

5-2010

## Relative deprivation, relative gratification, status, and health

Samuel A. Holland

University of Nevada Las Vegas, sam\_a\_holland@hotmail.com

Follow this and additional works at: <https://digitalscholarship.unlv.edu/thesesdissertations>



Part of the [Health Psychology Commons](#), and the [Public Health Commons](#)

---

### Repository Citation

Holland, Samuel A., "Relative deprivation, relative gratification, status, and health" (2010). *UNLV Theses, Dissertations, Professional Papers, and Capstones*. 314.

<https://digitalscholarship.unlv.edu/thesesdissertations/314>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by Digital Scholarship@UNLV with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Dissertation has been accepted for inclusion in UNLV Theses, Dissertations, Professional Papers, and Capstones by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact [digitalscholarship@unlv.edu](mailto:digitalscholarship@unlv.edu).

RELATIVE DEPRIVATION, RELATIVE GRATIFICATION, STATUS, AND  
HEALTH

by

Samuel Holland

Bachelor of Science  
James Madison University  
2001

Master of Science  
James Madison University  
2003

A doctoral document submitted in partial fulfillment  
of the requirements for the

**Doctor of Philosophy in Psychology**  
**Department of Psychology**  
**College of Liberal Arts**

**Graduate College**  
**University of Nevada, Las Vegas**  
**May 2010**

Copyright by Samuel Holland, 2010  
All Rights Reserved



**THE GRADUATE COLLEGE**

We recommend that the dissertation prepared under our supervision by

**Samuel A. Holland**

entitled

**Relative Deprivation, Relative Gratification, Status, and Health**

be accepted in partial fulfillment of the requirements for the degree of

**Doctor of Philosophy**

Psychology

Jeffrey M. Kern, Committee Chair

David Copeland, Committee Member

Cortney Warren, Committee Member

Timothy Bungum, Graduate Faculty Representative

Ronald Smith, Ph. D., Vice President for Research and Graduate Studies  
and Dean of the Graduate College

**May 2010**

## ABSTRACT

### **Relative Deprivation, Relative Gratification, Status, and Health**

by

Samuel Holland

Dr. Jeffrey M. Kern, Examination Committee Chair  
Professor of Psychology  
University of Nevada, Las Vegas

The role of perceived inequity in health was assessed and compared to other social constructs predicted to be relevant in the relationship between inequity and health. Four studies were conducted that demonstrated that a social comparison-based measure of relative deprivation (RD) and relative gratification (RG) can predict health and continues to do so even after accounting for perceived stress, personal income, perceived control, and social participation. A measure of RD/RG was developed that was based on prominent social comparisons and objects of comparison. This measure was reliable and predictive of health in two samplings of different populations. Comparisons of the RD/RG measure to other measures are explored and future directions in health inequity research are discussed.

## TABLE OF CONTENTS

ABSTRACT .....	iii
LIST OF TABLES .....	vii
CHAPTER 1 INTRODUCTION .....	1
Purpose of Study .....	6
CHAPTER 2 REVIEW OF RELATED LITERATURE .....	7
Social Comparison Theory .....	7
The Reactions to Social Comparisons .....	36
Stress and Health .....	71
Measurement of Stress .....	100
Measurement of Health .....	103
Relative Deprivation/Gratification .....	110
Hypothesis .....	122
CHAPTER 3 STUDY 1: IDENTIFICATION OF PROMINENT COMPARISON GROUPS AND OBJECTS WITH A COLLEGE SAMPLE .....	123
Introduction .....	123
Methods .....	124
Results .....	126
Conclusion .....	130
CHAPTER 4 STUDY 2: PERCEIVED INEQUITY SCALE DEVELOPMENT AND VALIDATION WITH A COLLEGE SAMPLE .....	137
Introduction .....	137
Methods .....	138
Results .....	144
Conclusion .....	150

CHAPTER 5	STUDY 3: IDENTIFICATION OF PROMINENT COMPARISON GROUPS AND OBJECTS FROM AN EXPANDED POPULATION .....	155
	Introduction .....	155
	Methods .....	157
	Results .....	159
	Conclusion.....	163
CHAPTER 6	STUDY 4: MEASUREMENT OF RD/RG IN THE GENERAL POPULATION AND THE ROLE OF SELF-REPORTED INEQUITY IN HEALTH PREDICTION.....	168
	Introduction .....	168
	Methods.....	169
	Results .....	176
	Conclusion.....	182
CHAPTER 7	DISCUSSION.....	188
APPENDIX 1	IRB APPROVALS.....	196
REFERENCES	.....	216
VITA.....		232

## LIST OF TABLES

TABLE 1	Descriptives and t-Test Values for Study 1 .....	197
TABLE 2	ANOVA Values for Study 1 .....	198
TABLE 3	Means and Descriptives for Study 2 .....	201
TABLE 4	Correlations of Indices for Study 2 .....	202
TABLE 5	Descriptives and t-Test Values for Study 3 .....	203
TABLE 6	ANOVA Values by Gender for Study 3 .....	204
TABLE 7	ANOVA Values by Age for Study 3 .....	207
TABLE 8	Age Differences for “Members of the Same Religion” for Study 3 .....	210
TABLE 9	Means and Descriptives for Study 4 .....	211
TABLE 10	Correlations of Indices for Study 4 .....	212
TABLE 11	Correlations of Indices Split by Sex for Study 4 .....	213
TABLE 12	Summary of Regression Analysis for Study 4 .....	214
TABLE 13	Correlations of Exercise Indices for Study 4 .....	215

## CHAPTER 1

### INTRODUCTION

Social comparisons may have a significant impact on health. Social research is rich with examples of inaccuracies in self-perceptions. Davis (1966) found that when comparing students' career aspirations across colleges, career aspirations were lower at the prestigious schools. This is despite the reality that career prospects are often better for students attending more prestigious colleges. Academy award winners might live, on average, four years longer than the runners up (Redelmeir & Singh, 2001) despite the fact that academy award winners and the runners up are all successful in their field and are all likely to be quite wealthy. Marmot's (2005) review of his Whitehall and Whitehall II studies of British civil servants found that as job position increased, so did life span. One could do a lot worse than having a well-paid and benefit-rich position with the British civil service. The relationship between perceived social standing and health may be through observations of others.

People are surrounded by *social information*- the available communications about others. People are constantly exposed to opportunities to receive information about others- directly and indirectly, deliberate and not. For example, when one sees a man driving an expensive car this may be indirectly stating that he has a lot of money or maybe that he is very shallow (perhaps one derives both conclusions from seeing this). The interpretations of social information vary. One person may derive a very different conclusion than another person from the same social information.

In studies comparing industrialized countries the amount of equity within the country increases comparably to mortality rates (Wilkinson, 2005). In all of the

industrialized countries used in the study, a threshold of wealth had been achieved (i.e., the standard of living was sufficient to attain basic services such as food, clean water, and medical care). On a smaller scale, neighborhoods in the United States show a similar effect (Diez-Roux, Nieto, Muntaner, Tyroler, Comstock, Shahar, Cooper, Watson, & Szklo, 1997). As the disadvantage of the neighborhood increased, so too did the incidence of coronary heart disease (CHD). This effect remained after controlling for personal income, education, and occupation. Macintyre, Ellaway and Cummins (2002) contended that the perception of a neighborhood was one of the major factors contributing towards the residents' health.

This status health disparity may not be uniquely human. The relationship between status and health is found even among primates (Abbott, Keverne, Bercovitch, Shively, Mendoza, Saltzman, Snowdon, Ziegler, Banjevic, Garland, & Sapolsky, 2003). Higher status primates live longer than primates of lower status. Status seems to influence health in non-human primates primarily through access to social support, which increases safety and resources. With status comes additional access to material and emotional resources. In humans, this status effect may be compounded by the stress of not meeting one's own expectations, since people generally do not dream of being low on the totem pole. This stress may lower the immune system's efficiency, increasing illness and death (Cassell, 1976).

Marmot (2005) contended that there exists a *status syndrome*- one's perceived control (*autonomy*) and the ability to participate in socially important activities (*social participation*) has profound effects on health (both increase as one's social standing improves). As an example of social participation, members of the American middle class

might consider vacations and new clothes as entitlements. When one is unable (or rather, when one *believes* one is unable) to engage in these activities, health may be compromised from reduced social networks and attenuated immune functioning. This is not a poor vs. rich dichotomy but a health gradient. Incremental changes in status have corresponding health effects. While other factors contribute to health, Marmot argued that these factors are related to status. For example, education and parent's health all contribute to one's status (as well the inverse; status contributes to one's ability to afford education and the health of family). Money is another important predictor of health, but this relationship tapers off in industrialized societies once a general affordability of basic needs and services is achieved (Wilkinson, 2005).

While a relationship between social standing and health has been identified, we do not yet understand how two affluent people in a wealthy country could have different health just because one has a little *more* money. The next step in fully understanding the relationship between health and status is identifying the causal mechanism by which a psychological variable (perceived status) affects a biological variable (health). While there has been considerable progress in the last ten years in identifying the biological responses to psychological phenomena (Maier & Watkins, 1998) little progress has been made in understanding how and why these responses occur. Understanding how status affects health, however, requires understanding how a sense of status is derived. The act of determining one's status- how and where one fits within a given group- is based on interpretations of social information.

Social theories have emerged to explain the various uses of social information and provide insight regarding how social information may influence health. Heider

(1944) stated that there is a need for people to be able to predict their environment to maximize outcomes. Social information may be a reflexive act that can affect health, both positively and negatively. This process can and likely does occur without a conscious awareness as people typically don't look at someone in an expensive car and think: "He has a nice car and therefore is of a higher class than I am. I will now feel worse about myself and start smoking two packs a day."

