required to gather the information necessary to contact and follow-up on a sample, predictive validity will effectively be done in reverse. Instead of predicting a reaction

based on an individual's order dominance score, the score would be predicted from an assumption about a population. If a population of people could be identified who are highly order or chaos dominant, testing such a population and discovering trends between the two would be another method of gauging predictive validity.

Construct validity will be measured in three different ways. The first, more general type of construct validity will be measured in a similar way to how predictive validity would, except without the time gap between the two measurements. The second of these two measurements would be presented in a different format to the rest of the ODS with the aim being to prove the robust nature of trait, by discovering a strong relationship between the ODS score and the second measurement. The last two forms of content validity (convergent and discriminant) are the two forms most readily within the scope of the research. A scale mentioned numerous times in the original research and used in the production of the Order Dominance scale itself is the Negativism Dominance scale or NDS (McDermott, 1988). The NDS along with the Telic Dominance scale (TDS) (Murgatroyd, Rushton, Apter & Ray, 1988) use the same underlying theory that the Order Dominance scale used in construction and the two scales both measure personality traits that are in theory closely related to the idea of order dominance. The telic dominance scale measures a person's tendency towards being either telic or paratelic. A telic person 'is primarily oriented towards, or feels the need to be primarily oriented towards, some essential goal or goals' (Apter, 1982, p. 47). A paratelic person 'is primarily oriented towards, or feels primarily oriented towards, some aspect of his continuing behaviour and its related sensations' (Apter, 1982, p. 47). The scale is a 42-item scale that measures three critical subscales of a telic individual: serious-mindedness, planning orientation and arousal avoidance

(Boekaerts et al., 1988). High scores on these subscales mean an individual is telic dominant and low scores mean an individual is paratelic dominant (Boekaerts et al., 1988).

While both the NDS and the TDS are similar, they possess differences to the Order Dominance scale that should assist in assessing both convergent and discriminant validity. Oddly enough, these two measures could play a role in discovering a level of concurrent validity for the scale. The assumption is made that a scale such as this has not yet been constructed, so high correlations with either the NDS or the TDS would suggest that this has not actually been achieved. Ideally, the correlations between the NDS and the ODS will be higher than the correlations between the TDS and the ODS, as the former two scales are assumed to be closer related than the latter two.

Content validity, referring to the extent to which the scale represents all facets of a social construct (Terre Blanche, Durrheim & Painter, 2006), is not part of the design of the current research but may have already been addressed in the reliability analysis and construction of the scale. As discussed in the literature, order dominance is multifaceted and is likely to present on a number of different levels. The items in the original scale were chosen rather specifically to identify the aspects of order and chaos dominance on as many different levels as possible. The original factor analysis procedure showed clearly enough that the first factor was directly related to that of a conflict between the social norm and the general law. This attribute formed the core of order and chaos dominance and is why a large percentage of items loading on factor one were carried through to the final scale. On top of this, content validity is particularly difficult to genuinely establish especially when no existing scales cover the order dominance idea. It is on

these grounds that the aim of the current research was to establish predictive validity, construct validity and not content validity.

3.2. Need for the scale

Broken Windows Theory is a prominent theory in psychological literature and the fact that it has been applied in large-scale experiments is a testament to that. It clearly identifies an active personality trait in a person that has yet to be measured. The trait is similar to negativistic dominance, but does not actually require that an individual take part in any rebellious activities. Thus, developing a scale that can accurately measure this trait would strongly aid the proponents of BWT. Finding this trait in individuals would mean that what occurs in the premises of BWT is not just a random phenomenon, but also a measurable trait of human behaviour.

The scale could help to identify at risk populations. With access to their order-dominance scores methods could be put in place that could assist in reducing the risk of those populations involving themselves in crime. Criminal activity is as likely to occur within the same area that the at-risk individuals live as it is in neighbourhoods outside their own. In this respect, it would be of a far greater benefit to the law enforcement to be aware of the order-dominant scores of an individual than the criminal activity within an area, as this may not accurately reflect the behaviour of the community. In other words, the scale can target people instead of places and by doing this, insure that the problems are more directly and fairly solved.

In the same respect, the scale can be used in institutions that do not have access to the background information of the individual. Chaos-dominant children can be closely monitored at school. These children will not appear to be difficult children at all, because their personalities will not indicate as such. Chaos-dominance is not a form of rebellion, but another form of conformity, usually in a more negative light. These children will witness social norms in conflict with the general rules of the school and be more likely to follow the social norm because the behaviour seems acceptable. By identifying these students, efforts can be made to insure that there is as little conflict between the social norms and the general rules as possible. These children would be most at risk amongst rebellious peers and may find themselves conforming to another child's rebellion.

The scale could also be used with prisoners. While it is highly likely that the prisoners may answer the questions in a socially desirable way, it may still be possible to use the scale on prisoners who have served their full sentence and are due for release. Gauging the scores of these prisoners would help in identifying to what extent the prisoners have been rehabilitated and how likely they are to reoffend when reintroduced to society. A prison is a place where order is a constant theme and exposure to this theme day in and day out may be a genuine reason for the rehabilitation of the prisoner, however, a measure of order dominance will be able to measure how likely being reintroduced into a potentially less ordered environment would affect their chances to reoffend. The prison system could likely be a working example of BWT.

It is clear then that there is not only a need for this gap to be filled in the literature, but that the order dominance scale may also be useful in a number of institutions, including law enforcement,

school, work environments and prisons. Initiating this type of research in the future would also constitute one of the major aims of the research.

4. Methodology

4.1. Research Design

4.1.1. Construct Validity

The research is comprised of two different types of validity assessment in order to establish the scale as a working measurement. This version of construct validity requires that the ODS is able to successfully measure a level of this personality trait that can be used to ‘predict’ how an individual would react to a given scenario. In an ideal situation each person who was given the ODS to complete would be required to perform some task at a later time where the ODS score could be used to predict the reaction (this would constitute a genuine predictive validity measure). However, an experiment of this size would have been beyond the scope of the research and the idea was instead condensed into something that could be established on paper. As a result, the measure more closely resembled a measurement to gauge the validity of the order dominance construct than that of its predictive power. Each participant was required to offer their opinion on two vignettes at the end of the test battery they were to complete. The research made use of ten vignettes; five which leaned in an order dominant direction and five which leaned in a chaos dominant direction. Each participant was given one of each on their questionnaire. The desired result from this test was that people who were scored as highly order or chaos dominant would not be affected either way by these two vignettes, i.e. they would react in an order or chaos dominant fashion to both scenarios. However, those whose score on the

ODS was relatively mid-ranged would react differently depending on which way the scenario leaned, i.e. order dominant reaction if the scenario was order dominant and vice versa.

4.1.1.1. Designing the Vignettes

The vignettes (Appendix 9.5) were designed in relationship to the ODS construct itself. However, the items in the ODS are designed to give the participant as much freedom as possible in designing the scenario they are hypothetically involved in. Ideally, the participant completing the ODS will not dwell too long on a hypothetical scenario, as they are purposely designed to be vague and general so that special circumstances for an answer cannot be given. The vignettes were designed to let the participant think about the scenario they are involved in and offer as much detail about the scenario as possible without it becoming redundant. Crucially, the ODS items have a set range of answers and the participant is forced to give an answer that most closely relates to what their actual response may be, but no freedom to express that actual response. The vignettes gave the participants an opportunity to engage in a realistic situation and offer up whatever answer they liked. The idea behind this was that the participants' level of order dominance could also be accurately gauged by analysing the content of their replies. The method for how these replies were coded is discussed in greater detail later. All of the vignettes were proof read and piloted on a small scale to check that they were easy to understand and did at least look like they were measuring what they were meant to be measuring. This change in the presentation of order dominant questions would ideally establish the robust nature of the order dominance trait and give the scale a level of general construct validity.

4.1.2. Predictive Validity

As mentioned previously, a realistic measure of predictive validity was beyond the scope of this research, but it is still possible to measure predictive validity although theoretically in reverse. This is done by first assuming where a difference in the order dominant trait may lie in a population and then administering the ODS to opposite ends of the population and testing for differences. This would be in contrast to using ODS scores to define different populations and predict their behaviour at some future time. The ODS would then effectively be administered a long time after the predicted behaviour of the person has been established. In the case of the order dominance the assumption that older people would be more order dominant was made based somewhat on the experiments by Keizer et al (2008). For example, the first study done was in an area where people kept their bicycles and the area was defaced with graffiti. While these two activities are not beyond the older generation, it is still the younger generation that was being targeted. This suggests that it is the younger generation where switches were more likely to occur, i.e. less order dominance. Evidence for this in similar scales has already been established. The TDS (telic dominance scale) classifies a person as either telic (goal-oriented) or paratelic (primarily concerned with their current behaviour). Research with the TDS discovered that the older generation was far more likely to be telic than paratelic (Apter, 1982). Considering the parallels between this trait and the trait of order dominance it would be safe to assume that a similar difference would exist. In order to test this assumption the sample gathered aimed at reaching about 50 participants who were over the age of 50 and comparing them to the remaining sample to test for differences.

4.1.3. Convergent and Discriminant Validity

The two scales that are most theoretically associated to the ODS have been discussed already; the NDS (negativism dominance scale; Appendix 9.3) and the TDS (telic dominance scale; Appendix 9.4). Both of these scales, in their entirety, were given to each participant with the ODS. The NDS measures a very similar trait to the ODS but is scored in the opposite direction to the ODS; negativism is given higher scores whilst order is given higher scores in the ODS. In this respect, the theoretical similarities (both being constructed using reversal theory) and the similarities of the two traits themselves implied that the correlation between the two scales would be quite high, but negative. The TDS measures a different trait to the ODS, although vague similarities could be identified, and is scored in the same direction as the ODS. As the two scales share the same theoretical basis, but are measuring essentially different personality traits the assumption was made that the correlation between the two scales would be significantly smaller than the one between the ODS and NDS and that it would be positive.

4.1.4. Establishing Validity magnitude

Running these analyses can indicate validity, but it does not indicate how much validity, i.e. what the magnitude of the validity is. In order to truly measure the magnitude of the validity achieved, Westen and Rosenthal (2003) suggest calculating the effect sizes of the analyses run to establish whether or not the scale is valid. For each of the four types of validity that will be established, an effect size will be calculated. This statistic will establish the magnitude of the

validity that has already been established in each instance and would act as a better measure of validity than the result of the individual analyses themselves.

4.1.5. Re-Establishing Norms

The scale norms will be reassessed and recalculated as a secondary measure of reliability, as the best way to ensure that the conclusion that the scale is a reliable one made by Jones (2009) is to retest it, thus, decreasing the chances of error. The new set of norms will be given with the original set for evaluation on the reliability of the scale by those using it in the future.

4.2. Sample

It was difficult for the research to attain separate samples for each of the above mentioned forms of validity to test individually due to the short timeframe within which the research needed to be completed. It was for this reason that all the major types of validity were assessed with just one sample. 250 copies of the ODS, NDS and the TDS were made while 50 copies of each of the vignettes pairs (5 in all) were made to distribute, all of which were bundled together so that each participant was required to answer all tests and a random vignette pair. The majority of the test battery reached the tutorial of a 2010 Psychology 201 course in research statistics at UKZN for which permission from the course and tutorial coordinator was granted verbally. This was done for convenience as the sample was easily accessible and completing data for research gave the students some insight into the research process. This, along with the addition of a small incentive, was hoped to motivate a response rate well over the 70% accepted norm for response

bias. In addition to the students, the test battery was distributed to people over the age of 50. Of the 250 copies that were originally distributed 235 were returned completed in their entirety. The remaining 15 were either not handed back or not completed. This gave the research a response rate of 94%, far over the norm suggesting no bias in response.

The age of the sample ranged from 18 – 96, with a mean of 29.82 and a median of 21. However, 65.96% of the sample was 21 or younger and the disparity between mean and median is easily fixed when 45 people over the age of 50 are removed. The age of the sample would then range from 18 – 42 with a mean of 20.55 and a median of 20 suggesting that the age of the sample is more normally distributed. The majority of the sample was female (178 people or 75.74%), conducive to university norms, with the male population forming the remaining 24.26% (57 people). Each person was asked for their race and the sample showed that the majority of the sample was Black (129 people or 54.89%), with the remainder being made up of White (71 people or 30.21%), Indian (28 people or 11.91%) and Coloured (7 people or 2.98%) individuals. The last bit of demographic information given was the individual's religious beliefs and people of Christian faith formed 74.89% (176 people) of the sample, with the remainder being made up by ambivalent individuals (this included atheist, agnostic, spiritual, etc. Essentially anyone who did not form part of any organised religion) (34 people or 14.47%), Hindus (18 people or 7.66%), Muslims (6 people or 2.55%) and one Other (0.43%).

4.3. Procedure

4.3.1. Data Collection

Much of the data was gathered from a 2010 UKZN Psychology second year statistics tutorial. The questionnaires or test batteries, along with the incentive (a small chocolate), were given to twelve tutors, who were each responsible for a group of students from the statistics class. Each tutor was also given a set of instructions to ensure that the majority of the sample took the questionnaire more seriously. They were required to verbally inform the potential participants of their rights to not partake in the research as well as their right to opt out of the research once completing the questionnaire. The tutors were also required to tell the participants how long the questionnaire would take to complete should they work through it at a steady pace. Lastly, the tutors were also responsible for handing out the incentives once they were handed back a completed questionnaire and encouraged not to let the participants take the test home with them as it makes it considerably harder to retrieve. Although all of these instructions were given to the tutors, the method by which they chose to collect the data could not be monitored beyond this point. Some tutors allowed participants to complete the questionnaire at a later stage and all of these were returned at some point. The sample of participants over the age of 50 was gathered independently of the psychology students and was done through a kind of snowball/convenience sampling methodology. The older contingent of the sample was accessed through older members of the principal investigator's family, who subsequently distributed the questionnaires amongst their friends and so on. Of the 50 distributed this way 45 were returned completed.