Several social theories were developed to account for how and why people use social information (Suls & Wheeler, 2000). Over time, many of these theories have expanded to the point of overlapping into interconnected processes. In this review of the varying theories, the primary focus is on the development of one of the more popular social theories- social comparison theory (SCT). Under SCT, people compare themselves to others for the purposes of determining if their opinions about the world are accurate and to get a sense of their abilities (Festinger, 1954a). SCT predicts that people all share a desire to improve their own abilities and look to others to get a sense of what abilities can and should be improved.

SCT predicts that social comparisons can result in changes in our emotional state. Furthermore, modern SCT holds that the act of making social comparisons may be unavoidable. If people are presented with social information they do not like, then they may change their perceptions to accommodate self-esteem. For example, if one makes a comparison and does not like the conclusion, that person may feel stressed that their abilities are not up to par. From social comparisons several consequences may occur because expectations are heavily influenced by what is observed in others. Besides changes in self-esteem, social comparisons also affect one's perception of justice in the

world. Like the theories of social information, the consequences of social comparison are interconnected. Many studies support the notion that status is derived from social comparisons, and distress derived from such comparisons is greater in lower status individuals, thus compromising health.

Health inequity researchers have encountered considerable difficulties in identifying a stable and consistent relationship between status and health (Marmot, 2005). A reason for this has been a difficulty in determining how status is derived. In modern society status is not determined by birth or communicated by title. Furthermore, an individual's status today varies by context. One may have a high-status at work but a low-status in the family. In modern societies without formal class distinctions, status is understated but still present. Modern status is social in nature; one's ranking within a social hierarchy. Status is defined in part by occupation, as one's profession contributes to autonomy and social participation. As Marmot stated "You may be the world's worst parent or spouse, son or brother, but if you are the local doctor, you are somebody in the community. You have something going for you that the street sweeper does not" (p. 121).

Marmot (2005) considered the link between social status and health to be a consequence of higher social standing leading to greater autonomy and social participation. These two perceptions then act on health by creating psychological states that either hinder or enhance the immune system. I argue that status is derived from social comparisons because these perceptions (i.e., social participation and perceived control) are determined by seeing what one has and does.

This paper outlines social comparison theory as the process by which status is evaluated. The consequences of these social comparisons are proposed to be the primary factor responsible for a social gradient in health. This argument is structured into three sections. Section one is a review of SCT, section two outlines the consequences (reactions) to social comparison, and section three is a review of the mechanism of stress and health. This study is designed to measure the relationship between perceived inequity, status, and health.

## CHAPTER 2

### REVIEW OF RELATED LITERATURE

#### Social Comparison Theory

If humans in modern society evaluate status from social comparisons, then understanding how and why social comparisons occur is imperative for predicting a health gradient in status. The formal theoretical conception of social comparison is attributed to Festinger (1954a). He postulated that social comparison occurred naturally and for the purpose of self-enhancement. People prefer non-social comparisons, comparisons made on objective standards such as scaled scores. Social comparisons, however, are likely to be used when objective standards of comparison are unavailable and when a model that is similar on some specific characteristic is available. For example, a person determine if they passed a test simply by having a score above a specific cutoff score the instructor deems as “passing” (objective comparison) but will often ask other people how they did on the test (social comparison) to get a sense of how they are did, and this will likely influence how they feel about their performance. If a person barely passed a test they might feel good if no one else passed, but not if everyone else scored very well.

SCT is intertwined with Festinger’s previous theory, *informal social communication* (1950). Informal social communication theory attempts to explain the pressures towards conformity of opinion among a group. Festinger contended that this desire for conformity exists because people have a need for social reality and group locomotion. Stated another way, people try to agree on what is real and have a sense of moving forward in thought with a group. The group sends “communications” to

members, and these communications are especially salient to an individual that is out of step. For example, imagine that someone is dressing outside the social reality (or social norm, as we are more apt to call it today) and, consequently, given subtle cues that this behavior is not appropriate for the group (e.g., sneers, condescending talk, and a general lack of respect). They are behaviors, to be sure, but motivated by opinion and serve as a pressure towards uniformity to dress in more socially-acceptable ways.

SCT does not conflict with informal social communication theory (Wheeler, 2000). Rather, SCT is an expansion. While informal social communication theory concerned attitudes, SCT includes abilities and behaviors as well as attitudes (Festinger, 1954a). This is the reason for changing the term *communications* to *comparisons*. Social comparisons occur from an internal motivation to know that one's opinions are correct and to understand one's capabilities. This motivation is not social in nature, but is satisfied with social comparisons in the absence of personal experience.

Festinger's original publication on SCT (1954a) is considered difficult to interpret. Suls and Wheeler (2000) described the effort as "brutal reading" (p. 5). The more popular treatment of SCT is a symposium paper published in the same year by Festinger (1954b). In it, social comparison theory was outlined under five tenets:

1. This social process arises when the evaluation of opinions or abilities is not feasible by testing directly in the environment.
2. Under such circumstances persons evaluate their opinions and abilities by comparison with others.
3. This comparison leads to pressures toward uniformity.

4. There is a tendency to stop comparing oneself with others who are very divergent. This tendency increases if others are perceived as different from oneself in relevant dimensions.
5. Factors such as importance, relevance and attraction to a group which affect the strength of the original motivation will affect the strength of the pressure towards uniformity. (p. 217)

The symposium paper left out the notion of the unidirectional drive upward for abilities, the consistent desire for improvement. This notion that there is a drive upwards for abilities was discussed Festinger's formal publication (1954a). This drive upwards was only for abilities. That there is not a drive upwards for opinions is logical given the assumption that social comparison is used differently for abilities (gauge of potential) than opinions (gauge of accuracy).

Shortly after introducing SCT, Festinger published a paper describing *cognitive dissonance theory* (Festinger, 1957). Cognitive dissonance is a state of discomfort experienced when one realizes an opinion they held may not be true or engages in a behavior that conflicts with their beliefs. The theory attempts to predict the efforts one will engage in to reduce this discomfort. Cognitive dissonance theory is similar to SCT in the evaluation of opinions. While SCT supplanted informal social communication theory as an explanatory model, dissonance theory does not supplant SCT because it does not consider the role of other people (Wheeler, 1991). Cognitive dissonance is a complimenting theory to SCT, not a successor as SCT was to informal social communication theory.

Though Festinger (1954a) introduced SCT, it was the efforts of many other researchers that tested the theory. The original conception of SCT provided a framework for understanding the need and the nature of what may be an unavoidable behavior. The theory, however, was vague about the dynamics of these comparisons. Indeed, following Festinger's publication, (1954a), the theory went uninvestigated. Festinger himself pursued the development of the study of cognitive dissonance. As Wheeler (1991) noted, "social comparison theory went nowhere" (p. 6). Wheeler (2000) credits the development of SCT to the efforts of Schachter, who taught a yearlong graduate seminar at the University of Minnesota in 1960. These students began testing the tenants of SCT; further developing and refining the theory.

Festinger (1954a) posited that one makes comparisons to others to understand how one's abilities and opinions align with the world and provide a reference for improvement. Consequently, SCT stops when a group is very divergent from the self. Social researchers challenged these predictions. Specifically, early research explored whether SCT is performed upwards or downwards (Wheeler, 1991). Another issue was similarity or the degree to which the comparer identifies with the compared. Subsequent research explored when and why one might make a comparison with someone deemed dissimilar.

Major advances have been made in the study of SCT since Festinger's (1954a) initial offering, such to the point that many tenets of SCT now incorporate the tenets of other social process theories. Studies exploring the direction of comparisons led to SCT expanding to include emotional consequences from comparisons (Wheeler, 1991). This occurred because the reason one makes comparisons varies, in part, by affective state

and this information was necessary for predicting the direction of comparisons. The similarity assumption between comparer and compared underwent revision as well, mostly in reconsidering who is “similar”. SCT initially described a process of earnest self-assessment to similar others and was revised to describe an unavoidable process manipulated to maintain and enhance self-esteem as well evaluate opinions and abilities. These changes are so vast that some argue that the current interpretation of SCT is a different theory than the original conception put forward by Festinger. Wheeler (1991) referred to Festinger’s (1954a) social comparison for opinion validation and ability estimation as *classical social comparison theory* and the later researcher’s emphasis on positive affect and ego enhancement as *neo-social comparison theory*. These changes are reviewed here to establish a basic understanding of modern SCT and to conceptualize how SCT can influence health.

The major advances in SCT since Festinger’s (1954a) paper will be reviewed here in the following order:

- Direction of comparisons
- Similarity
- Influence of Emotion
- Individual differences
- Measurement Issues

#### *Direction of Comparisons*

One primary issue in social comparison literature is the direction of comparison. Festinger’s (1954a) SCT was initially interpreted as meaning that comparisons are upwards in nature (Wheeler, 1991). *Upward comparisons* suggests that one is

comparing themselves on a comparison dimension to people or thing that they deem to be higher in rank or status. This was one of the first predictions tested by the Minnesota students, who developed what is now known as the rank-order paradigm. This is a test of comparison direction by which participants believe they are being measured on some trait and given a numerical score of performance and told where they ranked compared to other participants (all participants were told the same score and rank). They are then offered the opportunity to see the score of another by selecting a different rank. Wheeler (1991) described his own development of the paradigm as one of the first published (Wheeler, 1966). Wheeler tested participants in groups of seven on a measure of ability to benefit from a psychology seminar. He manipulated student motivation to take the seminar by the initial description of the class (either a very worthwhile class or not). All subjects were told they ranked in the middle (ranked fourth) and given the rough scores of the top and bottom. Participants were then asked which other participant's score they would like to see. Participants overwhelmingly chose those of higher ranks, and this effect was stronger for the highly motivated participants. This supported the notion of upward comparisons.