No participant was monitored during the administration of the questionnaire and there were no stipulated situational or behavioural conditions that the participant needed to meet. Each participant was allowed as much time as possible to complete the questionnaire but, as mentioned, some participants were allowed to leave with the questionnaire and return it completed at a later stage. The conditions for administration were lax only because the items could not allow for cheating and the participants were aware that the questionnaire did not measure any level of performance. However, that being said, most participants did complete the questionnaire within 10 or 15 minutes and were rewarded for their cooperation with the small chocolate (incentive). They were aware of the incentive before they were given the questionnaire, as this was one of the only ways to insure that they would complete a lengthy questionnaire without giving back empty scripts, however, most of the older sample were happy to complete the questionnaire without any incentive.

4.3.2. Data Coding

The questionnaire contained both qualitative and quantitative data, which, after coding, would all be transformed into something that could be analysed quantitatively. This meant that the qualitative data needed to be coded in such a way as to create comparable quantitative statistics.

4.3.2.1. Coding Vignettes

The vignettes were placed right at the end of the questionnaire and were open ended scenarios in which the participant could offer up any response they wish. In order to code these responses

into something usable for analysis a form of content analysis was used. The vignettes were analysed by two independent judges who were both educated on the ideas of order and chaos dominances. Ideally, in a situation with two judges one would like to use Cohen's Kappa in analysis to check the level of agreement between the judges, but it was at this stage that a compromise needed to be made. While Cohen's Kappa is robust and can be used if the judges are given even a large number of options, it becomes tricky to use when the options are no longer independent of one another (summarised in table 1). A lack of independence would mean the calculation of Cohen's Kappa would not give an accurate representation of agreement. For example, if one judge gave a 4 for one participant's answer, while the other judge gave a 3 Cohen's Kappa would register this as a disagreement when it actually represents a partial agreement (table 1). With this in mind the data would be coded with the broader, more Likert-scaling system for the main analysis and then condensed for the calculation of Cohen's Kappa. 0 and 1 would be combined into one group, 3 and 4 into another, and 2 would remain independent.

Table 1
Coding Template Used for Vignette Data

Code	Order Dominance Level	Description
0	Chaos Dominant	Adherence to a contradictory social norm despite the presence of an overarching law
1	Relatively Chaos Dominant	Adherence to a contradictory social norm, but siding somewhat with the overarching law
2	Neutral	Shows no obvious preference for either an order or chaos dominant response or discovers a means of reacting to the scenario without being order or chaos dominant
3	Relatively Order Dominant	Adherence to the overarching law, but siding somewhat with the social norm
4	Order Dominant	Adherence to an over arching law despite the presence of a contradictory social norm

In order to better illustrate how this was done an example of one of the scenarios used has been given along with 5 different answers that justified an agreed upon code (by both judges) of 0-4:

Imagine you and a friend are going out to the movies and you notice R50 falling out of someone's pocket in front of you. They don't notice the money falling out because they appear to be in a hurry. You and your friend also appear to be the only ones who have noticed this. If you are quick, you can still catch the person who dropped the money.

What do you do next in this situation? (Appendix 9.5.1)

Answer 1 (code 0): 'I just take the money and keep it for myself'

Answer 2 (code 1): 'This depends on how close I am with this friend of mine. If we are not that close, I can not give or tell her that she lost her money'

Answer 3 (code 2): 'We pick the money, and vote whether to give it back or take it'

Answer 4 (code 3): 'Quickly find the person and give it to them, but to be honest we surely would hope that he gives us something in return, but if he doesn't, disappointing but its ok'

Answer 5 (code 4): 'Pick it up and catch them'

The overarching or general law in this circumstance is stealing. The money does not belong to them and the law would dictate that the right thing to do would be to return it. However, the law is not visibly being enforced in this circumstance as the participant and their friend are the only ones who have seen this happen, so the social norm is determined by themselves and their reaction decides whether that social norm is in conflict with the general law (chaos dominant) or

in line with it (order dominant). The first answer is coded as 0 because the participant's reaction means that their social norm is in direct conflict with the general law and they would effectively be stealing.

The second answer was trickier as it is clear that the participant has misunderstood the question. However, the participant says that their reaction depends on the relationship that exists between friends. This would lead closer to a neutral response, but they decide to add that if the friendship was not close they would take the money (instead of offering up the alternative of "if the friendship was close"), suggesting that the response is chaos dominant, but that special circumstances have been considered where the ordered option may be applicable.

The third answer was coded as 2 because the reply seemed to answer the question without suggesting whether or not their reaction would be order or chaos dominant. In other words, the answer that they have given is inconclusive and finds a way to answer the question without being order or chaos dominant.

The fourth answer suggests that it should be coded as a 4 because the participant is doing the right thing by returning the money, but even though they appear to be adhering to the general law they do not appear to be all that happy about it and is instead coded as a 3. A reward for returning a denomination of that amount would either be insultingly small or be an impractically large proportion of the original money. It would be unlikely that the participant would be expecting an insultingly small reward, but rather one that the person getting their money back would think is impractically large. The participant is then adhering to the general law, but on their own terms, suggesting that their social norm is somewhat in conflict.

The fifth and final answer is coded as a 4 for the plain and simple reason that the participant is doing exactly what the general law (don't steal) is dictating to them. The social norm they have created falls directly in line with the general or overarching law.

4.3.2.2. Quantitative

The ODS had already been designed to match the coding schemes of the NDS and the TDS and so the remaining three quantitative scales were all given codes of 0, 1, or 2 depending on the nature of the answer. The ODS was scored in an order dominant direction (i.e. an order dominant response was coded as a 2, a 'not sure' response was coded as a 1 and a chaos dominant response received a 0), the NDS was scored in a negativism dominant direction (i.e. a negativistic or rebellious response received a 2) and the TDS was scored in a telic direction (i.e. a telic response received a 2).

Demographic data was given as open ended answers, but the responses given were very limited making the coding process for demographics relatively easy. Table 2 summarises how this data was coded. All data was entered into MS Excel as it was given by the participant. The data was coded through the use of a series of IF formulas into the codes given above. The data was then prepared for analysis using the statistical program R (version 2.12.1), by removing all specious information (i.e. unreturned scripts, incomplete scripts) and editing all data that had been entered incorrectly. The majority of the analysis was done in R (version 2.12.1) and the remaining exploratory data analysis was done using the Excel spreadsheets.

Table 2
Demographic Codes

Age	Race	Gender	Religion
1 = 18-20	1 = Black	1 = Male	1 = Christian
2 = 21-25	2 = White	2 = Female	2 = Muslim
3 = 26-30	3 = Indian		3 = Hindu
4 = 31-35	4 = Coloured		4 = Ambivalent
5 = 36-40	5 = Other		5 = Other
6 = 41-45			
7 = 46-50			
8 = 51 and over			

4.4. Ethical considerations

No major ethical problems arose in the data collection or data coding process as the scale itself does not ask penetrating questions of its participants. The ODS, NDS, TDS and the vignettes are not sensitive questionnaires, even though the items centre around petty crime most participants would be aware that the punishment for such crimes is minor at best and, thus, they are not being asked to potentially incriminate themselves in any way.

However, should the participant feel that the questions being asked of them are damaging or offensive, measures were in place to insure that they did not have to participate in the research process. The participants were required to sign a form of informed consent to participate. They were informed of their rights to choose to participate in the research process and told how they may opt out of the study at any time. This was done by giving each of the 250 questionnaires a number from 1 to 250 with which the participant could identify themselves. They were also given the number and e-mail address of the principal investigator and told to send their reference number via sms or e-mail if they wish for their data to be removed from the study. Over and above this the participants were not required to give out any information that could potentially

identify them; only their age, gender, race and religion. Only the principal investigator had access to any of the completed questionnaires and once the data was coded into Microsoft Excel, the participants were essentially unidentifiable. No participant withdrew from the study.

Although the addition of an incentive is a heavily debated topic in social science research as to the effect it may have on the data gathered, the nature of the incentive in the research is not one which could affect outcomes. The participants would not have been a part of the research because the incentive would help them in any way as it can be assumed that the majority of the participants are financially stable to some extent and do not suffer from any problems that a small chocolate would quickly fix. In this regard the incentive was no more than a small enticement that offered no long-term benefits for the participant and, therefore, did not attract participants based on a need for the incentive.

It is also important to consider the nature of the data collected. The data is not false and no data was fabricated in the coding process. The data was only excluded from the data analysis procedure if the questionnaires were incomplete. Of the 250 questionnaires distributed 235 were returned complete and of those 15 remaining questionnaires only two were removed from the study because they were incomplete, the remaining 13 were not returned.

5. Results

5.1. Demographics differences

The first part of the analysis procedure involved discovering if there were any significant differences between any of the gathered demographic information in their ODS scores. This was done only on the ODS scores as it was important to identify potential biases in the data, which could be controlled for in further analysis. The ODS score of the participant is the dependant variable in all cases.

5.1.1. Age

The eight categories of age classification were scaled down to six when it was discovered that no one between the ages of 36-40 and 46-50 had completed the questionnaire. The remaining six categories (represented by figure 4) were heavily clustered into three major age groups, namely 18-20, 21-25 and 51 and over, while the remaining three, 26-30, 31-35 and 41-45 totalled four participants.

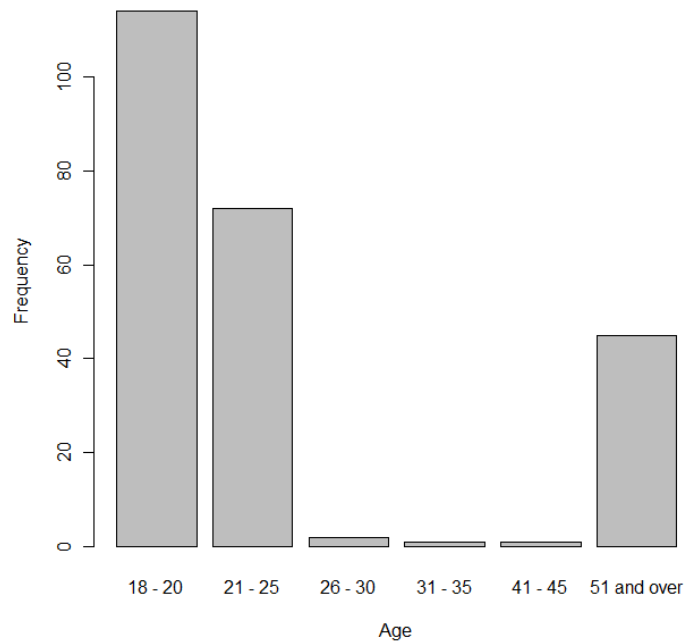


Figure 4: Bar graph of Age distribution

The means and standard deviations of the six categories are given in table 3. A simple examination of the means as they are would suggest that significant differences may lie between the younger and older age groups.

Table 3
ODS descriptives by Age

Age Group	Mean	SD	N
18-20	29.73	7.2	114
21-25	28.74	7.07	72
26-30	38.5	6.36	2
31-35	39	NA	1
41-45	33	NA	1
51 and over	39.29	4.73	45

After confirming that the assumption of homogeneity was not violated ($\text{Pr}(>F)=0.114$) and running the ANOVA procedure in R (version 2.12.1) it became clear that there were significant

differences between at least one of the groups well beyond an alpha of 0.05 (F value=16.654; $\Pr(>F)=5.053 \times 10^{-14}$). Comparisons crossing the central line (representing a mean difference of 0 for any comparison) are highly unlikely to show significant differences and the first impression one gets by looking at the plot of confidence intervals (figure 5a) is that the participants over the age of 50 differ significantly from the participants between the ages of 18 and 25. The tiebreaker plot (figure 5b), which works on the same principles (similar to the subsets created by a Tukey HSD), confirms that this is indeed where the differences lie (the red lines, as well as the fact that they do not cross the central line, confirm that these are significant differences).

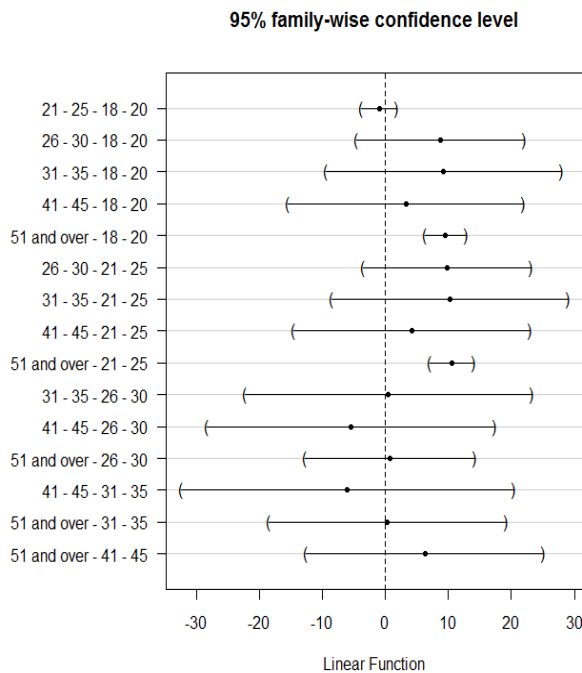


Figure 5a: Age Plot of Pairwise Mean Comparisons

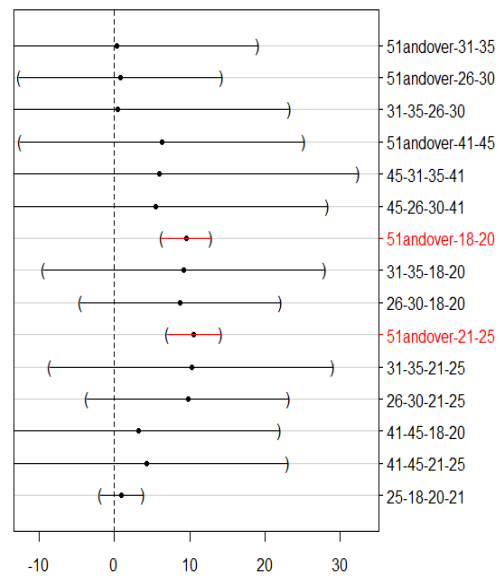


Figure 5b: Age Tiebreaker Plot

5.1.2. Gender

The gender demographic showed a far larger proportion of females than males, creating space for the assumption of homogeneity to be violated. The Levene's test for homogeneity did not

reveal that the variances between the two groups were significantly different ($\Pr(>F)=0.3581$) and the analysis could continue forward. A simple independent samples t-test was employed to test for these potential differences of which none were found ($\Pr(>F)=0.2436$). Looking at the means of each group, 31.72 for females and 30.33 for males, it becomes even clearer that even with a much larger sample it is unlikely a difference would exist.

5.1.3. Race

The ‘race’ categories that emerged from the data were Black, White, Indian and Coloured; represented by figure 6 and are discussed in more detail earlier.

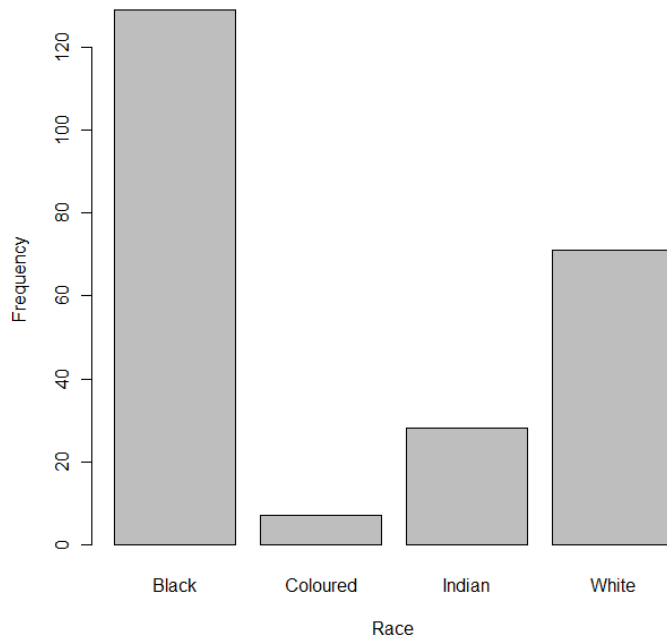


Figure 6: Bar Graph of Race Distribution

The race category ‘Coloured’ represented a small and unrepresentative section of the population and was treated with care throughout the remainder of the ANOVA. The mean ODS score for the

race category 'Black' was 28.77 with a standard deviation of 7.23, the mean ODS score for the 'White' race category was 36.77 with a standard deviation of 6.06; the mean ODS score for the race category 'Indian' was 31.18 with a standard deviation of 7.01; while the 'Coloured' race category had a mean of 25.71 and a standard deviation of 8.71.

After confirming homogeneity ($\Pr(>F)=0.207$) and running the ANOVA it was very clear that significant differences existed between at least one of the groups (F value=22.118,

$\Pr(>F)=1.259 \times 10^{-12}$) with the plot of confidence intervals (figure 7a) suggesting that these differences may lie between the white race category and the remaining categories. However, in order to identify if these differences were genuine (especially considering the nature of the coloured sample) it was necessary to use the multiple comparisons tiebreaker plot (figure 7b).

Even with the small sample of people from the race category 'Coloured' the differences between the race category 'White' and the remaining categories all proved significant (again indicated by the red lines and the fact that they do not cross the central line representing a mean difference of 0).

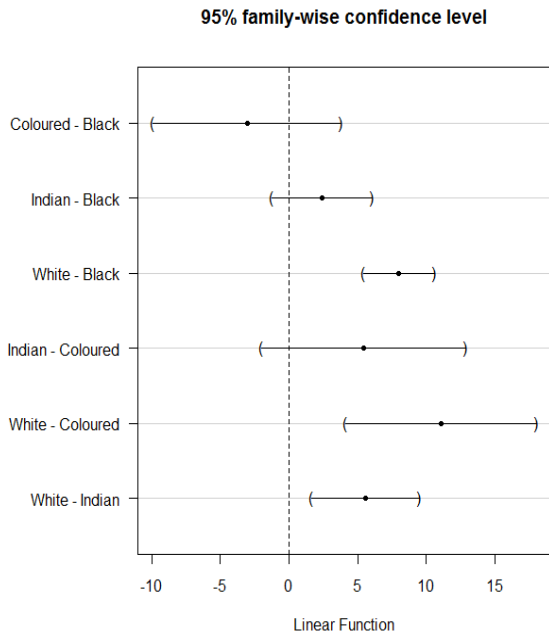


Figure 7a: Race Plot of Pairwise Mean Comparisons

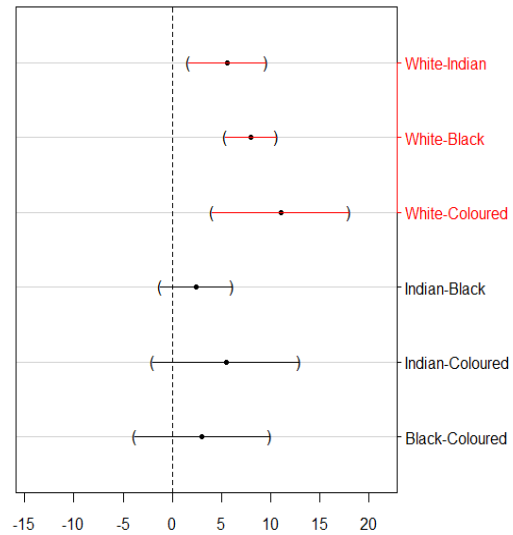


Figure 7b: Race Tiebreaker Plot

5.1.4. Religion

The participants were not restricted by pre-defined categories for religion as South Africa has a multitude of diverse religions to which people subscribe. Religious views that included atheist, agnostic, spiritual and any other view that does not adhere to any form of organised religion were collectively named ‘Ambivalent’. The other major religions that emerged from this were Christianity, Muslim, Hindu and one other (Judaism). Many participants gave their relevant denominations as their religious view, but these were all gathered into these four major categories. The sample was primarily Christian as figure 8 shows; the mean ODS scores and their standard deviations are given in table 4.

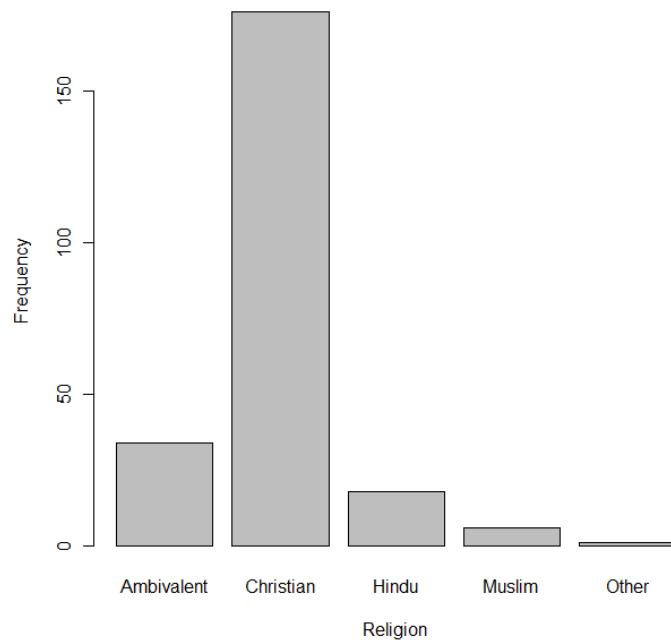


Figure 8: Bar Graph of Religion Distribution

Table 4
ODS Descriptives by Religion

Religion	Mean	SD	N
Ambivalent	32.53	6.94	34
Christian	31.19	8.06	176
Hindu	30.94	7.83	18
Muslim	31.5	5.24	6
Other	34	NA	1

Levene's test was not significant ($\text{Pr}(>F)=0.2048$). The observed means do not appear to be significantly different from one another (barring the one other, whose sample size is too small to bear any significance) and the ANOVA confirmed this ($F \text{ value}=0.251$, $\text{Pr}(>F)= 0.9088$). The plot of the confidence intervals (figure 9a) and the accompanying multiple comparisons tiebreak plot (figure 9b) suggest that no particular comparison shows any signs of approaching significance should the sample size be increased.

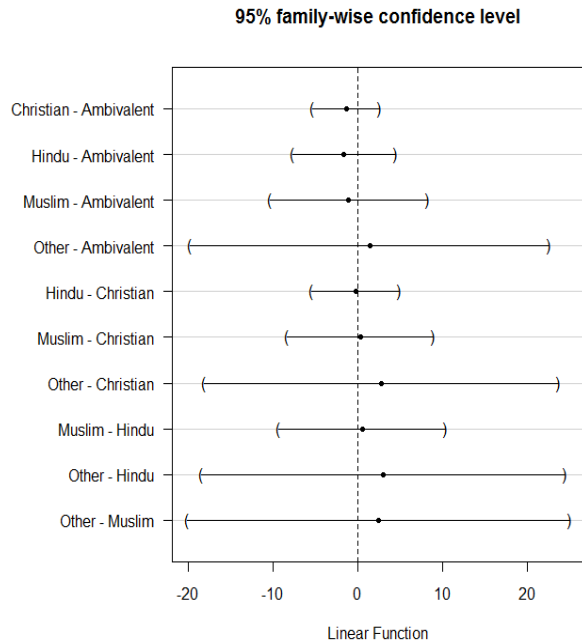


Figure 9a: Religion Plot of Pairwise Mean Comparisons

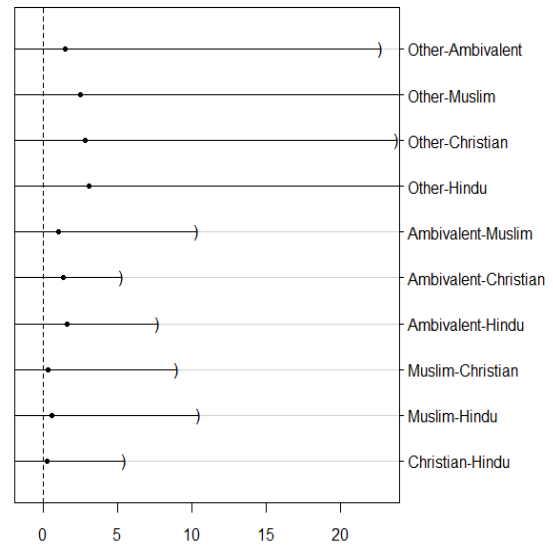


Figure 9b: Religion Tiebreaker Plot

5.2. Construct Validity

In order to establish that the scale had some form of construct validity the ODS results of each participant needed to be able to accurately predict the nature of the answers they would give on the vignettes. The assumption was that order dominant individuals would remain order dominant and vice versa for chaos dominant individuals, but that people who were in the middle range would be more likely to switch between order and chaos dominance depending on the way the vignette had been written (i.e. in an order or chaos dominant direction). The procedure to establish this used exploratory data analysis (EDA).

The first thing that needed to be established is whether or not the two judges were in agreement over the way each participant's vignettes were coded. The Cohen's Kappa for the first scenario was 0.7985 and the distribution of the two judges' agreement is given in table 5a. The Cohen's

Kappa for the second scenario was 0.7772 and the distribution of the two judges' agreement is given in table 5b. Most literature suggests that a Kappa over 0.75 (Fleiss, 1981) is considered excellent agreement between the two judges and both of these Kappas fall above that benchmark.

Table 5a
Agreement in Scenario 1

Scenario 1	Judge1		
	1	2	3
1	34	1	5
2	7	15	4
Judge2	3	4	3
			162

Table 5b
Agreement in Scenario 2

Scenario 2	Judge1		
	1	2	3
1	126	4	4
2	0	4	3
Judge2	3	16	3
			75

The second thing that needed to be established was whether or not significant differences existed between the five different scenarios pairs. In order to do this the total score from each judge on a scenario pair was added up, creating a potential range from 0-16. The means and standard deviations of the scenario pairs are given in table 6.

Table 6
Descriptives of Total Judge Score by Scenario Pair

Scenario Number	Mean	SD	N
1	9.07	5.76	46
2	8.96	4.02	47
3	9.96	2.6	47
4	8.37	3.02	46
5	9.92	4.25	49

The resulting ANOVA was not significant (F value=1.3028, Pr(>F)=0.2698) and the plot of confidence intervals (figure 10a) and tiebreaker plot (figure 10b) confirmed this and that none of the scenarios would reach significance soon if the sample size were increased.

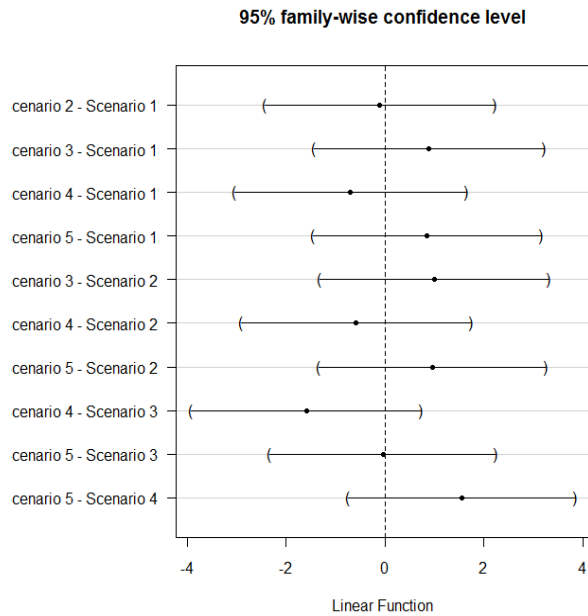


Figure 10a: Scenario Plot of Pairwise Mean Comparisons

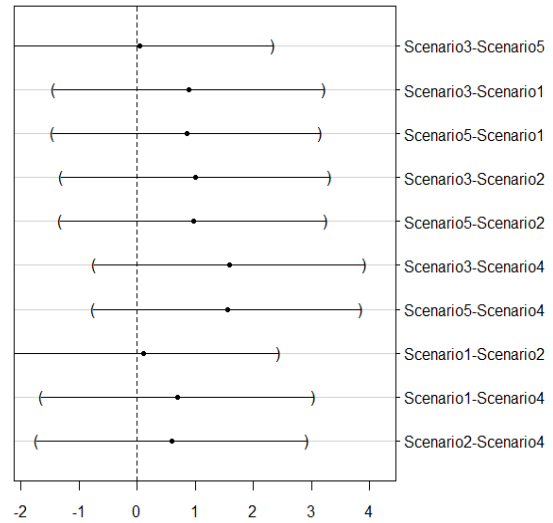


Figure 10b: Scenario Tiebreaker Plot

With these two factors established it was possible to continue on with the exploratory data analysis to discover whether construct validity existed. A simple correlation would not be sufficient in determining whether or not this validity existed as the idea was to discover whether or not the people in the middle range were switching more often than the extreme scorers. It was first necessary to decide how the scores would be divided into low, middle and high scorers. This was done by using the descriptive statistics of the scale. The mean for the ODS was 31.38 with a standard deviation of 7.79. It was decided that participants who fell outside the mean score plus one standard deviation would be considered the high and low scorers, which meant that participants with scores less than or equal to 23.59 or 24 and scores greater than or equal to 39.17 or 39 would fall into those categories. With this in mind the low scoring group consisted of 48 people, the middle scoring group consisted of 146 people and the high scoring group consisted of 41 people.