*Downward comparisons*, those comparisons made towards someone or something lower on status or ability, were also identified from Hawkmler's (1966) rank-order paradigm. Similar to other rank-order studies, participants completed a test of ability and asked whose score in the ranking they would like to see. Hawkmler manipulated this design by informing participants they had performed poorly on the test. In this condition, participants would often favor seeing the score of someone who performed worse. This was considered by Hawkmler to be defensive downward

comparison, and is some of the earliest work that led to the study of the interaction between affective states and comparison tendency. Goethals and Darley (1977) cited this research as evidence for the utility of social information for self-esteem (the role of research exploring downward comparisons in the incorporation of emotion into SCT is discussed further in the section on emotion).

Wheeler (1991) commented that early SCT researchers' interest with directionality may have been misguided. He noted that the issue of the direction of comparison is derived from Festinger's (1954a) universal drive upwards. Festinger seems to arrive at this conclusion from research prior to his publication of SCT, which identified that people set their level of aspiration as slightly better than average (Lewin, Dembo, Festinger, & Sears, 1944). This is curious since beside the Lewin et al. (1944) paper there is not much evidence for the drive upwards. Festinger himself did not seem to support a universal drive upwards (Wheeler, 1991). The concept is absent from his 1954 Nebraska Symposium on Motivation (Festinger, 1954b) and was mentioned only briefly in the formal publication of SCT (1954a). Regardless of the original intentions of comparison directionality, current SCT posits upward, lateral, and downward comparisons.

### *Similarity*

A second issue in social comparison literature relates to how comparison objects are chosen and the influence of perceived similarity. Goethals and Darley (1977) restated SCT to incorporate tenets of attribution theory to better describe the how an individual evaluates his or her opinions and abilities and how comparisons are chosen. The tenets of Festinger (1954a) that addressed these issues were Hypothesis I, III, and

VIII. Hypothesis I stated that “There exists, in the human organism, a drive to evaluate his opinions and his abilities” (p. 117). Goethals and Darley (1977) agreed that one does have a drive to reduce uncertainty about correctness of opinion and adequacy of abilities, but one also has a desire to believe positive things concerning their opinions and abilities as well as a need to have high self-esteem.

This need for a high self-image takes precedent over objective appraisal. One will choose a comparison strategy where they see themselves as doing well over doing poorly. Festinger (1954a) stated in hypothesis III stated: “The tendency to compare oneself with some other specific person decreases as the difference between his opinion or ability and one’s own increases” (p. 120). Additionally, hypothesis IV stated: “If persons who are very divergent from one’s own opinion or ability are perceived as different from oneself on *attributes consistent with the divergence*, the tendency to narrow the range of comparability becomes stronger (p. 133, emphasis in original). Goethals and Darley (1977) contended that these relate to the similarity hypothesis of SCT which describe one’s choice of a comparison other. They argued that similarity had been interpreted by other researchers too literally, as one with similar scores on abilities or opinions. Rather, the similarity hypothesis should be considered “when persons are perceived to be similar to oneself on attributes related to an opinion or performance, the tendency to compare with them increases” (p. 265). People decide who is a similar other first, then weight the comparisons accordingly- the greater the estimated similarity the greater the weight of the comparison. The previous interpretation of a similar other as similar on the compared ability suffered from a paradox because seeking the

performance of another was a comparison, yet it is not until one learns of another's performance that one can gauge if the person is indeed similar.

Goethals and Darley (1977) reconsidered SCT from the perspective of attribution theory. This was deemed important because in the actual evaluation of opinions and abilities one is really evaluating *dispositions*, not objective performance. Ability is non-observable and therefore not directly comparable. Because of this lack of direct comparison, the extent to which a person judges another to be performing well is attributional- one attributes a performance as part ability and part situational factors. Attribution theory provides insight into how one may "divvy up" performance, and therefore is helpful in understanding SCT. Attribution theory attempts to explain the methods one uses for evaluating the causes of others' responses as well as one's own responses (Kelley, 1973). Opinion comparisons are also under attribution but not in the same way as abilities because opinion statements are directly observable. However, evaluations of one's opinion systems- a person's values and intelligence- are essentially opinion comparisons but involve a greater degree of attribution about the causes of the opinion statements than single opinion comparisons. So while opinion comparisons do not involve much attribution, opinion systems do have a substantive attribution component.

Attribution theory and SCT have several commonalities. Both include a need to understand and predict the environment and base the quality of opinions and abilities on social information. The theory's emphasis on the evaluation of performance led Goethals and Darley (1977) to recommend the above-described modifications to SCT. They proposed that SCT's drive to evaluate opinions and abilities be expanded to

include a desire to validate them as well. The similarity hypothesis was expanded such that an individual will prefer comparing with those who are perceived to be similar on attributes that are related to their opinion or performance level. Similar others are not based on performance, but attributed ability. One who is perceived as dissimilar and performed better would have other traits that could be used as an explanation for the success (discounting). However, if the dissimilar other's performance was worse; this could be used to boost one's image if they considered the differences as inhibitory. Therefore, a dissimilar other could be used for comparison.

During the 1970s, the tenets of SCT were revised to incorporate affective states and, in doing so, revised the assumption of similarity (Wheeler, 2000). Goethals and Darley (1977) emphasized the role of self-esteem in selection of social information. SCT therefore expanded from self-assessment to also the use of social comparisons for self-evaluation. As described previously, people will make downward comparisons when they were told their performance was sub-standard (Hakmiller, 1966). Goethals and Darley (1977) suggested that this was done to avoid drawing the conclusion that one is inferior. They extended this notion to argue that there is a tendency to view others' as more advantaged in non-ability factors than reality, in order to maximize satisfaction with success and minimize upset with failure. The inclusion of downward comparisons led to the incorporation of attribution theory into SCT, to explain why downward comparisons would occur if people were attempting to objectively evaluate abilities (people are, but only to the point that self-esteem becomes threatened). SCT and self-esteem were now related in theory, and research explored the emotional influences of social comparisons as well as the influence of social comparisons on emotions.

### *Influence of Emotion*

A third development of SCT research is the exploration of the emotional consequences and antecedents of social comparisons. Again, studies of comparison direction facilitated this development. Wills (1981) identified downward social comparisons that were made for the purposes of self-enhancement. Wills theorized that subjective well-being can be enhanced when a comparison is made to a less fortunate other (Wills, 1991). He also suggested that people make lateral comparisons. He defined downward comparison as “comparison with a target who is worse off on at least one dimension” and lateral comparison as “comparison with one who is also experiencing problems, but is at the same level as the self on a given dimension” (p. 56-57).

Wills (1991) did not disagree with the notion of upward comparisons; but downward comparisons also occur, particularly when one is distressed and the objective situation is not immediately remedied by action. Therefore, if one cannot change their condition or position, one will instead change perception until one’s well-being is perceived as improved. For example, if one finds they have run into a financial crisis, an “it could be worse...” thought might be used to help ease the pain of the situation. A comparison direction, then, is based on the context. Upward comparisons occurred under select situations where the individual does not have much at stake in the outcome. Downward comparisons occur when one is stressed or threatened, and self-esteem is linked to comparison process. Lateral comparisons were hypothesized to occur in similar situations as downward. Wills (1991), however, identified two reasons that might influence preference of one type of comparison over the other. First, that there is a natural primacy of downward comparisons available. People generally prefer downward

comparisons to lateral ones. Second, the intended purpose of the comparison also influences which type of comparison will be made, such as if the purpose is for improving life-satisfaction or self-perceptions. Downward comparisons are likely superior than lateral ones for improving life satisfaction because they provide evidence that one's situation could be worse, and is therefore not as bad as it could be. Lateral comparisons are superior for self-perception issues because of a validation one's problems are common and occur in others. In the example of having a financial crisis, a lateral comparison would be thinking of other peers that also have financial problems, thus delegating the crisis to a "par for the course" mentality.

Not all downward comparisons are equal in impact. A downward comparison to a friend who is perceived as doing worse does not carry the same weight as a comparison to a homeless person seen on the street. The degree of discrepancy between a person and their downward comparison influences the amount of improved self-evaluation (Wills, 1991). Generally, downward comparisons improve self-evaluation no matter how different the person or group but there is a point where the compared is so far inferior in ability that empathy is felt instead. Typically, one makes downward comparisons with individuals who are slightly-to-moderately worse on a comparison dimension to improve life satisfaction. In the short-term, negative mood is reduced but positive mood is not necessarily increased (though Wills contended it is possible). Downward comparisons, therefore, are negatively reinforcing.

Wills (1991) did not make any specific postulates about the long-term effects of downward comparisons. However, life-satisfaction comparisons are likely based on life-situations, enduring factors such as income and family. As such, an enduring effect may

be created from short-term effects. Wills posited that upward comparisons are necessary for one's assessment of ability to be grounded in the real world, and that people are more likely to make such comparisons when the outcome is expected to be positive. Should this not occur, one might engage in downward comparisons to protect or enhance self-esteem.

Perhaps one of the most important follow ups to Festinger's (1954a) work on social comparison theory is Brickman and Bulman's (1977) exploration of the relationship between social comparison and emotion. A "google scholar" search conducted in May 2008 identified 211 relevant references to the author's work in multiple research topics including self-esteem (Rasmussen, Willingham, & Glover, 1996), career-oriented behavior (Buunk, Peiró, & Griffioen, 2007), race relations (Okami, 1992), body image (Eisert & Kahle, 1982), and self-perceived health (Powers, 2004). The reason for the popularity of this publication is the authors' emphasis on shifting the focus of social comparison to the consequences derived and the natural conditions under which the phenomenon occurs.

The initial interpretations of Festinger's (1954a) postulates of social comparison were perhaps an optimistic interpretation of the phenomenon. Social comparison was outlined by Festinger to be used for opinion and ability evaluation with similar others. Later researchers assumed a preference for upward comparisons because these were towards superior others and therefore one could look upon them and attempt to understand how one's own abilities might be raised. Brickman and Bulman (1977) acknowledged these conditions exist and furthered the expanse of social comparison theory by contending that the opposite of each of these postulates was also true: people

will also make social comparisons for self-evaluation, make comparisons downward, and will make comparisons to dissimilar others. A person is exposed to a multitude of social information and social comparison is possibly an inevitable element of social interaction. The social comparisons made and their interpretation is influenced by the perceived cost of the social comparison. Therefore, social comparison is selected based upon the intended goal of the person- self-evaluation or self-improvement.

The natural conditions by which social comparison occur were formally brought to the forefront of social researcher interest with Brickman and Bulman's (1977) work. The traditional approach of studying social comparison at the time was the rank order paradigm where participants take a test, are told where they fall in the distribution of scores, and are then asked who else's score they would like to see. Such research supported the desire for upwards comparisons (Wheeler, 1966) though had also demonstrated downwards comparison (Hawkmler, 1966). Brickman and Bulman (1977) criticized this approach for neglecting the role of social interaction and creating situations of "low cost"- comparisons in which participants are not particularly invested. For example, a somewhat well documented effect in social comparison research is the tendency of people to prefer to wait alone after receiving a test score than when they have not received the score (Gerard, 1963). This finding had been interpreted as a satiation of social comparison, but Brickman and Bulman (1977) countered this could also be evidence of social comparison avoidance.

Brickman and Bulman (1977) attributed the failure to identify avoidance of social comparison to the assumption that social comparison is only used to attain accurate appraisals of performance (Festinger, 1954a). Social comparison research had

been biased by an over-emphasis on rational processes. Another reason for the lack of studies of social comparison avoidance is the relevance on experimental settings, where avoiding a comparison is difficult or impossible. For example, in rank-order paradigm experiments the ability to avoid social comparison has been essentially removed, as the comparison processes have been made salient by virtue of the procedure. Social facilitation may also have contributed, as participants may have been unwilling to admit they were not interested in making comparisons.

As evidence of avoidance of social comparison, Brickman and Bulman (1977) provided general anecdotal evidence as well as experimental. As anecdotal evidence, they noted that social norms have emerged in society that limits one's ability to attain social information, such as the stigma about asking another's income or asking about sex. Such norms are evidence that people do sometimes avoid social comparison. This information, however, is still attainable. Information obtained through gossip has no self-disclosure to the referent. Common gossip topics are the same topics that are social stigmatized in polite conversation- money and sex. Gossip, then, is bootleg social information. Though impolite, gossiping is preferable to mutual disclosure because the social cost is lower. As experimental evidence, the authors reviewed research performed by Brickman and Kessler (1977), which tested preference for disclosure of test scores by performing a variant of the rank order paradigm. Instead of assuming a score revelation, participants were placed in groups that did or did not anticipate having to reveal their test score and whether or not the participant anticipated learning others test scores. This created the groups of mutual disclosure, self-disclosure without receiving social comparison information, receiving social comparison information without self-

disclosure. It was this emphasis on naturalistic conditions that facilitated the emphasis on the emotional consequences of social comparison. Specifically, how individuals seek social information while minimizing the psychological and social costs involved.

According to Brickman and Bullman (1977), people will affiliate themselves with certain groups, form coalitions, and compete based on maximizing a sense of self. This conception was neither supported nor rejected by Festinger's (1954a) original conception of SCT. Festinger did describe social comparison as occurring naturally and for the purpose of self-enhancement and people have a drive upwards for improving their abilities. It is unlikely, however, that Festinger lacked the cynicism to consider people as more interested in being seen well than objectively doing well. Festinger (1954a) stated, "those who could not view themselves positively would have deep feelings of failure" (p. 157). His subsequent dissonance theory did address a similar notion that consensus was more desired than accuracy (Goethals, Messick, & Allison, 1991). Dissonance theory posits that people will justify their feelings or behavior by adjusting cognitions, and is founded on the notion that perceptions will be altered for the purpose of maintaining or improving self-image. As with comparison directionality, it is possible that the development of SCT to include the influence of emotion may have been originally considered by Festinger. Regardless, incorporating affect into SCT expanded the purpose of social comparisons from being useful for attaining candid self-assessment to defending self-image. With self-esteem as an influencing role in the types of social comparisons as well as the conclusions drawn from them, other individual differences were considered for how they relate to SCT.

### *Individual Differences*

A fourth issue in social comparison literature is the role of *individual differences*. Individual differences refer to the ways individual people differ in their behavior and include personality traits (such as self-esteem) and perceptions (such as worldview). The variations in social comparison stem from a desire for individuals to view themselves positively. Modern SCT suggests that the type of comparison, the direction, and the conclusion will be selected to maximize positive assessment. This conceptualization implies that individuals will favor self-serving comparisons over realistic assessments up to the point a reconcilable conflict with reality is created. Put simply, people would rather believe they are doing well more than they want to do well. The term “well” is defined here loosely as “better than average”.

The utility of social comparison in self-esteem is constrained by reality. Klein and Kunda (1993) found that people would revise their own estimates of how often they engaged in a desired behavior when informed that the majority of people similarly engage in the behavior. Thus, people will create self-serving perceptions of others and themselves. Goethals and Klein (2000) stated that there is a “balance between the careful logic and rational that people sometimes bring to social comparison versus the irrational and biased thinking that just as often affects the way they compare to others” (p. 23).

The study of downward comparison and the emotional consequences associated with it also contributed greatly to an increased interest in the study of individual differences in SCT (Wheeler, 2000). Goethals (1986) explored reality constraints of SCT as related to the *prisoner’s dilemma*. The prisoner’s dilemma is a social experiment where a person role-plays they are a prisoner of the police after committing a crime and

is separated from their cohort and both are offered a deal (Poundstone, 1992). Whoever talks to the prosecution will be offered a reduced sentence. If one testifies (*defects*) for the prosecution while the cohort remains silent, the betrayer goes free and the silent accomplice receives the full 10-year sentence. If both remain silent (*cooperates*), both prisoners are sentenced to only six months in jail for a minor charge. If both choose to defect, each receives a five-year sentence. Participants in the prisoner's dilemma determine which the best course of action is. Goethals (1986) found that people playing the prisoner's dilemma would underestimate the proportion of people playing cooperatively, regardless if the participant played cooperatively. Both groups considered cooperation to be a positive ability. Thus, Goethals concluded that people underestimate the proportion of individuals doing well and overestimate the uniqueness of their own strengths, since those who cooperated also committed the underestimation, which he termed *uniqueness bias* (UB).

Goethals et al. (1991) defined UB as "reality-constrained self-deception" (p. 162). The UB was stronger for traits that are difficult to objectively measure (such as kindness and cooperation, "would" behaviors) than those more easily ascertained (such as intelligence, "could" behaviors). Thus, self-evaluations are self-serving but reality constrained. The level of reality constraint varies because the level of precision in the measurement of an ability or opinion will vary. For example, college students have more assessments of intelligence, such as standardized test scores and GPA, than many other individuals. Therefore, a college student would likely have greater difficulty in manipulating intelligence ability comparisons than a non-student who has fewer measures present in his environment. The college student would not be bound to

accepting the truth, per se, as shifting comparison groups could provide an escape to the conclusion that one is not intelligent. The reality constraint therefore also varies by comparison groups.

UB has also been identified in other populations including primary school attending children, high school children and middle management bankers (Goethals et al., 1991). Several variations in UB were identified. The degree of UB concerning intelligence across studies and was highest in the high school study. Overall, however, the UB on intelligence was lower than the UB for moral and athletic items. Gender effects were also identified. The males tended to differentiate themselves more (commit a greater UB) for athletic and intellectual behaviors while females tended to have greater UB for moral behaviors. The gender differences were consistent across the studied age groups. According to Goethals et al. (1991), realistic social comparison is more common when one is forced to confront comparative information. Laboratory experiments, such as the rank-order paradigm, manipulate comparison and therefore make differences salient. This laboratory design, however, may be too contrived to simulate day-to-day comparisons.

Wheeler and Miyake (1992) attempted to measure social comparison as it occurred naturally through diaries that participants kept for a two week period. The participants were told before the study began that noticing a similarity or difference with another is not necessarily a comparison unless followed by a psychological reaction (a contentious point explored later). When a participant felt they had made a social comparison they recorded the referent and dimension as well as their mood before and after the comparison. Prominent referents (>5% of total responses) were close friends,

ordinary friends, acquaintances, strangers, family members, imaginary persons, significant partner, and famous persons. Prominent dimensions included academics, personality, physical appearance, lifestyle, abilities, social skills, wealth, and opinion. All of these comparison groups and dimensions of comparison were provided by the authors with one exception. The “significant other” comparison group was identified as a prominent comparison group through an “other (specify)” option in the questionnaire. Sex differences were identified in comparisons. Women were statistically significantly more likely to compare with a family member and men were more likely to compare with a famous person. Women were also more likely to compare themselves on appearance and men were more likely to compare on opinion. In terms of affect on mood, participants were more likely to make downward comparisons on friends than anyone else.