Ideally these groups would be distinct from one another, which would prove that the ODS score predicts the vignette score. An ANOVA was performed to check for this. Below are the means and standard deviations for the ANOVA (table 7) as well as the pairwise mean comparison plot (figures 11).

Table 6
Low, Middle and High Group Descriptives by Vignette Score

Group	Mean	SD	N
High	10.93	3.66	41
Low	6.4	4.49	48
Middle	9.74	3.61	146

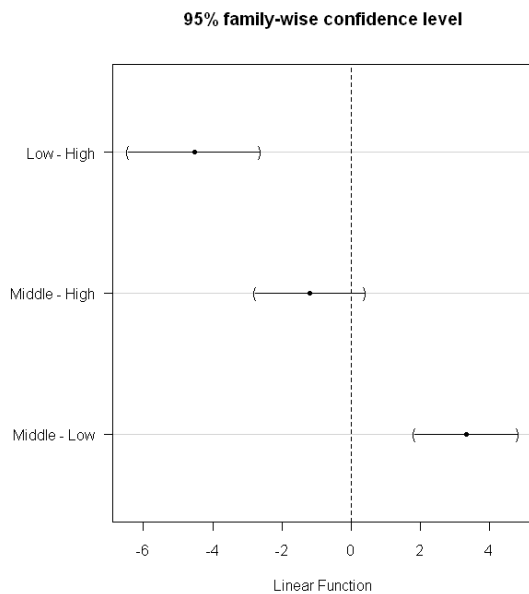


Figure 11: Group Plot (by Vignette Score) of Pairwise Mean Comparisons

The ANOVA on the vignette score came up highly significant (F value=18.64, $\Pr(>F)=3.11 \cdot 10^{-8}$). The tiebreaker plot (figure 12) clearly shows that all three groups differ significantly from one another, with the exception of the high and middle scorers comparison.

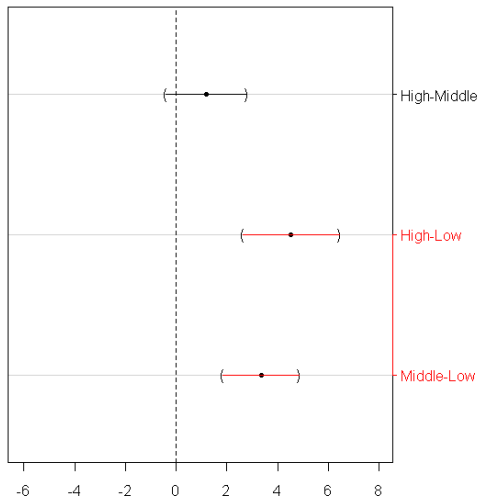


Figure 12: Group (by Vignette Score) Tiebreaker Plot

The last thing to establish was the percentage of people who switched from order dominance to chaos dominance from one scenario to another in each of the three groups. Each scenario pair was set up so that the first scenario of the pair would lean in an order dominant direction and the second scenario would lean in a chaos dominant direction. Thus, adding the score of each judge for each scenario resulted in two data points ranging from 0-8 for each participant. These data points were treated like X and Y co-ordinates. The assumption would be that each group would cluster differently on the theoretical scatterplot these points would produce (a real scatterplot is very little help and difficult to interpret as there are a limited number of possible points on the graph). The order dominant scenario was plotted on the y-axis, while the chaos dominant scenario was plotted on the x-axis. The distance from the origin (0,0) to each point was calculated (Hypotenuse variable; the significance of which will be discussed later) as well as the angle it created off the x-axis (Angle variable). This meant that the greater a person's angle is the more order dominant they were in their approach to the vignettes.

Ideally, in this situation, one would use an ordered logistic regression and use these two variables to model the high, middle and low ODS scorers. There were two major reasons why this was not done. The first of these was the small sample size; a logistic regression with so few participants would battle to tease out any underlying relationships that may exist. The second was that the relationship between the two variables is completely different and consequences of each require that they be interpreted independently of one another. The alternative was to tackle each in two separate ANOVAs. The results of the first ANOVA on the hypotenuse variable yielded highly significant results (F value=20.619, $\text{Pr}(>F)=5.728 \times 10^{-9}$). Although only just beyond the significant 0.05 mark of alpha, the second ANOVA on the angle the hypotenuse line created was also significant (F value=3.3985, $\text{Pr}(>F)=0.0351$). The means and standard deviation of the ANOVAs are summarised in table 7 below. Their respective pairwise mean comparison plots as well as their tiebreaker plots are also given (Figures 13a,b and 14a,b).

Table 7
 Low, Middle and High Group Descriptives by Hypotenuse and Angle Variables

Group	Hypotenuse		Angle		n
	Mean	SD	Mean	SD	
Low	5.33	3.54	19.2	24.8	48
Middle	7.85	2.27	27.91	25.74	146
High	8.38	2.25	32.47	22.68	41

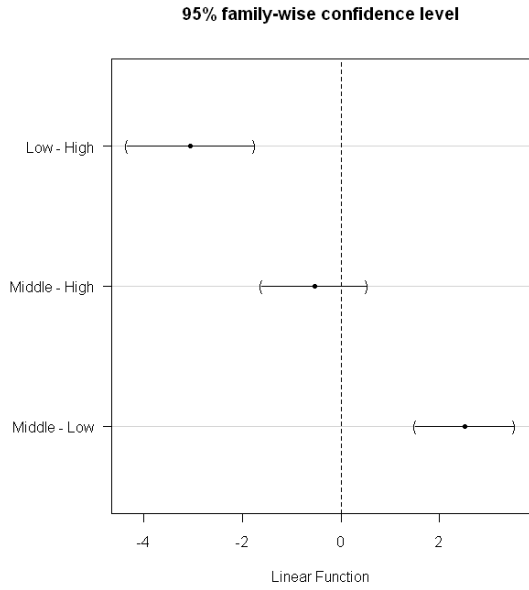


Figure 13a: Hypotenuse Plot of Pairwise Mean Comparisons

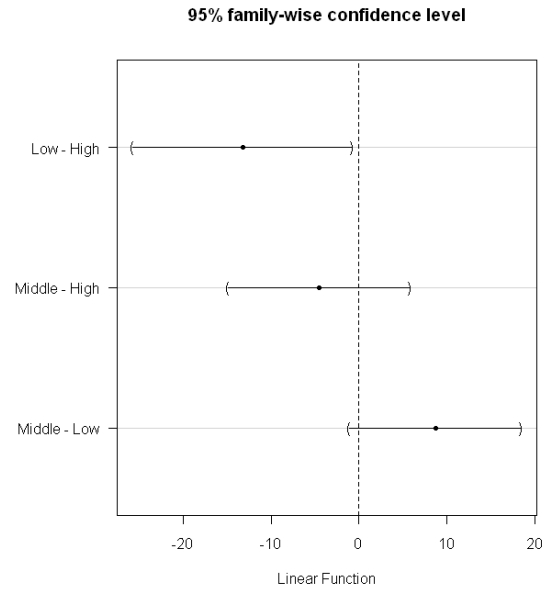


Figure 13b: Angle Plot of Pairwise Mean Comparisons

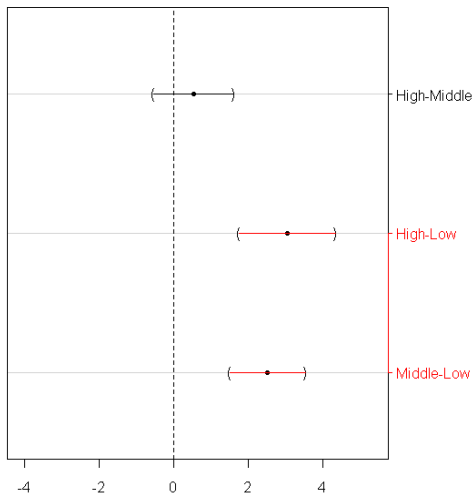


Figure 14a: Hypotenuse Tiebreaker Plot

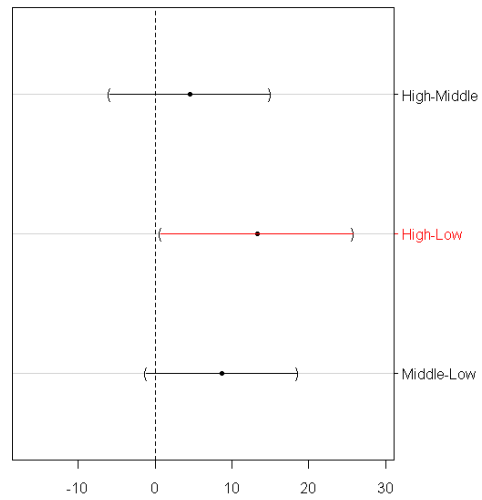


Figure 14b: Angle Tiebreaker Plot

5.3. Predictive Validity

As it was impossible to get a separate sample in which to test the hypothesis that older people will score higher on the ODS than younger people, a subsample of the main sample was taken instead. The 45 people over the age of 50 were used as the sample for the older generation and, as the demographic differences had already proven significant, race and religion was used to match the remaining sample to the older group. This group consisted of 26 participants under the age of 50, all white and held either Christian or ambivalent religious beliefs. The two samples were simply subjected to a test for homogeneity and a one-tailed independent samples t-test to check for significant differences. The Levene's test for homogeneity (F value=0.6831, Pr(>F)=0.4114) was not significant at an alpha level of 0.05 and the t-test revealed highly significant differences between the two age groups (p-value= 3.428×10^{-07}) with the younger sample's mean (32.42) being significantly lower than the older sample's mean (39.29).

5.4. Convergent and Discriminant Validity

The NDS and TDS were included in the questionnaire to establish these two forms of validity, each for their own reasons. The NDS was included for establishing convergent validity as it is theoretically very similar to the ODS and the measured trait is also fairly similar. The TDS was included to help establish both convergent and discriminant validity as it is theoretically very similar but rather different with the measured trait. This was done by simply correlating the ODS with the NDS and the ODS with the TDS. The correlation between the ODS and the NDS was -

0.5823 (p-value= 2.2×10^{-16}) while the correlation between the ODS and the TDS was 0.214 (p-value=0.0009597), both easily reaching significance at an alpha level of 0.05.

5.5. Establishing Validity magnitude

The effect sizes for all the different analyses were calculated using GPower 3.0.10. The hypotenuse variable in the analysis establishing content validity only has an interpretable value when explained in conjuncture with the angle variable. The angle variable can, however, be interpreted alone and an effect size calculation on this ANOVA would yield an interpretable result. The effect size for this procedure was 0.1707. This is considered a small effect size for this test, according to GPower, and is below the calculated required effect size of 0.258 for this comparison.

The effect size statistic on the independent samples t-test for the measure of predictive validity was calculated at 1.3121. GPower considers anything beyond 0.8 as a large effect size for this test and it is above the calculated required effect size of 0.8185 for this comparison.

The effect size statistic on the Pearson's correlation co-efficient between the NDS and the ODS for the measure of convergent validity was calculated at 0.7631. GPower considers anything beyond 0.5 as a large effect size for this test and it is above the calculated required effect size of 0.2104 for this comparison. The co-efficient between the TDS and the NDS was calculated at 0.4626 which is considered a medium effect size, but is above the required effect size of 0.2104.

5.6. Re-establishing Norms

The reliability study (Jones, 2009) establish the scale norms at that stage as having an absolute range of 0-50 and an observed range of 4-46; the scale's mean was 29.52 with a standard deviation of 7.96 and a reliability coefficient 0.728. By recalculating these norms with the new dataset the scale had an absolute range of 0-50, an observed range of 8-48, a mean of 31.38, and a standard deviation of 7.79. The recalculated reliability coefficient was 0.7507 and was standardised to 0.7527 with only one item, 11, bringing the alpha level down (alpha increased to 0.7593 when item 11 was removed, but the increase was far too small to justify dropping the item). In comparison with the two published scales' reliability coefficients the ODS performed the best, with the NDS having an alpha of 0.7137 (standardised to 0.7145) and two items bringing alpha down and the TDS having an alpha of 0.6642 (standardised to 0.6613) and nine items bringing alpha down. The mean for the first set of scale descriptive was a lot lower than the new mean of 31.38. This was potentially due to the addition of 45 people over the age of 50 who were shown to have significantly higher means. So these participants were removed from the data and the norms were recalculated with a range of 8-45, a mean of 29.51, and a standard deviation of 7.18. The first set of descriptives will be given as the most accurate ones, as the original sample had sampled from a wider spread of the population (i.e. included 45 members of the population outside the university environment).

6. Discussion

The focus of the discussion is on two primary aspects of the scale; namely, what conclusions can be drawn from the set of results that the research achieved and what are the implications for the various theories outlined earlier, according to these results and the final scale. The discussion will also briefly look at the weaknesses of the research and the possibility of future research using the now fully validated scale.

6.1. Interpreting results

6.1.1. Demographics differences

The demographic information gathered in the reliability study (Jones, 2009) was collected in the hope that no large or significant differences may occur in the data, as the focus at that point was on the construction of the scale and the items, not necessarily the participants, and this could suggest that discriminating items discriminated based solely on a particular dominating demographic. For example, a large portion of the sample was Christian a large difference between the Christian population and the remaining sample could mean that items which had previously discriminated well would no longer, when the religion variable was controlled for. However, the reliability study identified significant differences between male and female participants with female participants scoring significantly higher than the male participants. At the time this difference was put down to chance, regardless of the test significance, because the

sample of males and females was not balanced. The remaining demographics did not come up significant either and for some reason there was no significance test attempting to discover differences between people of different ages. The assumption in the validity study was that the majority of the demographic information would again return no significant differences, although the category of age was an unknown variable.

6.1.1.1. Age

The reliability study had left the age demographic in a raw data form and analysis on age data in a sample that small (177 participants) would have resulted in outlying ages being considered as their own separate sample and the resulting ANOVA would yield highly unreliable results. However, for the purposes of this research the gathered age data was coded to fit into one of 8 categories and the result meant that the data was testable for differences. The results of this analysis had interesting implications for the analysis of predictive validity that would be done later. The descriptives of this category alone were of interest to the research. Although it was clear that the means of the 18-20 year old and the 21-25 year olds were likely to be significantly different from the 51 year olds and over, the standard deviations revealed another interesting story. The deviations for the first two groups were stable at around 7 while the older group had a far smaller deviation of around 4.5. This suggested that the older generation was consistently achieving higher means and that their response format may vary very little between participants. The ANOVA showing that these significant differences did indeed exist between the older generation and both the 18-20 year olds and the 21-25 year olds just seemed to confirm what already appeared to be evident from the descriptives.

6.1.1.2. Gender

As gender had been the only significant demographic in the original reliability study, it was assumed that a similar pattern could emerge in this research and it was nearly included as a potential variable where a difference would exist for measuring predictive validity. The proportion of males to females was surprisingly similar to the original reliability study, even with the inclusion of 45 participants outside the university contexts, further backing up a repeat occurrence. The t-test did not, however, show any significant differences even though the mean pattern was similar (females had a higher mean).

There are two potential reasons for this observed result. The one is that the demographic differences seen in the original study were calculated using the full array of items at that stage. Only 25 of those 65 original ODS items were carried through to the final scale, meaning that the observed difference could be as a result of the other 40 items that did not make the final scale. The other reason this could have happened is based on the assumption the research made about older participants scoring higher than younger participants. The mean for older males (38.44) and females (39.76) alone also showed a similar pattern to the sample as a whole, but the older contingent of the male part of the sample was 28.1%, while the older contingent of the female part of the sample was only 16.29%. This means that a larger percentage of older participants were in the male part of the sample, effectively pushing the overall male mean up more than the overall female mean. An independent samples t-test (Levene's $\Pr(>F)=0.7576$; t-test p-value=0.01813) of only the younger contingent of the sample revealed the same significant

differences that were observed in the reliability study, with the male mean (27.17) being significantly lower than the female mean (30.15). Although the initial analysis did not reveal these differences, it may be worthwhile to investigate why they seem to occur in future research.

6.1.1.3. Race

The original reliability study did not find any differences between the different race categories and it was surprising to discover such highly significant differences in this research. The patterns were similar to those of the original research with 'White' having the highest mean and 'Coloured' having the lowest mean. In the original study 'Black' and 'Indian' were nearly identical, but in this research the 'Black' mean was quite a lot lower than the 'Indian' mean. The only differences arising from the ANOVA were between the 'White' contingent and the remaining categories, with all these being highly significant. The reason this was happening was exactly the same as the reason the gender differences did not show initially; the older participants. Upon closer inspection, the entire sample of older participants was discovered to fall in the 'White' category, which would push the 'White' mean up significantly. Repeating the ANOVA without the sample of older participants (F value=3.1423, $Pr(>F)=0.0265$) revealed that although there were still significant differences between the race categories, it was unable to identify any significant comparisons. Table 8 shows the means and standard deviations of the race categories without the older contingent and its respective significance plots (figure 15a and 15b).

Table 8
 ODS Descriptives by Race (no participants over 51)

Race	Mean	SD	N
Black	28.77	7.23	129
Coloured	25.71	8.71	7
Indian	31.18	7.01	28
White	32.42	5.69	26

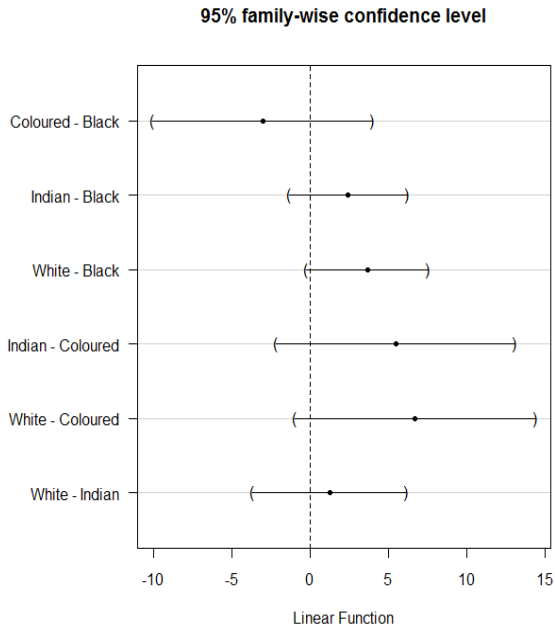


Figure 15a: Race Plot of Pairwise Mean Comparisons (no participants over 51)

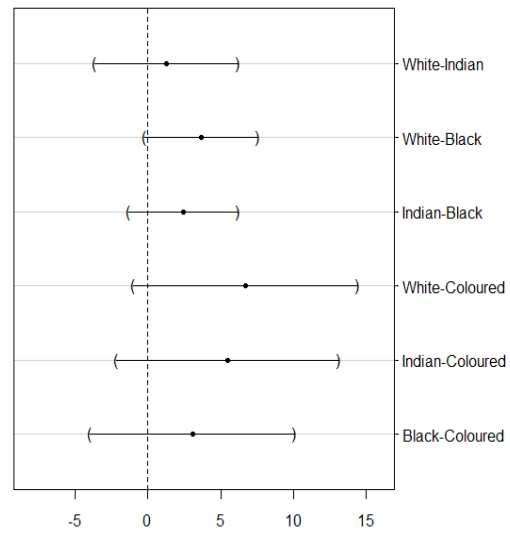


Figure 15b: Race Tiebreaker Plot (no participants over 51)

Although none of these comparisons are significant, the two plots can identify that the overall ANOVA model is most probably significant due to the differences between the ‘White’ participants and the ‘Indian’ and ‘Black’ participants. Statistically, however, there are no significant differences between these groups once the older contingent has been removed from the sample.

6.1.1.4. Religion

The spread of the sample by religion was fairly similar to the spread observed in the original study and appeared to be similar to the spread one would expect in a country like South Africa. The means of each group did not vary greatly and the ANOVA showed that there were no significant differences between even one of the groups. However, the one interesting thing to note from these means is that the ‘Ambivalent’ mean suggests this is the most order dominant category. One would assume that subscribing to a religion comes with a more extensive set of rules to obey and an equally extensive list of punishments and that this alone would cause religious people to be more order dominant. With that being said, the observed differences are not significant and no such assumptions can be made without better substantiating evidence, both theoretically and statistically. It would be an interesting topic to pursue in the future as the differences may be approaching significance with larger and more equally distributed (in terms of religion) samples.

6.1.2. Construct Validity

Establishing construct validity was by far the trickiest process in the analysis procedure. If one were to simply total up the judges scores and correlate them with the total ODS score this would get nowhere towards establishing whether the ODS had predictive validity at all. One might expect that a high ODS scorer would get a high score on the vignettes at the end so that there would be high correlations between the two, but this assumption would only partially be true. For reference sake, the correlation between the two (ODS scores and the vignettes) is 0.337 (p-

value= 5.922×10^{-8}) and although it is significant it fails at establishing anything, because the nature of this calculation is far more complex. There is no indication of who is scoring high or low and whether the high and low scorers maintain that pattern when exposed to the vignettes (i.e. no matter which way the scenario leans, they remain in their respective dominant states). It also gives no indication of which part of the population is switching states and how those people who are switching states did on the ODS.

Before continuing with the procedure it was important to check that both the judges were in agreement with one another and that the scenario pairs did not differ significantly from one another. As this procedure was relatively exploratory, these two parts of the procedure would be the equivalent of testing for homogeneity of variance in a t-test. If the judges did not agree on the coding then the scenarios may have more than one pole and offer up more than just order and chaos dominance as the levels with which one could react to the scenario. In the same light, if the scenarios were significantly different from one another, then they would not all be measuring the same levels of order dominance, i.e. one scenario may lean so far in the order dominant direction that it would be nearly impossible for the participant to react in any other way. The reason that multiple scenarios were given was because the point of this exercise was to predict an order dominant reaction according to the participants' ODS score. Having only the one scenario pair would have essentially been a test of the scenario pair rather than the trait of order dominance itself. Having the multiple scenario pairs also establishes the existence of order dominance on multiple levels. As the agreement between the judges was very high over a sample of 235 with 3 options for each code and there were no significant differences between any of the scenario pairs, the effort of discovering construct validity continued.

The nature of statistics means that the sample is always assumed to come from one population, until proven otherwise, and it was this assumption that was used to determine what defined a high and low scorer. Two standard deviations away from the mean is diving into territory where the high and low groups would both be very small and considered significantly different (at an alpha level of 0.05) from the population and the assumption remains that the data gathered came from just the one population. Thus, a standard deviation of 1 was decided upon and all participants in the sample scoring higher or lower than the mean \pm one standard deviation were considered the high and low scoring groups respectively. The remaining group were considered the middle-scorers; the group with the least obvious dominance and, therefore, the group most likely to switch between scenarios.

Even though the ANOVA with this new variable of high, middle, and low scorers on the judges ratings came up significant it still did not answer the question of whether or not the middle part of the sample was switching states more than the high and low scoring groups were and, therefore, one more bit of data analysis was required to properly establish construct validity. By treating the judges scores for each scenario as X and Y co-ordinates the distinct patterns between the high, middle and low scoring groups could be more readily identified. With the chaos dominant scenario plotted on the x-axis and the order dominant one on the y-axis it could be hypothesized that a high ODS scorers would produce a greater angle off of the x-axis. This was because of the hypothesis that a high ODS scorer would be unlikely to switch states, i.e. they would score high on the order dominant scenario and low on the chaos dominant one. In the same respect, a low scorer or chaos dominant individual would be assumed to have a far smaller

angle off the x-axis. The magnitude of the switch is determined by the length of the line from the origin to the x,y co-ordinate. The two had to be addressed separately (not in the same model) because they would have adverse effects on each other and do not measure the same things. A middle scorer who is switching with a large magnitude, could end up with similar results to a high scorer who isn't switching with a smaller magnitude. They are still, however, interesting in their own right. Nonetheless, while the relationship that exists between the two is complex to show statistically, each model can be explained separately to reach one conclusion.

The ANOVA on the angle of the overall regression line for each group (forced through the origin) was significant and showed the pattern that was hypothesized. The high scoring group had the biggest angle, the low scoring group had the smallest angle and the middle group fell in between suggesting that they had the greatest propensity to switch. Although only the mean difference between the high and low scoring groups was significant, the observed pattern is evidence that a larger sample would have found significant differences between all the groups. This confirms the hypothesis that high and low scores switch less frequently than the part of the population who are not strongly dominant in one way or another.

The second ANOVA on the hypotenuse variable also yielded significant results. The high mean observed on the high scoring participants suggested that they were not only consistent in not switching, but that they remained highly order dominant on both scenarios. This further supports the differences in angle already discovered between the high and low scoring groups. The high mean observed with the middle group was very interesting, as this suggested that the switches that were being observed in this group were most likely from one extreme to another (say scoring

as high as 8 on both scenarios), as opposed to small less meaningful switches (say getting a judges of score of 2 on both scenarios). The magnitude that was observed for the low scoring group was significantly lower from the other two. This suggests that low scorers, while chaos dominant, were not actually very chaos dominant in their replies. This is possibly because the sample was primarily taken from a university population who are unlikely to be as openly chaos dominant. If one could identify a population that was particularly chaos dominant the magnitude of their chaos dominance might increase and further support the consistent nature of their responses.

6.1.3. Predictive Validity

Trying to test for just a part of the multitude of available types of validity with just the one sample has already proved problematic and having to re-sample from the gathered data to gauge the level of predictive validity is another weakness in this research. The t-test given in the results section of this research was one of a few that was run before the role the demographic information was accounting for was included. The younger groups were all created the same way and each one was highly significant when compared to the group of older people. The older contingent has already played havoc on the analysis in some respect, proving time and time again that their considerably higher mean is more than just significantly different from the remainder of the sample. It would have done just as well to include any one of the t-tests done in this way in the final results as they all were significant. In the end the sample was match to the older contingent, by means of race and religious affiliation. The t-test performed on this data still

yielded a significant result and is, thus, safe to assume that this form of validity has been established; information that is further backed up by the effect size calculation for this test.

6.1.4. Convergent and Discriminant Validity

There is no existing measure of order dominance. There are currently no published scales available that measure a directly comparative personality trait, but the NDS and TDS are similar scales in their own respects. The NDS is a very similar scale in both its theory and its personality trait and was included in the research for the purposes of measuring convergent validity. The hope was that it would correlate quite highly with the ODS, but not so high as to make it difficult to distinguish the differences between the two personality traits. It was also hoped that this correlation would be negative as the two scales are scored in opposite directions (i.e. for the ODS to correlate positively it would need to be scored in a chaos dominant direction). The observed correlation of -0.5823 was nearly perfect in satisfying the conditions to establish convergent validity as this meant that the results of the NDS could account for 33.91% of the variance in the ODS scores.

The TDS was included to measure both convergent and discriminant validity as it was theoretically similar (convergent), but the measure itself was quite different (discriminant). The correlation in this case was hoped to be about half of the correlation between the ODS and the NDS and in a positive direction. The resulting 0.214 was fairly close to this and accounted for only 4.58% of the variance in the ODS scores, suggesting that the common ground that the two scales share is only in their theory, something that is unlikely to be highly emergent in a simple correlation.