The overall experience of recording social comparison was reported to be non-intrusive and reasonably valid. In post experimental follow up questions, participants reported little difficulty in recording comparisons, somewhat accurately, with little interference in their daily lives and they did not experience an increased sensitivity to social comparison in such a way as to change the number of comparisons. Participants also did not report hesitation to report comparisons due to guilt. Self-report data of such mental processes as social comparison may be difficult to measure, but this does suggest that comparative processes can be measured through self-reports.

The finding by Wheeler and Miyake (1992) that downward comparisons were made with friends is in conflict with Wills (1991) downward comparison theory which predicts the opposite. Also, upward comparisons were more likely to occur when one

felt bad and downward when they felt good, also in conflict with downward comparison theory. Wheeler and Miyake (1992) suggested errors in reporting could have influenced the results but proposed instead that mood states influence social judgments, and that mood influenced self-assessment more than assessment of others. In their study, self-esteem was measured before the 2-week period started. Those with higher SE made more downward comparisons and responded with more positive affect than those with low SE.

This incongruence in prediction of the effects of downward comparison was accounted for through individual differences (Wheeler, 2000). Wills' (1981) work on downward comparison suggests that low self-esteem individuals would make downward comparisons for self-enhancement. Beck's (1976) cognitive model of depression, however, holds that individuals with depressive symptoms are perceptually biased against themselves and use upward comparisons to maintain this bias. These are dysfunctional comparisons used to keep an individual from improving their self-worth. These two theories make different predictions, and neither is incorrect. Rather, one's state of self-esteem and depression both influence the kind of comparisons one makes. Wheeler and Miyake (1992) found that one's self-esteem prior to a comparison influenced the direction of the comparison. Those with high self-esteem made more downward comparisons and were less bothered by lateral and upward comparisons (their research is described in greater detail later).

Wheeler (2000) explored studies of social comparison and self-esteem and identified an enduring trend: Downward comparisons are more likely to be made by those with high self-esteem, are highly extroverted, and generally experience positive

emotions. Upward comparisons, conversely, are more likely to occur among those with low self-esteem (stable or unstable) and are *highly agreeable*. Highly agreeable people are those that are not skeptical of others and do not believe themselves to be superior to others and as such are more cooperative than competitive.

Neuroticism was the only trait found to predict affect of downward comparisons: the positive effect of downward comparisons was greater for those high in neuroticism, but neurotics did not make more downward comparisons than non-neurotic individuals (Wheeler, 2000). Wheeler argued that neurotics have a stronger need for validation that someone is worse off. Besides- or perhaps related to- neuroticism, other personality traits influenced comparison choices. People that are dominate, competitive, and ambitious compare downward when given a choice, as do extroverts. Though not as consistent, upward comparisons are made by those with unstable self-esteem.

Self-esteem appears to be the predictor as well as the outcome of the choice in upwards or downwards comparisons (Wheeler, 2000). High self-esteem is related to downward comparison. People that have high self-esteem are more likely to use social comparisons to maintain a high self-esteem when confronted with failure than those with low self-esteem. Conversely, low self-esteem people will be enhanced after success instead of failure and will make upward comparisons after poor performance (high self-esteem people will avoid these). Depressives are perhaps more willing to attain accurate information at the cost to self image, or will seek this information deliberately (in line with Beck's prediction). This incongruence is reconciled by acknowledging the role of self-esteem in choice of comparisons.

Beyond choice in comparison, self-esteem plays a role in the affect of comparisons as well. Self-esteem dictates not only the choice in comparisons but the conclusions one draws from them. Those high in self-esteem responded less negatively to upward social comparisons. Low self-esteem people did not tend to make downward comparisons to use social comparison to reduce distress, unless they were not given competing information. In these conditions, those with low self-esteem responded more favorable than high self-esteem. It seems that those manipulating social comparison to be seen in a favorable light are less likely to be depressed. Depressed individuals benefit emotionally from downward comparisons, but tend to make upward comparisons when available. Ahrens (1991) argued that in general, there are numerous social comparisons that can be made and people selectively attend to the large amount of information available in the world.

Social comparison is used to maintain self-esteem. This is in line with Beck's (1976) concept of depression. However, one may also manipulate comparisons to enhance self image, as Wills (1981) contended. The exact nature of why and how one uses social comparisons is not understood, though Festinger (1954a) contended that there is a drive to see oneself positively (again, these modern developments in SCT may be implied in the original consideration by Festinger, 1954a). For the current purposes, the important findings here are that comparisons vary along two dimensions- direction (upward, lateral, or downward) and comparison outcome (favorable or unfavorable). Individual differences determine when one seeks either dimension.

While the dimensions of comparisons are meaningful, the most important aspect of SCT may be the emotional consequence of the comparison (Alicke, 2000). When one

is forced to cede superiority to another, they will re-evaluate the other's ability level as much a higher than their own, known as the *genius effect*. Such findings illustrate that besides not seeing the world objectively; people try to see the world less objectively to maintain their self-image. In Goethals et al. (1991) review of UB, the phenomenon was highly correlated with self-esteem and negatively correlated with depression. Apparently, it is healthy to think of oneself as unique, even if this is not true.

### *Measurement Issues*

By the late 1990's, SCT had expanded to the point where identifying what was *not* part of the theory needed to be established. Wood (1996) gave a formal definition of social comparison derived from a meta-analysis of SCT research. This was not easy, as underscored by the author with a story of the Nags Head Conference on Social Comparison in May 1992, where 16 social comparison researchers continually asked what social comparison is without successfully agreeing upon an answer. Still, there was a "core value" of social comparison that was generally held- that social comparison is "the process of thinking about information about one or more other people in relation to the self" (p. 520-521). This definition contains two contentious phrases; *thinking about* and *in relation to the self*. The first, which individuals think about information, does not imply careful or conscious thought. The second phrase, where individuals consider information in relation to themselves, concerns at a minimum only identifying a difference or similarity between oneself and another.

The degree to which these two phrases can be expanded does not have a general consensus in the literature and therefore Wood (1996) described major processes that occur in social comparison to explore these disagreements. These are: *acquiring* social

information, *thinking* about it, and *reacting* to it. When one is acquiring social information, one may do so directly with another person or indirectly through such things as television and gossip. Wood noted that researchers varied considerably as to whether the acquisition of information involved focusing on a comparison target and examining similarities or differences to simply identifying those similar to the self. When thinking about social information, the process leads to making a judgment about the comparison and determining the implication of the comparison. When reacting to social information, one concern among SCT researchers is whether or not the comparison must affect the comparer to be considered a social comparison.

The process of acquiring information is an issue of contention among researchers because it calls into question the definition of a social comparison referent (Wood, 1996). While the core definition states that the comparison is at least with someone the individual makes direct contact, the definition can be extended. Other comparisons include indirect comparisons, those comparisons with people the individual indirectly attains information (such as from a coworker about another coworkers promotion) to fictional characters, stereotypes, and products of the imagination. As described previously with the research of Goethals, et al. (1991), social reality is not perceived accurately and a perception of this reality is constructed to serve one's comparative purposes. Because of the importance of indirect comparisons on social reality, Wood (1996) argued in favor of including such referents within the study of SCT, succinctly summarized by the explanation: "imaginary comparison others have a psychological reality" (p. 522).

The process of thinking about a comparison raises the issue of whether a comparison must be a deliberate effort to seek social information in order to be considered a social comparison (Wood, 1996). People may directly seek comparison information or encounter it in their environment. Social comparisons may not be deliberate or even conscious efforts. As reviewed previously with the work of Brickman and Bulman (1977), people may not be aware of the comparisons they make. Though not every encounter is a comparison, not every comparison is deliberate. Therefore, Wood (1996) did not believe comparisons must be deliberate efforts.

The process of reaction to a comparison brings into question the issue of whether a social comparison must have an impact to be considered a comparison. One may identify someone as doing better on a relevant dimension but not be affected. Whether or not this is considered a social comparison was identified by Wood (1996) as a divergence in the literature. Arrowood (1986) stated that the criterion of a comparison be that it changes the self-evaluation. Wood (1996) disagreed with this position, offering that in cases where one wants to confirm instead of change self-evaluation. Because of such situations, a comparison should not be validated by its effects.

Having argued for a rather inclusive mindset towards what is to be considered a social comparison; Wood (1996) outlined what he considered is not a social comparison and explored such boundaries through a summary and critique of the varying methods of measuring SCT. The methods of social comparison measurement fall into the categories of *selection*, *reaction*, and *narration*. Each of the three approaches has several paradigms and measures.

In the selection approach, social comparison is the dependent variable and the focus is on the processes involved in seeking information. Social information is made available to the participants by the researchers and comparison conditions are manipulated. For example, in a rank-order paradigm experiment participants take a test then are given a scored rank. They are given the opportunity to see the score of a different rank. If they choose someone who scored higher then this is an upward comparison, while a comparison towards someone scoring lower qualifies as a downward comparison. Early rank-order paradigm studies were the first to demonstrate downward comparison, though people usually selected to see the scores of those who ranked higher (Wheeler, 1966). Wood (1996) noted that this approach suffers from a lack of focus on the reasons for the comparisons.

The reaction approach uses social comparison as the independent variable and the focus is on the effects of the comparison. Davis' (1966) study of career goals in prestigious versus non-prestigious schools (mentioned in the introduction of this review) is an example of the reaction approach. Students who earn high grades at a college where it is easier to earn high grades will often have higher goals than a student who is just as qualified but attends a more difficult college. Reaction studies suffer from a similar problem to the selection approach in that it is difficult to ascertain if the changes found between groups are truly due to social comparison (Wood, 1996).