While these two correlations played their role in establishing convergent and discriminant validity for the scale, it was the interaction of both that confirmed it. By using the existing knowledge that the NDS and TDS are theoretically similar but different in what they measure and by identifying that chaos dominance is very similar to negativism dominance, it was possible to hypothesize how the two would manifest their relationships with the ODS. In achieving not one, but both of these ideal hypothetical relationships the ODS has established that it has both convergent and discriminant validity.

6.1.5. Establishing Validity magnitude

The effect sizes calculated are all large enough to establish a medium or large magnitude for the various types of validity, barring that of general construct validity. According to Westen and Rosenthal (2003), using the effect sizes to conclude on validity would mean that predictive, convergent and discriminant validity had been safely established, but not a general level of the validity of the construct. However, the analysis of construct validity was exploratory and the effect size statistic was only an estimate of the final test in a long string of tests. Every outcome in the tests preceding this final test had had a result which allowed the analysis to continue. The preceding tests were not a series of tests testing the effect of the same outcome each time, but a series of tests contingent on the test before it in order to achieve one outcome. Although the effect size in the calculation of construct validity was small, the analysis was exploratory and the fact that the preceding tests all satisfied the required assumptions suggests that there was a level of construct validity present.

6.1.6. Re-establishing Norms

The norms of the scale were recalculated for two major reasons. The first of these is that by establishing similar norms on more than one occasion the resulting figures are far less due to a random occurrence than if this has only been done once. The second of these is that the first set of norms was essentially calculated on edited statistics. The scale was distributed with 65 ODS items as well as 15 items from the NDS and SDS. The 25 questions in the final scale were also in an entirely different order in that initial questionnaire. All of this can have a serious effect on the score of each person's final scale. Although the scale was still distributed amongst other tests in this research, it was the first scale the participant had to tackle and could effectively be thought of as being distributed alone for the first time. With that being said, it is also very difficult not to concede that distributing a scale with a series of other tests may also have an effect on the way the items are answer in comparison to how they might be if the scale was genuinely distributed separately.

So then a repeated calculation of the scale norms would further cement those observed figures. The figures were surprisingly similar to the first set of scale norms set down in the original research, suggesting that the items in the scale are not answered differently depending on the order of the items or the amount of other potentially distracting items in-between. The alpha level had actually increased slightly, meaning the items worked better together than they had before and the fact that only one item dragged the level of alpha down (and minimally so) means that the right set of 25 items made the final scale. It was also positive to see that even with the

great care taken to score the NDS and the TDS correctly the alpha level of the ODS was still superior.

The only major difference between the two sets of norms was the mean statistic and this was the result of the repeated problem of including the older generation in the sample; the means were nearly identical when they were removed. The higher mean was left as the official mean as both samples had been made up primarily from university students; adding the contingent of older people to the mean gave the norms a more diverse and representative range of the general population.

6.2. Theoretical Implications

6.2.1. Broken Windows Theory

Harcourt (2001) points out a crucial error of broken windows research: the results of the research were too quickly attributed to broken windows itself while the potential for other programs or simple economic factors to play a role in the observed results was far too large to ignore. The research done by Keiser et al (2008) seems to suggest that BWT works better in a more controlled environment, but that it is still a social phenomenon and could easily be effective in large scale implementations. Harcourt (2001) does not appear to argue against the idea of BWT, but simply suggests that the theory itself cannot be the exclusive or even major reason for the observed changes in the relevant environments.

It is difficult to argue against the idea that such a trait exists. Once a person is presented with a situation where there is a conflict between the general law and the social norm and they are required to act, a choice will be made. No matter which choice is made, it is possible to ascribe some form of order or some form of chaos to the nature of that choice and by setting up an iterated format of this choice it is inevitable that a person will begin to favour one type of choice over the other. This, of course, suggests the idea of dominance. A percentage of the population will choose order or chaos regardless of the surrounding environment or the extenuating circumstances, but the percentage of the population whose choice is altered based on the surrounding environment or the extenuating circumstances are what make the theory of broken windows something that is tangible and measurable. The answer to Harcourt's (2001) criticism of BWT is not that it does not work, because the effect definitely exists as long as there is a choice, but that the effect of BWT may not be as pervasive as it was initially thought to be.

The idea of individualising the effect not only helps in bypassing Harcourt's (2001) argument, but also helps in measuring the size of the effect on an individual level. Traditionally BWT's level of success has been measured by the change in the environment (i.e. decrease crime rates), but the ODS would take the perspective of the individual instead, giving BWT an idea of how the people in any given environment would react to a change in order. In other words, a highly chaos dominant population would be resistant to the idea of change in favour of order as they are unlikely to choose an order dominant option at any point in time and the reverse would be true for a highly order dominant population. The study would have the most success with a population of 'fence-sitters' (people who cannot consistently choose which way they would act)

as this group of people are most likely to change when presented with new surroundings or circumstances.

The major problem for a scale of this kind is how applicable a measurement of order dominance is to a real life scenario where those choices are presented. This problem does not manifest itself when the environment is the primary measure of success, but it becomes problematic when the individual is the source of data. The idea is to gauge a participant's level of order dominance and not the level of what the participant views as order dominance. By doing this, the scale will measure a general level of order dominance which will hopefully be applicable outside the conditions (time and place) that the test was taken in. To stop participants from creating elaborate hypotheticals with each scenario the instructions at the beginning of the scale emphasize the importance of the first answer that comes into the participants head. In this way the participant answers the question as quickly as possible without giving them enough time to imagine more and more extraneous and confounding variables for the scenario. The more specific the scenario the participant creates, the less applicable the final order dominance score is outside of the test conditions. A researcher may find that a participant who takes too long on the items may be giving a more accurate impression of the order or chaos present in their immediate environment as opposed to their own personal measure of order dominance, but this would be something for future investigation.

6.2.2. Reversal Theory

The choice that each person is presented with when faced with the scenarios of the scale adheres perfectly to the idea of reversals. The choice is bipolar in its extremes, i.e. choosing to litter or not, and the resulting action either adds to the order or to the chaos. Figure 2c is the best representative model of how a person may act when given a choice of this kind. Inside them would exist a level of order dominance (represented by the off-centre placement of the fulcrum) and if enough external force (the environment or circumstances) was exerted on the board a switch may occur between the states, otherwise the individual would continue to act in the way that their dominant state suggests. The ODS (it is hoped) is not simply measuring how a particular situation may cause an individual to act, but how the individual will behave in general. This is best represented by figure 16.

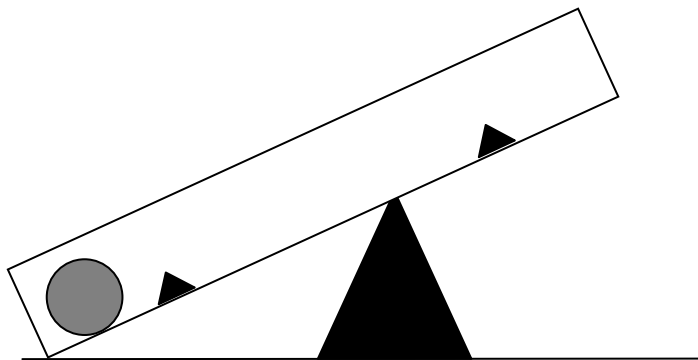


Figure 16: Visual Representation of an ODS Score

Figure 16 only differs from figure 2c by the addition of two small triangles on either side of the board. The off-centre fulcrum still represents an individual who is dominant in one way or another, but the triangles represents the stability enjoyed by one of the two states. In a once-off scenario (figure 2c), force would need to be exerted on the board for a short time only to initiate

a change in state. However, to change an individual from being order dominant to being chaos dominant would require a similar amount of force exerted for an extended period of time. The individual is only likely to change their dominant state if the environment makes a permanent change.

Crucially, however, is the position of the fulcrum. It is not known whether the fulcrum shifts in order to accommodate this new state of dominance or whether it remains the same. If the fulcrum were to shift it would suggest that the state of order dominance (or any kind of mental state dominance) is determined on a purely environmental basis. However, if the fulcrum remains in the same position it would suggest that dominance is genetic. This would mean that dominant states could be altered, but it would always be easier to slide back into one of the two dominant states. The implications for this extend well beyond the scope of the research. For example, if one were to consider sexual preference as one of two (or more) dominant states. A fulcrum that remains the same would suggest that sexual preference is genetic, while one that moves would suggest that it is acquired.

It is important to note that while each item in the scale is bipolar in nature the ODS score is not. People are not simply order dominant or chaos dominant otherwise there would be no point in giving a score. Many different levels of order dominance exist and in the case of the ODS scale the number has been simplified to 51 potential levels, but many more may exist. A scale turns the bipolar see-saw movements of dominance into far less fluid ones. Imagine a cog-like joint attaching the board to the fulcrum and the movement occurs in increments instead of one relatively fluid motion from order to chaos dominance (or vice versa). In the same respect as

before, an offset fulcrum would result in the board preferring one state to another. It is relatively easy for the board to move and settle with a bit of force to various other increments around it (i.e. shift around within the preferred state of dominance), but relatively difficult for the board to settle on the opposite side even with a large amount of external force.

The models of reversal theory are quite clear in the idea that dominance is an individual experience. The motivation to switch may be caused by external social means, but the ultimate change is in the individual, not in the surrounding environment. However, as will now be discussed, enough individual change will eventually have significant and noticeable effects on the surrounding environment.

6.2.3. Secondary Theories

6.2.3.1. Evolutionary Game Theory

Evolutionary game theory acts less as an alternative theory to the two main theories used to create the scale and more as means for explaining how the phenomenon of order dominance may have arisen. In this concept of evolutionary game theory every individual would adopt either the hawk or dove strategy (obviously far more than just the two strategies exist; only these two extremes are used to assist in explaining). In the case of order dominance this would mean that every individual would adopt a chaos or order based strategy. This strategy would give the surrounding environment a dominant strategy (either order or chaos) and it would continue to regress in this way until one could conclude on whether or not the world was in fact order or

chaos dominant. If one could accurately calculate the pay-off for 'winning' and 'losing' a resource battle (slightly more obscure in this case) for both order and chaos dominance then one could also accurately work out the expected ratio at which order and chaos dominance strategies would balance out.

Once these strategies had reached an equilibrium the ball on the board of the immediate environment or indeed the entire world would be free to oscillate to the increments immediately surrounding it, but a complete change in dominance would require a major external force to act on it and, as the environment is acting as its own individual, this change would have to be at the influence of other environments. If the pay-offs remain the same for adopting one of the two strategies it is unlikely that one strategy will dominant the other entirely, similar to the mutant hawk gene when it first appeared in the hawk-dove game. It appears then that as long as there are at least two strategies of order dominance to adopt, the trait of order dominance would be an evolutionary stable strategy.

This idea would have to be scrapped if the fulcrum argument came into play. If the order dominance trait were a genetic trait then individual shifts, let alone environmental shifts, would be significantly harder to achieve. This would not mean that they were impossible, but the genetic tendency of an individual to favour one state over the other would mean an order dominant individual would be likely to remain an order dominant individual in the same way an order dominant environment would remain the same. Again if one were aware of the ratio of order to chaos dominant individuals that reached equilibrium, one would essentially be able to

inject a chaos dominant environment full of order dominant individuals and monitor whether the individuals or the environment adapted.

6.2.3.2. Social Identity Theory

Social identity theory shows how an individual may use the environmental cues to identify the appropriate means of behaving in that environment. If one were taking the perspective of social identity then the individual would make decisions based on what they have available to them socially. In this respect, introducing an order dominant individual to a chaos dominant environment is unlikely to spark change in the environment. The individual may choose to remain order dominant, but their socially available information will be telling them otherwise. Without the support of a likeminded social group it is highly unlikely that the individual will be able to remain in a contradictory state as the pressure from the environment will make this difficult. The environment in question need not mean that the individual have any direct contact with other members of the community. The nature of the environment itself will dictate to the individual how the community behaves in respect to order and chaos dominance.

The second important thing to remember is that order dominance is not only bipolar. In the very extremes it may be, but an unknown number of different levels may exist (the scale only measures a theoretical 51). This means that an individual is almost always going to enter an environment, unless it is their own (and even then it is subject to change), with a level of order dominance that is not equivalent to the level of that environment. If the differences in level are small then it is likely that the individual will adapt to the environmental changes, but the larger

the differences become the more resistant the individual will become to that change, regardless of what the social norms may be dictating. These dynamics change further if the individual was acting as one member of a likeminded social group in a new environment. A group of individuals in a conflicting environment are highly unlikely to change as the immediate group dynamic is more important to maintain than the standards of the environment. A group of likeminded individuals in a conflicting environment is likely to be the scenario where change or switching is most resisted.

6.3. Weaknesses of the research

The sample is always one of the most crucial parts of a piece of research, as an unrepresentative sample would mean that the results do not accurately reflect the population dynamic. This was proven outright when it was discovered what an essential role age plays in the trait of order dominance. An unrepresentative sample is far more damaging to the development of a psychometric scale than to normal social research, as it means that the scale has been constructed with data that limits the cultural reaches of the scale. The disadvantages of using a sample comprised primarily of students are the advantages that the students have over the general population. The first is that they have the insight necessary to understand the underlying personality traits that the scale is attempting to measure and this gives them the ability to either answer in a very socially desirable way or completely abuse the test and come out as chaos-dominant as possible. The second advantage, or rather trait, which university students have is that they are likely to be either an order-dominant or chaos-dominant population, as they share

many similar attributes. This means that the scores obtained for the ODS in this research could be skewed in one direction or another.

The research also suffers from a lack of literature. There is plenty of literature on all the relevant theories used in the research, but the trait of order dominance was not strictly realised as a trait which could be measured prior to the research done by Jones (2009). This left the research in uncharted waters regarding what to expect during the construction, reliability analysis and the validity analysis. It was created using the theory available, but whether or not it would actually work properly and adhere to those theories was an unknown.

The major weakness of the research though was the timeframe within which it needed to be completed. Due to the short timeframe the sample was small and convenient. Ideally the research would have also made use of more than one sample, so each type of validity could be measured independently. However, with that being said, the theory is highly applicable to the research at hand, the data was analysed with great care and accuracy, and, considering the allotted time for the research, was very thorough in exploring all potential options.

6.4. Future research

Now that the ODS has proven to have a sufficient level of reliability and validity, the scope for research opportunities with the scale is vast. The research done to assess the validity of the scale has raised a few potential areas for future studies in the demographic section alone. The difference in scores of different age generations was assumed to exist, but the vastness of the gap

could never have been predicted. This suggests that there are systematic differences between people on the order dominance trait with a large enough age gap. While this is obvious from the current piece of research, it is less evident whether these systematic differences exist between items (although the standard deviation of the older contingent in comparison to the younger contingent does give a clue towards this hypothesis). The current sample is too small and has been over analysed already if one wants to consider seriously testing this assumption. If the older generation could be identified to be answering a specific set of questions in the scale consistently different to the younger generation, it would imply two things. The first of these is that the scale does not apply properly to the older generation because the conflicts presented to them in the scale are not even viewed as conflicts, due to a continuous change in standards of etiquette. The second of these is that these items would be able to identify which of these etiquette standards had changed drastically enough to not even be considered relevant to the older generation anymore. The implications for the scale may easily be outweighed by the other potential research opportunities finding these differences would open up.

Evidence supporting the fact that males are more risky investments than females (insurance, for example) is more than sufficient to suggest that differences may exist between males and females on the order dominance trait. However, the original study (Jones, 2009) suggested that the differences found between males and females could be due to the severely unequal sample sizes. The differences between males and females in this research was only discovered when the older participants were removed from the sample and even then the sample size was still skewed in favour of females. The gender differences have not been tested on an even playing field as of yet and this makes it both difficult to conclude that the differences really exist as well as what the

origins of these differences may be. A study designed to exploit these differences (whether they exist or not) may make great strides towards discovering the reasons for why they exist.

The scale could also be used in schools as an assistant to the educators. Children with higher chaos dominant tendencies would be monitored closer than those who are more likely to uphold the rules that are given to them. However, the scale is not currently adaptable for use on children as the design of the research restricted the sample to participants over the age of eighteen and as a result a proportion of the questions were related more to the behaviour of adults than children. Should research on order dominance be conducted on children then the scale would need to be adapted appropriately.

These are but a few ideas of directions for which the order dominance scale could be taken in the future. The scale would apply in all scenarios where opportunities to be order or chaos dominant exist. The question for future researchers wishing to use the scale is which of these situations or scenarios would be of most interest to the field of psychological research.

7. Conclusion

The goal of this research was different to the goals of traditional psychological research. While a problem is being solved or a gap in the literature is being filled, the outcomes of the research do not reveal new pieces of knowledge about a particular population of individuals. Although the research has unearthed some potentially interesting pieces of knowledge that could be pursued in future, the focus of this research has never been on the individuals in the study, but the psychological instrument that was being constructed.

The idea for the Order Dominance Scale was a combination of two previously existing psychological ideas, but the trait had never been hypothesized or measured previously. The outcomes of this research and the research done by Jones (2009) suggest that the scale has proven a reliable measure and has been validated to prove that this measure is indeed order dominance (with order dominance being defined as a tendency towards the general law in the presence of a conflicting social norm and chaos dominance being defined as a tendency towards a social norm in conflict with the general law). However, the scale was proven reliable and valid on small and unrepresentative samples, negatively affecting the conclusions drawn by those reliability and validity statistics. The Order Dominance Scale that this research concludes with as its psychometrically sufficient measure of the proposed personality trait can only be called a taste of what a large scale research venture would discover with more time and a bigger budget. The current scale may act as a measure of concurrent validity in such a study, but the scope past that is limited.

Scales created with the classical test theory methodologies are becoming rarer in psychological research and literature. The focus is becoming on the shortest and most effective way of diagnosing an individual, such is the case with the PHQ-2 (Kroenke, Spitzer, & Williams, 2003) which assesses an individual's depressive state in two questions. Upon discovering the individual is depressive the PHQ-9 (9 questions) is administered to assess the severity of the depressive state, although the PHQ-2 has done some of this assessment already. Psychometric tests consisting of 25-items, such as the ODS, are becoming less desirable as measurements in the research field. Applying the ideas of item response theory to the ODS in a larger scale experiment, may make the scale more user friendly to any institutions wishing to use it.

With all of that being said, the resulting scale is still a step in the right direction and a number of positives have emerged from the research. This was the first time the observations of Keizer et al (2008) were hypothesized to be conflicts occurring at the individual level and that the varied increases in crime were due to both the varied level of social conflicts and the varied levels of order dominance in each person. Although the environment plays a huge role in the order dominance trait, the idea introduced by the scale suggests that the environment is not a sufficient indicator of this phenomenon; the type of individual is contingent on the observed result too.

One of the biggest positives to take out of the research was the scale's consistent norms. The original norms were calculated on a very similar population (when the older contingent was removed), and even with the administration procedure being quite different (i.e. the number of items in the scale, the order of those items, and the number of other tests) the norms were nearly

identical. This bodes well for the scale's reliability and suggests that the methodologies undertaken to construct the scale, flawed as they may have been, did not result in a measure of order dominance that could not be used, but rather that these methods were stringent enough to create a reliable scale. The limited time and budget is always going to prove the downfall of a research venture of this size, but with that in mind, the resulting Order Dominance Scale has proved to be a reliable and valid construct according to classical test theory ideologies. The descriptives of the final scale include a range of 8-48, a mean of 31.38 with a standard deviation of 7.79, and a reliability coefficient of 0.7527.

8. References

- Apter, M. J. (1982). *The experience of motivation: The theory of psychological reversals*. London: Academic Press
- Boekaerts, M., Hendriksen, J., & Michels, C. (1988). The assessment of telic dominance in primary school pupils. In Apter, M. J., Kerr, J. H., & Cowles, M. P. (Eds.), *Advance in psychology: Vol. 51. Progress in reversal theory* (pp. 265-274, 369-372). Amsterdam: North-Holland.
- Brook, D. (2006, February 19). *The cracks in 'broken windows'*. The Boston Globe. Accessed on 12 March 2009 at http://www.boston.com/news/globe/ideas/articles/2006/02/19/the_cracks_in_broken_windows/
- Corman, H., & Mocan, N. (2005). Carrots, Sticks, and Broken Windows. *Journal of Law and Economics*. Vol. 48.
- Cressman, R., Morrison, W. G., & Wen, J. (1998). On the Evolutionary Dynamics of Crime. *The Canadian Journal of Economics*. 31, pp 1101-1117.
- Dawkins, R. (2006), *The Selfish Gene*. Oxford University Press: Oxford.
- DeVellis, R. F. (2003). *Scale development: Theory and applications*. (2nd ed.). Thousand Oaks, CA: Sage.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. (2nd ed.). New York: John Wiley.
- Gault, M., & Silver, E. (2008). Spuriousness or mediation? Broken windows according to Sampson and Raudenbush (1999). *Journal of Criminal Justice*, 36, 240-243.

- Harcourt, B. E. (2001). *Illusion of Order: The False Promise of Broken Windows Policing*. Cambridge, MA: Harvard University Press.
- Harcourt, B.E., & Ludwig, J. (2006). Broken Windows: New Evidence from New York City and a Five-City Social Experiment. *University of Chicago Law Review*, 73.
- Jones, K. (2009) *The Order Dominance Scale: Construction and Reliability Analysis*. Unpublished honours thesis. University of Kwa-Zulu Natal, Pietermaritzburg.
- Keizer, K., Lindenberg, S., & Steg, L. (2008, December). The spreading of disorder. *Science*, 322, 1681-1685.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2003). The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Med Care*. 41, pp 1284-1292.
- Lachenicht, L. (1988). A critical introduction to reversal theory. In Apter, M. J., Kerr, J. H., & Cowles, M. P. (Eds.), *Advance in psychology: Vol. 51. Progress in reversal theory* (pp. 1-43). Amsterdam: North-Holland.
- Loewenthal, K. M. (2004). *An Introduction to Psychological Tests and Scales*. (2nd ed.). New York: Psychology Press LTD.
- Maynard-Smith, J. (1964). Group selection and kin selection. *Nature*. 201, pp 1145-1147.
- McDermott, M. R. (1988). Measuring rebelliousness: The development of the Negativism Dominance Scale. In Apter, M. J., Kerr, J. H., & Cowles, M. P. (Eds.), *Advance in psychology: Vol. 51. Progress in reversal theory* (pp. 297-312,). Amsterdam: North-Holland.
- McDermott, M. R., & Apter, M.J. (1988). Negativism Dominance Scale. In Apter, M. J., Kerr, J. H., & Cowles, M. P. (Eds.), *Advance in psychology: Vol. 51. Progress in reversal theory* (pp. 373-376,). Amsterdam: North-Holland.

- Murgatroyd, S., Rushton, C., Apter, M. J., & Ray, C. (1988). Telic Dominance Scale. In Apter, M. J., Kerr, J. H., & Cowles, M. P. (Eds.), *Advance in psychology: Vol. 51. Progress in reversal theory* (pp. 373-376,). Amsterdam: North-Holland.
- Sampson, R. J., & Raudenbush, S. W. (1999). Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American Journal of Sociology*, 105, 603–651.
- Sherif, M. (1966). *In Common Predicament: Social Psychology of Intergroup conflict and Cooperation*. Boston: Houghton-Mifflin.
- Strahan, R., & Gerbasi, K. (1972). Short, homogenous version of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 28, 191-193.
- Tajfel, H. & Turner, J. C. (1979). An Integrative Theory of Intergroup Conflict. In W. G. Austin & S. Worchel (Eds.), *The Social Psychology of Intergroup Relations*. Monterey, CA: Brooks-Cole
- Weston, D., & Rosenthal. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology*, 84, 608-618.
- Wilson, J. Q., & Kelling, G. L. (1982). Broken windows. *Atlantic Monthly*, 249, 29–38.

9. Appendices

9.1. Consent Form

Consent form

The order dominance scale: Research into a usable measure for order dominance in individuals

This research project aims to develop a scale that would measure the trait of order dominance within an individual. The project requires you to fill in a small amount of information about yourself and to complete a series of four independent scales

In order to ensure complete honesty the test will remain completely anonymous and all information will be kept confidential. Please be honest, it is vital to the test.

Should you wish to withdraw from the research project at any point you may do so by using the contact details provided below and giving the researcher your reference number. There are no penalties for withdrawing from the research project and doing so would ensure that your data is excluded from the remainder of the project.

If you intend to participate in this project, please sign the declaration below and hand it in with your completed tests. Should you have any queries about anything regarding the research project, do not hesitate to make contact through the details below.

Declaration

I hereby confirm that I understand the contents of this document and the nature of the research project, and I consent to participating in the research project.

I understand that I am at liberty to withdraw from the project at any time, should I so desire.

SIGNATURE OF PARTICIPANT:..... **DATE:**.....

Kyle Jones
Cell: 0836621145
E-mail: 206508190@ukzn.ac.za

Prof Lance Lachenicht
E-mail: lachenicht@ukzn.ac.za

9.2. The Order Dominance Scale and Scoring sheet

Instructions

This is a measure designed to ascertain how you would react in certain social situations. There are 25 questions to answer and each has three possible responses. These responses by no means cover all the possible options, but please accept this and make your choices from the options given to you. The 'not sure' option in every question can double in a few cases as a 'sometimes' option, however, please try to use this option as little as possible. Please do not think about your answers for too long. It is your first reaction that is most important. Should you complete the test in this way, it should take you between 5-10 minutes

This is not a test of intelligence or ability and there are no right or wrong answers and complete honesty would be greatly appreciated.

Thank you for your help.

1. Do you believe smoking marijuana should be legalised?
 - a. Yes
 - b. No
 - c. Not sure
2. Do you obey to signs like "DO NOT TOUCH" or "KEEP OFF THE GRASS" if it's clear that no one else around you is obeying to them?
 - a. Yes
 - b. No
 - c. Not sure
3. Do you believe that breaking the law sometimes does more good than bad?
 - a. No
 - b. Yes
 - c. Not sure
4. You see someone littering in front of you in a public area. Do you
 - a. Do something about it
 - b. Ignore it
 - c. Not sure
5. You and your colleagues are required to prepare for a meeting. If no one looks like they are going to prepare. Do you
 - a. Not prepare
 - b. Still prepare
 - c. Not sure
6. It is alright to litter in an area that is already full of litter
 - a. No
 - b. Yes
 - c. Not sure
7. You are in a night club with a bunch of friends when someone breaks a glass near your group. Do you
 - a. Leave it and move away
 - b. Attempt to clean up a little, so it's safer for your friends
 - c. Not sure
8. You see writing on a desk or bathroom wall that offends you. Do you
 - a. Write a response
 - b. Leave it
 - c. Not sure
9. Your friends tell you to bunk out work on a Friday for a long weekend. Do you
 - a. Agree, it's only once
 - b. Disagree, work is your responsibility
 - c. Not sure
10. Your friend forgets they lent you R50 to go out. Do you
 - a. Forget about it too
 - b. Pay them back
 - c. Not sure
11. Have you ever driven when you have had a little too much alcohol?
 - a. Yes