Like the selection approach, the narration approach also attempts to assess the social information attained in everyday comparisons, but it uses participant reports. For example, in a self-recorded comparison diary technique, participants keep a log of their comparisons. The experiment by Wheeler and Miyake (1992), in which participants

recorded their social comparisons, is an example for this approach. Wood (1996) felt the self-rated approach was unacceptable as a measure of SCT because of the inability of the researcher using this method to distinguish social comparison and social judgments. According to Wood, three factors affect the validity of narration techniques: lack of awareness, social desirability/self deception, and problems in selectivity and recall. The three problems of narration would be greatest with self reports because participants must first recognize and admit they make comparisons and the retrospective nature of self-report measures increase the risk of committing selection, recall, and aggregation errors. In aggregation errors, one's answers are skewed when they aggregate the information. Wood considered the self-rating approach of assessing social comparison as inappropriate because they are judgments derived in part from social comparison, but also from other processes such as self-serving attributions. Furthermore, comparative ratings may be made without actually considering social information. Even when asked to rate themselves to a comparison target they may simply answer following a heuristic where they always see themselves as better (Alicke, Klotz, Breitenbecher, Yrak, & Vredenburg, 1995).

Wood (1996) provided several recommendations for measuring social comparison. One was that social comparison measures should avoid global self-report measures because social comparison is mired with social desirability, self-deception, and lack of awareness of engaging in social comparison. If a method requires participants to admit they engage in social comparison, conveying permission for making social comparison is prudent. Another recommendation was to use the least constraining methods as possible. Wood suggested providing comparison targets that are available in

the participants' daily life, prominent people/groups that the individual likely uses for comparison. Wood concluded that overall, the study of social comparison has traditionally emphasized selection studies and as such social comparison literature has little knowledge of how people react to social comparisons.

Festinger (1954a) posited that the drive to compare is strong and the desire to have the superior performance is as great as the desire to achieve uniformity (this was his concept of a unidirectional upward drive). Subsequent research has suggested that in situations where one does not excel, an adjustment of perception will occur until a feeling of being better does occur (Alicke, 2000). The purpose of SCT is not only accurate information of social position but positive self-evaluation. People manipulate their interpretation of social information to boost self-esteem. Such manipulations take place by expanding a sense of self or marginalizing the deficiency. This occurs because people want to feel good. Successfully validating opinions and abilities is more satisfying but adjusting perceptions is often quicker and easier.

Festinger's SCT (1954a) suggests an inherent uncertainty in people's knowledge of their own opinions and abilities, which drives a desire to seek social information. Brickman and Bulman (1977) later argued that people seek social information for self-esteem management. Under these situations, social comparison does not lead to a desire to improve but a changing of perception until one comes out on top. Likely, social comparison is a reflexive mechanism that gives information of one's standing. When one is challenged by the findings, they may strive to do better or adjust their perception to accommodate a threatened self-image. The expansion of SCT to incorporate *affect* resulted in research of the *effect* of SCT. This shift created an overlap with SCT and

other social process theories that had been exploring the relationship between emotion and social information.

### The Reactions to Social Comparisons

During the 1970's, SCT formally shifted to include the study of social information and emotional influences. The interest in the consequences of social comparison marked what Buunk and Gibbons (2000) considered the "renaissance" of SCT. As discussed previously, the incorporation of SCT with attribution theory revised the predictions of who is selected for comparison (Goethals & Darley, 1977). When SCT researchers became interested in the effects of social comparison, an effort began to further connect SCT to existing social process theories. Following the formulation and extended research on SCT as the prominent theory that explains the way humans process social information, other alternate theories have emerged that focus on the consequences of social comparison self-evaluation maintenance (SEM) and relative deprivation/gratification (RD/RG).

The theories of SCT and SEM set the stage for explaining when and how one will utilize social information and feel satisfied (relatively gratified) or dissatisfied (relatively deprived). An extensive portion of section II explores the history of RD because the concept has enjoyed a longer but more contentious research history than SCT (RG has only recently been explored in depth). Additionally, RD and RG are argued here to function as intermediary variables between social comparison and health. A comprehensive understanding of the history of RD is necessary for understanding how

perceived deprivation can affect health and why in some situations it might not. This review of the consequences of SCT is organized in three sections:

- Self-evaluation maintenance
- Justice and belief in a just world
- Relative deprivation and gratification

### *Self-Evaluation Maintenance (SEM)*

As SCT expanded to include emotional states, the theory complemented a related social process, the self-evaluation model of social comparison (SEM). SEM adds that comparisons to others can create a sense of pride (Tesser, 1988). Tesser argued that SCT neglected to evaluate the relationship between social comparison and emotion. Specifically, he argued that social comparisons were important for defending and maintaining self-evaluation. When one encounters someone who is doing better, SCT predicts a threat to self-esteem. However, sometimes one will be filled with pride that another is doing better. Because of this, Tesser (1988) argued for a theory that explored *self-evaluation*. Self-evaluation refers to the way a person views him/herself and attempts to predict when a specific emotion will be felt as a result. For example, in some situations individuals feel jealousy when outperformed while in other situations pride is felt. SCT predicts that a person will compare themselves to others and conclude if the comparison is favorable or unfavorable. This results in SEM and SCT predicting different outcomes from a comparison. For example, a person may brag if their friend has a prestigious job or celebrate the success of one's favorite football team. Under SCT, one would be predicted to feel distressed because they do not have a prestigious job and

are not a successful football player. SEM accounts for the pride experienced by the success of another, termed *reflection*.

There are two requirements for reflection (Tesser, 1991). The first is that the accomplishment has to be very good. The second is there needs to be a psychological connection. Tesser referred to this connection as *closeness*. Comparisons are more likely as closeness increases. SEM is similar to SCT because it suggests that levels of closeness influence the expression of social comparison, analogous to the similarity hypothesis of SCT. Closeness is different than similarity, however, and people may make comparisons based on groups that are perceived as similar or close (Suls & Wheeler, 2000). SCT and SEM are at odds, Tesser (1991) contended, because in situations where an individual underperforms, SCT lowers self-esteem while SEM would raise self-esteem through closeness. An individual's sense of self weighs in the processes and individuals will engage in either SCT or SEM to maintain a positive self-evaluation. When encountering a comparison of lowering self-evaluation one either strives to improve ability or works to reduce the weight of the comparison.

Under SEM, comparisons are made with a close other; someone the comparer is tied to emotionally (Tesser, 1991). The degree of emotional connection to a person is the *distance*, which a greater distance being for those one has less emotional connection. This close other does not be similar in ability. This distance to comparison groups was also discussed in research by Miller, Turnbull, and McFarland (1988) which described two forms of evaluation: universalistic and particularistic. Universalistic evaluations are assessments of standing made towards others in general while a particularistic evaluation is an assessment made with others whom one feels a bond. Tesser (1991)

described SEM as being capable of measuring particularistic evaluations well because of the construct of closeness, something SCT could not accomplish.

SEM is based on the interplay of other's performance, closeness, and relevance (Tesser, 1991). These three relate with one another such that they function both as causes and consequences (Tesser compared SEM to Boyle's law which holds that temperature, volume and pressure interact in the dynamics of gas). For example, when relevance and other's performance is high, closeness will decrease. Under SEM, one would predict that when an individual is outperformed on a dimension, the relevance of that dimension is decreased. Tesser supported this multi-directional model through previous research. For example, Tesser and Smith (1980) found that in a guessing task, subjects gave their friends easier clues than to strangers but in a high relevance situation subjects gave their friends harder clues than strangers. Relevance moderates whether an individual is in a state of reflection or social comparison. If relevance is high- social comparison will emerge and a person will feel threatened. If relevance is low, reflection would emerge. In both situations, Tesser (1991) noted that one would find an increase in arousal but the direction of that arousal-the felt emotion- would be different.

To support this, Tesser (1991) described previous work in which researchers dichotomously manipulated the dimensions of relevance, performance, and closeness by assessing the feelings derived from situations where respondents either outperformed or were outperformed on a task that was either important or unimportant against either close or distant (Tesser & Collins, 1988). In situations where one was outperformed by a close other on a highly relevant dimension, jealousy was reported. The closeness was not statistically significant however, but relevance and performance interacted. Pride

was felt when one outperforms another and increases with relevance. Closeness did not have an effect on pride and was higher in the low relevance group. Tesser (1991) argued that this supports the notion of reflection. Performance, relevance, and closeness acted in response to one another, so as to maximize self image.

Classical SCT may be a sound theory for select contexts. Specifically, when an individual is interested in sincere efforts to improve control over their environment or understand their position. The concept of SEM works in many ways fill in the gaps left in SCT, such as providing the dynamics of how one desires to feel better from upward comparisons but is not motivated to improve. Beach and Tesser (2000) are diligent to explain that SEM is separate than SCT, not a competing theory. SEM is distinct from classical SCT in that when one engages in SEM the aim is to maintain self-esteem rather than reduce uncertainty about abilities. Both constructs are influenced by the degree of an interpersonal relationship with the comparison other. In SEM this dimension is closeness while in SCT it is similarity. Whether one engages in SEM or SCT depends on how strongly the dimension is important to the comparer's self-identity. The more important the dimension to one's self-image, the more likely SEM will be engaged to defend self-esteem. Additionally, as the degree of closeness with the individual increases so too does the likelihood that SEM will be favored over SCT. However, as SCT expanded to incorporate emotional evaluation, the distinction between SEM and neo-social comparison theory becomes difficult to identify. Suls and Wheeler (2000) agreed that SEM is separate, but that both theories have social comparison elements. The neo-SCT incorporates affect and downward comparisons, while SEM expands upon how one may derive positive self-image from any comparison.

### *Justice and Belief in a Just World*

The expansion of SCT to include emotions complemented other existing theories (as with SEM). SCT's development also provided explanatory mechanisms to other social processes, such as *equity theory*. Equity theory holds that a perception of inequity creates a negative emotional state (Austin, 1977). This negative feeling occurs for both the sufferers and the beneficiaries of the inequity, in the forms of anger and guilt, respectively. Though, the beneficiary experiences less distress than the sufferer. Austin related equity theory to social comparison through the argument that justice is defined by social comparisons. One evaluates the fairness of a situation by appraising situations of those deemed as similar. As an example, Austin described students comparing their grades in a class with their classmates to get a sense of whether or not the teacher is fair. When inequity cannot be removed and balance restored, one will attempt to justify the inequity. The inequity that cannot be effectively rationalized creates distress.

Belief in a just world (BJW) research follows people's desire to believe that people get what they deserve (Lerner, 1965). BJW theory holds that people develop general schemata that create general theories about reality. If one's theories suggest that the world is a just place then one has a higher BJW. This belief allows individuals to approach the world as if it is stable and orderly. Belief in a just world has been tied to mental health. Dalbert (1999) argued that BJW had a positive impact on mental health by creating feelings of competence and control, fostering the notion that good deeds will go rewarded (thereby allowing people to invest in their own good future), and fostering investment in long-term goals. By creating a stable and positive way to view the world, BJW enables people to feel that they have control over their lives.

BJW has undergone refinement and has been divided into a general belief in a just world and a personal one (Lerner & Miller, 1978). Personal BJW centers on the individual and is the belief that one's own fate is just, that one's own local environment is fair and just, that life is fair for oneself and one's own family and friends. General BJW is focused on the world at large and is the belief that the world in general is fair and just for all people. Lipkus et al. (1996) compared personal BJW and general BJW and found that personal BJW was a better predictor of life satisfaction, stress, and depression than general BJW.

#### *Relative Deprivation and Gratification*

As discussed previously, Wood (1996) contended that the consequences of social comparison were not understood. SEM and justice theories suggest SCT's relationship to self evaluation and worldview but still do not directly address the consequences of social comparisons. A possible set of reactions to social comparison may be relative deprivation (RD) and relative gratification (RG), two concepts that may be independent but related to modern SCT. RD and RG are discussed thoroughly here, as it is my position that RD and RG is a consequence of SCT. Furthermore, this model can be useful in accounting for perception-related effects such as status-related health inequities. While SCT had been progressing towards extending social comparisons to emotional consequences, RD theory originally focused on emotional consequences under select social comparisons (though this overlap was not originally identified). Research concerning the theory of RD arose to account for dissatisfaction with situations that did not seem to logically create dissatisfaction. This illogical dissatisfaction was in

time attributed to the source of comparisons, and was extended to account for a wide range of psychological phenomenon.

The text credited with beginning the formal study of RD was *The American Soldier*, a four volume series of books summarizing social research performed on the U.S. army throughout WWII by the Information and Education Division (IED) of the U.S. army. The IED commissioned Stouffer, Suchman, DeVinney, Starc and Williams to carry out a review of the studies and write up the results, published in 1949. At the time, this was seen as a natural extension of the continuing need for psychological understanding in war. For example, the Alpha test was adopted during WWI as a measuring device of ability. Stouffer et al. contended that the purpose of their work was to move forward from the study of abilities to perceptions. They stated: “Just as World War I gave new impetus to the study of human *aptitudes*, so World War II has given new impetus to the study of *attitudes*” (p. 5, emphasis in original).

Stouffer et al. (1949) contended that their work was intended for three audiences: Army officials, historians, and social scientists. They noted that social scientists are their primary audience and identified several benefits of army studies to the social sciences. One was a very low dropout rate, as soldiers were commanded to participate. Another was a study of class. Stouffer et al. described the time and sample of the study as very unique because with WWII beginning there were massive increases in soldier size. Also, because of conscription, many soldiers were not necessarily happy to be there so there was an excellent opportunity for the study of frustration. Furthermore, the environment was different than civil institutions in three ways. First, the authoritarian nature of the armed services is far greater than any civil institution. Second, the pronounced stratified

social system in the military. Third, in the armed forces there is an increased emphasis on traditional ways of doing things and a discouragement of initiative.

Stouffer et al. (1949) considered these conditions, though extreme, were not unique to the military. They gave the example of a factory worker during an industrial depression. Such a person would have little control and be in a very similar situation to a WWII soldier in terms of inability to leave. Therefore, the situation of the Army was an ideal environment to study the relationship between status and attitude. Like Marmot's Whitehall studies that found changes in employment grade were linked to health effects (Marmot, 2005), the soldiers in Stouffer and colleagues (1949) study had little perceived control and are limited in their social participation. SCT attempts to explain the relationship between comparisons and emotion while health inequities research explores the effect of status on health. The work of Stouffer and colleagues (1949) was an early effort to study the effect of status on attitudes and serves here as the research bridge between SCT and health inequity research.

Volume one of *The American Soldier* described the attitudes of soldiers across the span of the war as well as demographic variables such as education and race. Six hundred pages detailing variations in soldiers' attitude does not sound like enjoyable reading. To the contrary, Stouffer et al. (1949) demonstrated not only a literary skill but also an enviable ability to convincingly generalize findings to the general population. For example, the following excerpt illustrates the breadth to which Stouffer and colleagues were applying their findings:

Although it is a popular stereotype to view American civilian culture as a classless society, a host of sociological investigations have shown how important

the class lines are in our social system. Perhaps the greatest distinction between American and the most older European cultures is not the absence of class lines in America as compared with Europe but rather the greater extent of vertical mobility from one class to the next above. This vertical mobility has been made possible throughout American history by the progressive increase in standard of living, through the opening of new lands, and through increase in productivity per man hour by unparalleled use of machine power. It has been facilitated also by the differential birth rate, which encourages mobility from the fertile lower socio-economic strata to fill vacuums left by low fertility in the top classes. To climb the socio-economic ladder is the American dream, and those who do not succeed still are psychologically identified with the process through projecting on their children their own unfulfilled ambitions. (p. 244)

This passage demonstrates a generalization of Stouffer et al.'s (1949) research of soldiers' attitudes of promotion to the desire for upward mobility in the United States. When one perceives themselves as capable of improving in social class, they are often motivated to improve, and when they are denied it frustrates them. In the first volume, Stouffer et al. used the term *relative deprivation* in reference to frustration with promotion rate. Because of the structure of the Army, one's rank was analogous to status and therefore promotion was the mode of social mobility. As the following passage illustrates:

(The) only prospect of social mobility in the army was to climb the ladder of noncommissioned grades... Not only did they represent better pay, freedom from many

menial and irksome tasks, and frequently other special privileges as respecting quarters or passes, but they also were a badge of success. (p. 247)

Three samples of soldiers were surveyed in the fall of 1943 as they entered the service and again four months later to ascertain rank. Attitude was measured by the subset variables termed: *personal esprit*, *personal commitment*, *soldier-war worker job satisfaction*, and *rating of how well the Army is run*.

Personal esprit was measured by Stouffer et al. (1949) through ratings in 3 items covering overall attitude about army life (ex. "In general, how would you say you feel most of the time, in good spirits or in low spirits?" p. 95). Personal commitment was measured through 4 items (ex. "Considering everything, how do you feel about further service in the Army?" p. 96). *Satisfaction with status and job* was derived from 7 items (ex. "Do you consider your own present job or duty in the Army an important one in the war effort?"). Approval or criticism of the army was derived from 13 items (ex. "How much of your training or duty time is used in doing things that do not seem important to you?"). Possible answers varied across items, ranging from three to five options. Responses were deemed as favorable or unfavorable and analysis performed across education grades. The better educated (as defined by having completed high school) tended to have higher scores in personal esprit and personal commitment than less educated, while also holding less favorable ratings in satisfaction with the status and were more critical of the army. This effect was described in terms of RD: With higher levels of aspiration than the less educated, the better educated man had more to lose in his own eyes and the eyes of his friends by failure to achieve some sort of status in the Army. Hence, frustration was greater for him than for others if a goal he

sought was not attained- and this happened often, indeed, as a consequence of the kind of criteria in which the Army traditionally employed in selecting enlisted men for promotion. (p. 154)

Cross-class differences were found for the dissatisfaction (Stouffer et al., 1949). The military had many branches, and these branches were perceived differently both within and across one another. Military police (MP) were more satisfied with the promotion rate than Air Corps men, though the actual promotion rate was higher with the Air Corps. The men that had not been promoted in the Air Corps were more frustrated than non-promoted MP because the reward seemed more available. Those promoted in the Air Corps did not find the promotion as meaningful as those in the MP. There was greater dissatisfaction for the non-promoted and less satisfaction for the promoted in the Air Corps than the MP. This phenomenon was considered RD.

Within a branch of the military, those most critical of promotion rates showed the least job satisfaction (Stouffer et al., 1949). However, across branches job satisfaction was still higher in the Air Corps than MPs. This was attributed to the prestige of each branch. Air Corps was a higher prestige branch than MP. This was considered to be due to the greater opportunities for social mobility, and other factors such as the ability to learn skills useful in civilian life in the Air Corps than the MP. The authors also introduced RG, as the authors' caution against an attempt by the military to increase promotion rates in the hopes of job satisfaction. They stated that such a policy would reduce RG of the successful men instead of raising it.