- b. No
 - c. Not sure (if you were over the limit)
12. You see R10 fall out of someone's pocket and someone in front of you steals it. Do you
- a. Approach them and tell them to return it
 - b. Leave it, it is not your problem
 - c. Not sure
13. There is a traffic jam on the highway. A few people start using the emergency lane. Do you
- a. Follow their lead
 - b. Stay in your lane
 - c. Not sure
14. You leave quite a bad scratch in a new car, by mistake. Do you
- a. Leave your name and number, to pay for the damages
 - b. Ignore it, mistakes happen
 - c. Not sure
15. There is a fire in a building you are in. Do you
- a. Follow the exit procedure
 - b. Get out as fast as you can
 - c. Not sure
16. Would you rather travel through an unsafe neighbourhood if the route was shorter than the safe route?
- a. Yes
 - b. No
 - c. Not sure
17. Most people leave their rubbish in a cinema, probably because there are no bins. If rubbish bins were placed on the way out the theatre, would you
- a. Throw your rubbish away
 - b. Still leave it at your seat
 - c. Not sure
18. You happen to witness a number of people where you work taking office supplies. Do you
- a. Also feel you can take a few things
 - b. Uphold the rules
 - c. Not sure
19. Do you copy music instead of buying it?
- a. No
 - b. Yes
 - c. Not sure
20. You and your friend notice R20 lying on the floor on the street. Do you
- a. Take it
 - b. Try find out who lost it
 - c. Not sure
21. You see someone throwing glass into a paper recycling bin. Do you
- a. Put the glass into the correct bin
 - b. Leave it
 - c. Not sure
22. When you are reprimanded by a person of authority. Do you
- a. Get irritated and try find a new way to do what you were doing
 - b. Feel bad and apologize
 - c. Not sure
23. A group of your friends are talking in a library. Do you
- a. Ask them to keep quiet
 - b. Go talk with them
 - c. Not sure
24. Is your cellphone always on silent in movies?
- a. Yes
 - b. No
 - c. Not sure
25. You work for a small business and you give a customer too little change, but only notice once they have left. Do you
- a. Try to correct it
 - b. Leave it
 - c. Not sure

The scoring for the ODS is as follows:

Item	A	B	C
1	0	2	1
2	2	0	1
3	2	0	1
4	2	0	1
5	0	2	1
6	2	0	1
7	0	2	1
8	0	2	1
9	0	2	1
10	0	2	1
11	0	2	1
12	2	0	1

13	0	2	1
14	2	0	1
15	2	0	1
16	0	2	1
17	2	0	1
18	0	2	1
19	2	0	1
20	0	2	1
21	2	0	1
22	0	2	1
23	2	0	1
24	2	0	1
25	2	0	1

9.3. The Negativism Dominance Scale and Scoring sheet

Instructions

This is a measure of the way in which you react to certain social situations. For each of the following eighteen items three responses are given. For each item choose the response which is most true of you. These responses do not represent all the possible ones but please accept this and make your choice from those which happen to be given. Try to use the 'not sure' response as little as possible. Do not think too long about your answer. It is the first reaction that is important. Put a 'X' in the circle next to the response which is most true of you.

This is not a test of intelligence or ability and there are no right or wrong answers.

Thank you for your help.

1. When you are told that you are breaking a rule (for example, 'No Smoking'), is your first reaction to
 - a. Stop breaking the rule any further
 - b. Go ahead and still break the rule
 - c. Not sure
2. You have been treated badly by someone. Do you
 - a. Try to get back at the person
 - b. Hope that things will improve
 - c. Not sure
3. In trying to complete an exercise routine, you go through some pain. Do you
 - a. Continue
 - b. Give up
 - c. Not sure
4. "I enjoy the thrill I get from being difficult and awkward." Do you
 - a. Agree
 - b. Disagree
 - c. Not sure
5. If people are unkind to you, do you feel you should be
 - a. Unkind back
 - b. Understanding
 - c. Not sure
6. Do you find it exciting to do something 'shocking'?
 - a. Yes, often
 - b. No, hardly ever
 - c. Not sure
7. If you are asked particularly NOT to do something, do you feel an urge to do it?
 - a. No, hardly ever
 - b. Yes, often
 - c. Not sure
8. You are in a group of people who are drinking, but you do not like alcohol and are offered a drink. Would you
 - a. Refuse the drink
 - b. Accept the drink
 - c. Not sure
9. Do you tease people unnecessarily just so as to have some fun at their expense?
 - a. Yes, often
 - b. No, hardly ever
 - c. Not sure
10. A parking attendant tells you that you cannot park where you have just put the car. Would you
 - a. Apologise and move it
 - b. Argue with the attendant
 - c. Not sure
11. How often do you do something you shouldn't just to get some excitement?
 - a. Not often at all
 - b. Often
 - c. Not sure

12. You are asked to take part in an activity which secretly you dislike. Would you
- Say you have something else planned
 - Try hard to avoid an argument
 - Not sure
13. If you get yelled at by someone in authority, would you
- Get angry and argue back
 - Try hard to avoid an argument
 - Not sure
14. If a person your age was mean to you, would you
- Try to forget it
 - Try to get revenge
 - Not sure
15. Can you think of anything you oppose strongly
- No
 - Yes
 - Not sure
16. A charity will not accept you as a volunteer. Is your first reaction to
- Thank them for considering you
 - Tell them to 'go to hell'
 - Not sure
17. How often do others say that you are a difficult person?
- Rarely
 - Often
 - Not sure
18. If you ask a person at a party to dance with you who says 'no'

without offering any explanation, would you

- Get annoyed
- Accept it
- Not sure

The scoring for the NDS is as follows:

Item	A	B	C
1	0	2	1
2	2	0	1
3	N/A	N/A	N/A
4	2	0	1
5	2	0	1
6	2	0	1
7	0	2	1
8	N/A	N/A	N/A
9	2	0	1
10	0	2	1
11	0	2	1
12	N/A	N/A	N/A
13	2	0	1
14	0	2	1
15	N/A	N/A	N/A
16	0	2	1
17	0	2	1
18	2	0	1

9.4. The Telic Dominance Scale and Scoring sheet

Instructions

Here are some alternative choices. If you have an open choice, which of the following alternative would you usually prefer. Please complete all the items by putting a cross in the circle corresponding to your choice, making one choice for each numbered item. Only if you are not able to make a choice should you put a cross in the circle corresponding to 'Not sure'. Try to answer all of the items by putting a cross in one of the circles for each item, using the 'Not sure' choice as little as you can. Work quickly and do not spend too much time on any one item: it is your first reaction we want.

This is not a test of intelligence or ability and there are no right or wrong answers.

1.
 - a. Compile a short dictionary for financial reward
 - b. Write a short story for fun
 - c. Not sure
2.
 - a. Going to evening class to improve your qualifications
 - b. Going to evening class for fun
 - c. Not sure
3.
 - a. Leisure activities which are just exciting
 - b. Leisure activities which have a purpose
 - c. Not sure
4.
 - a. Improving a sporting skill by playing a game
 - b. Improving it through systematic practice
 - c. Not sure
5.
 - a. Spending one's life in many different places
 - b. Spending most of one's life in one place
 - c. Not sure
6.
 - a. Work that earns promotion
 - b. Work that you enjoy doing
 - c. Not sure
7.
 - a. Planning your leisure
 - b. Doing things on the spur of the moment
 - c. Not sure
8.
 - a. Going to formal evening meetings
 - b. Watching television for entertainment
 - c. Not sure
9.
 - a. Having your tasks set for you
 - b. Choosing your own activities
 - c. Not sure
10.
 - a. Investing money in a long term insurance/pension scheme
 - b. Buying an expensive car
 - c. Not sure
11.
 - a. Staying in one job
 - b. Having many changes of job
 - c. Not sure
12.
 - a. Seldom doing things 'for kicks'
 - b. Often doing things 'for kicks'
 - c. Not sure
13.
 - a. Going to a party
 - b. Going to a meeting

14. c. Not sure
a. Leisure activities
b. Work activities
c. Not sure
15. a. Taking holidays in many different places
b. Taking holidays always in the same place
c. Not sure
16. a. Going away on holiday for two weeks
b. Given two weeks of free time finishing a needed improvement at home
c. Not sure
17. a. Taking life seriously
b. Treating life light-heartedly
c. Not sure
18. a. Frequently trying strange foods
b. Always eating similar foods
c. Not sure
19. a. Recounting an incident accurately
b. Exaggerating for effect
c. Not sure
20. a. Spending R500 having an enjoyable weekend
b. Spending R500 on repaying a loan
c. Not sure
21. a. Having continuity in the place where you live
b. Having frequent moves of house
c. Not sure
22. a. Going to an art gallery to enjoy the exhibits
b. To learn about the exhibits
c. Not sure
23. a. Watching a game
b. Refereeing a game
c. Not sure
24. a. Eating special things because you enjoy them
b. Eating special things because they are good for your health
c. Not sure
25. a. Fixing long-term life ambitions
b. Living life as it comes
c. Not sure
26. a. Always trying to finish your work before you enjoy yourself
b. Frequently going out for enjoyment before all of your work is finished
c. Not sure
27. a. Not needing to explain your behaviour
b. Having purposes for your behaviour
c. Not sure
28. a. Climbing a mountain to try to save someone
b. Climbing a mountain for pleasure
c. Not sure
29. a. Happy to waste time
b. Always having to be busy
c. Not sure
30. a. Taking risks
b. Going through life safely

31. c. Not sure
 a. Watching a crucial match between two ordinary sides
 b. Watching an exhibition game with star performers
 c. Not sure
32. a. Playing a game
 b. Organizing a game
 c. Not sure
33. a. Glancing at pictures in a book
 b. Reading a biography
 c. Not sure
34. a. Winning a game easily
 b. Playing a game with the scores very close
 c. Not sure
35. a. Steady routine in life
 b. Continual unexpectedness or surprise
 c. Not sure
36. a. Working in the garden
 b. Picking wild fruit
 c. Not sure
37. a. Reading for information
 b. Reading for fun
 c. Not sure
38. a. Arguing for fun
 b. Arguing with others seriously to change their opinion
 c. Not sure
39. a. Winning a game
 b. Playing a game for fun
 c. Not sure
40. a. Travelling a great deal in one's job
- b. Working in one office or workshop
 c. Not sure
41. a. Planning ahead
 b. Taking each day as it comes
 c. Not sure
42. a. Planning a holiday
 b. Being on holiday
 c. Not sure

The scoring for the TDS is as follows:

Item	A	B	C
1	1	0	½
2	1	0	½
3	0	1	½
4	0	1	½
5	0	1	½
6	1	0	½
7	1	0	½
8	1	0	½
9	1	0	½
10	1	0	½
11	1	0	½
12	1	0	½
13	0	1	½
14	0	1	½
15	0	1	½
16	0	1	½
17	1	0	½
18	1	0	½
19	1	0	½
20	0	1	½
21	1	0	½
22	0	1	½
23	0	1	½
24	0	1	½
25	1	0	½
26	1	0	½
27	0	1	½
28	1	0	½
29	0	1	½
30	0	1	½
31	1	0	½
32	0	1	½
33	0	1	½
34	1	0	½
35	1	0	½
36	1	0	½
37	1	0	½

38	0	1	½
39	1	0	½
40	0	1	½
41	1	0	½
42	1	0	½

9.5. The Vignettes

9.5.1. Pair 1

SCENARIO 1

Imagine you and a friend are going out to the movies and you notice R50 falling out of someone's pocket in front of you. They don't notice the money falling out because they appear to be in a hurry. You and your friend also appear to be the only ones who have noticed this. If you are quick, you can still catch the person who dropped the money. What do you do next in this situation?

SCENARIO 2

Imagine you are attending an exhibition of sorts (something you would really enjoy) and you bring your camera with you. When you arrive at the exhibition, you find that no photography is allowed anywhere inside. The people hosting the exhibition have hired security to make sure no one misbehaves or, in your case, takes pictures. However, you find a part of the exhibition where there is no security and a few people who are attending have turned off their flash and begun taking photographs of a particular exhibit without being caught. You know no one will catch them because there is no security anywhere nearby. What do you do next in this situation?

9.5.2. Pair 2

SCENARIO 1

Imagine you and your friends are playing a card game at your house. You, along with one of your friends are doing much better than everyone else. You happen to look at that friend, during someone else's turn, and you notice them sliding one of the cards under their seat. No one else seems to have noticed this. What would you do next in this situation?

SCENARIO 2

Imagine you need to get something from a shop nearby your house and you decide to walk there. As you near the shop you can soon see that you will need to cross a very busy road to get to it. You can see that there are a set of robots with a pedestrian crossing about 50 meters past the shop where you can cross and walk back up to the shop. However, you also notice a few people jaywalking (crossing the street illegally) in front of you to get to the shop quicker. What would you do next in this situation?

9.5.3. Pair 3

SCENARIO 1

Imagine you go to the movies to watch a film you have been waiting a month to come out. You and your friends sit down to watch the movie, but halfway through the film a person in front of you gets a phone call and begins talking to his friend. You clearly remember the advert telling people to turn all their cellphones on silent, yet after two minutes this guy is still talking on his cellphone. You are beginning to miss parts of the movie. What do you do next in this situation?

SCENARIO 2

Imagine you are at a restaurant for dinner. The restaurant is absolutely packed, not one seat is available. After being forced to wait 20 minutes for a table, you have ordered your drinks and had to wait nearly 15 minutes for them to arrive. Now you are waiting for the food you have ordered to arrive. You have been waiting nearly 45 minutes already and you are waiting for a second round of drinks because of this. The waitron has told you that your food will arrive in around twenty minutes. What would you do next in this situation?

9.5.4. Pair 4

SCENARIO 1

Imagine that your boss at work says that he is going to be taking an extended lunch break to go out with his wife. When he leaves the office your colleagues become chattier and ease off of the work that they were doing a little. You have quite a lot of work to do before the end of the day. What would you do next in this situation?

SCENARIO 2

Imagine you made a promise to a group of friends that you would go to a house party. On the day of the party you decide that you aren't all that keen on attending the party anymore for some reason or another. What would you do next in this situation?

9.5.5. Pair 5

SCENARIO 1

Imagine that you normally drive quite fast wherever you go. You and your work colleagues have recently arranged for a lift club to work as you discovered you all live in a similar area. You have taken lifts with each of them to work and noticed that all of them drive quite slowly and often moan about speeding motorists on the way to work. It is now your turn to pick them up for work. What would you do next in this situation and why?

SCENARIO 2

Imagine you have been invited to a party and you and a group of your friends all attend. When you arrive at the party you all have a few drinks and settle down a little. A bit later, you decide that you want to dance and your friends agree with you. However, no one at the party is dancing at the moment and none of your friends will get up and dance until you do first. What do you do next in this situation?