The MPs felt that as a branch of the military they were discriminated against, while the Air Corps were held in high regard. Air Corpomen felt greater RD than MPs

when they were comparing themselves among one another (as they were when considering their satisfaction with promotion rate) but the reverse was true when they were comparing themselves against all other soldiers (as they were with rating job satisfaction). There are three main points to be drawn from this finding. First, who one uses for comparison varies by the comparison dimension. For promotion rate, soldiers compared themselves within their branch while for job satisfaction they looked at a “wider net” of soldiers. Second, the attitude regarding a single dimension may not represent an entire sense of satisfaction or dissatisfaction. If one chose to simply assess attitude about promotion rate, the Air Corpsmen would appear unhappiest of the military branches and predicted to be least satisfied with their job of any branch, when in fact the opposite is true. Third, which is a conclusion of the first two; is the role of prestige in evaluating one’s satisfaction. If one were looking for a general indicator of happiness, then job satisfaction is probably a more valid measure than attitude about promotion rate. Recall that prestige played a substantive role in assessing job satisfaction. Assuming this is accurate then what mattered to soldiers when evaluating satisfaction is what they believed *other* people thought about their status. These results suggest that one makes comparisons for themselves and as a member of a group (a concept that is explored later in this text with Runciman’s (1966) distinction between fraternalist and egoist RD).

The concept of RD was also employed by Stouffer et al. (1949) as an explanatory mechanism in a finding concerning the adjustment level of Blacks from the North to Army life. For those Black soldiers from the North stationed in the South, they reported resentment towards the discrimination they found in the South. Despite this,

they reported greater adjustment and satisfaction with Army life than fellow Northern Black soldiers stationed in the North. This finding was aggravating to the research team, as Stouffer would later describe:

Some of our analysts were almost in despair at this discrepancy. They actually held up the report on their study for over a month while they checked and rechecked in the vain hope of finding errors in the data or analysis to explain the paradox. (Stouffer, 1950, p. 199)

RD explained this paradox. The Northern Black soldiers were comparing themselves to Southern Black soldiers. The discrimination they encountered was felt as temporary for them, as after the war they would be returning home but recognized that the Southern Black soldiers they worked with would be staying in that environment. As Stouffer et al. (1949) stated:

Putting it simply, the psychological values of Army life to the Negro soldier in the South *relative to the Southern Negro civilian* greatly exceeded the psychological values of Army life to the Negro soldier in the North *relative to the Northern Negro civilian* (p. 654, emphasis in original).

Unfortunately, subsequent research confirming this belief was not available because the data collection had been performed years before this analysis. The failing was recognized as an unfortunate condition of the applied research (Kendall & Lazarsfeld, 1950).

The original academic work of Stouffer et al. (1949) describing RD was post hoc. The curious finding in military satisfaction led to the supposition that surrounding levels of reward influence dissatisfaction. This concept implies social comparison but

does not specifically state that comparisons are involved in the formation or maintenance of RD. Stouffer et al. did not formally define RD at all, and many subsequent researchers provided their own definitions. Merton, one of the researchers for *The American Soldier*, later described RD as a term to “account for feelings of dissatisfaction, particularly in cases where the objective situation would at first glance not seem likely to provoke such feelings” (Merton & Kitt, 1952, p. 52). This early description of RD suggests it is a condition of irony, where one feels the opposite of what their objective situation “should” create. From the onset of RD theory a relationship to SCT can be drawn. If one has an expectation that is fulfilled one would feel satisfied, while if an expectation is not met dissatisfaction would be the result. If RD is dissatisfaction, and this dissatisfaction is derived from expectations, and expectations are themselves determined by social comparisons, then RD is manipulated by social comparisons.

Later research formulated RD as related to expectations. For example, Davis (1966) applied RD to college students’ perceptions of career prospects. He analyzed a set of data that studied the careers of 35,000 college students from 135 colleges and universities gathered in 1961. Similar to branches of the military, Davis split colleges by level of prestige, based on the average scores of entering freshman on a standard test used for scholarship qualification. Schools were divided into four levels by distribution of scores. Davis found that GPA correlated strongly with choice in career after graduation and was stronger than the measure of scholastic aptitude correlated to career choice. Scholastic aptitude can be considered an objective measure of ability across all high school students while GPA is a measure of ability within college. GPA is the

objective standing in a school, but one may use GPA as a ranking of overall performance (this is perceptive since graduates in the job market would be competing against other students from other schools).

That career aspirations were not influenced as much by scholastic ability as GPA suggests that the students did not evaluate their ability by comparing themselves to all high school graduates or across other colleges but to other fellow students (Davis, 1966). Their objective standing was not as influential on career aspirations as their perceived standing. Other high school students that one has never met from schools one has never been to could be useful comparisons for assessing ability, but they are not as salient as fellow students. Davis stated: “students judge themselves by local standing” (p. 27). He argued that based on these findings, the popular notion of getting into the best possible school might not be the best approach in terms of attaining occupational mobility (occupation and occupational mobility are likely prominent components of one’s evaluation of status).

Among high school graduates within a specific level of scholastic aptitude, the more selective the college the lower the student’s GPA (Davis, 1966). This is not surprising; if two students from the same high school have identical GPAs the one attending Yale will probably have a lower GPA than the one attending the local community college. But, with this research in mind, consider the implications. The Yale student will have less aspiration, and while the community college student may not achieve all that they aspire to, the Yale student may not aspire to as much. This could create frustration from either one. The Yale student may be frustrated because her fellow students have achieved more, while the community college student may be frustrated

that she did not achieve what her previous self had hoped. Conversely, the opposite may occur if different comparisons are made. The Yale student, comparing herself to her high school friends, may feel satisfied that she is doing better, while the community college student may feel worse due to the same choice in comparisons. As with SCT, one's selection of comparisons, then, would be the deciding criteria and the implication of this data is that those surrounding us are our primary source of comparison.

Promotion rate and career aspirations are positive dimensions. Gurr (1968) also considered satisfaction and perceptions of social mobility as meaningful to RD, who used RD to account for the conditions that can cause civil strife. Gurr defined RD as: "Perceptions of discrepancy between their value expectations (the goods and conditions of the life to which they believe they are justifiably entitled) and their value capabilities (the amounts of those goods and conditions that they think they are able to get and keep)." (p. 1104)

Gurr (1968) argued that a reaction to perceived inequity is anger, and anger is a motivating state for aggression. However, participation in civil strife is mediated by the intensity of the RD. Using data from 114 political entities and data on civil strife in those areas between the years of 1961 to 1965, Gurr correlated civil strife to measures of persisting and short-term economic, political, and sociocultural deprivations. Thirteen deprivation measures were used; all considered a form of RD. Six of these measures were used to represent persisting deprivation and seven for short-term deprivation. Persistent deprivation included such variables as dependence on private foreign capital (measured by calculating payments to other countries as a percentage of the Gross Domestic Product) and religious cleavages (as measured by the number of organized

religious groups). Short-term deprivations included such variables as adverse economic conditions within a three-year span (evaluation of the number of new reports reporting such conditions as unemployment and market slumps) and inflation (as measured by measured by direction of cost of living). Eight of the deprivation variables accounted for two-thirds of variance in civil strife ( $R^2 = .65$ ) and intensity variables did account for civil strife significantly ( $R^2 = .43$ ). Gurr noted that this data was not a direct test of relevance since aggregate data was used but challenged one to explain the findings and considered this an excellent starting point. He argued that deprivation alone was not enough to cause strife. What was needed was a sense that one deserved something-entitlement.

Again, SCT provides an explanatory mechanism. One likely determines what they have coming to them, entitlement, largely by the groups they ascribe themselves. For example, one who considers herself working class probably does not believe she can go weeks without working and still pay her bills.

RD has also been previously studied in relation to status. One of the most cited studies of RD was Runciman (1966), who measured class distinction in Britain and attitudes towards inequality. As Marmot (2005) would later argue with social participation, Runciman (1966) believed that one's activities are a mark of status. Runciman stated:

There may be no inherent inequality in the preference of one man to watch polo instead of dog racing, to drink wine instead of beer, or to wear a bowler instead of a cloth cap when he goes to work. But such differences do directly affect the pattern of inequalities, since however neutral they may be as simple preferences

of taste, their effect is further to compartmentalize different strata into status-groups whose divergent styles of life inhibit their recognitions of each other as equals. (p. 95)

Runciman (1966) contended that education was both the mode of social mobility and a contributor to RD. Citing the interest of Stouffer et al. (1949) in education as a predictor variable of the relationship between promotion and attitude, Runciman (1966) believed that there was a parallel to this finding occurring in Britain. Prior to WWII, few children of manual laborers received secondary education. Therefore, people that had received secondary education were rarely used as reference groups (either comparative or normative). As education became more available, families had their children enrolled in secondary education, increasing the use of comparisons *by the parents* on individuals that had completed secondary education. The working class families felt more RD because they increased upward comparisons because their children were members of that dimension (closeness).

The magnitude and frequency of the RD for manual laborers had risen, even though objective inequality had actually decreased from greater access to education (Runciman, 1966). Furthermore, mass communication was increasing and was exposing the laborers to the possibility of a middle-class life. Runciman could not formally test this, as the research was not time series. Rather, he sought to measure class identification and perceptions of how one was doing in relation to others. He was curious about how manual workers and non-manual workers ascribe themselves and their perceptions of how they are doing in relation to others. He predicted that those at the top of their class would report the lowest level of RD. Manual laborers at the top

within their class would feel less RD than those at the lower level and those at the lower level of the non-manual labor workers because the manual laborer does not use non-manual laborers for contrasting comparisons.

To evaluate RD, Runciman (1966) relied on self-reported income. He did not favor objective measures of income, stating that estimates of income are more meaningful than objective measures for the purposes of assessing RD. The research consisted 1415 participants that had been classified as manual or non-manual by reported occupation (919 manual and 496 non-manual). He first attempted to determine who people used as a normative reference group with the question “What sort of people are you thinking of when we talk about “people like yourself?”” (p. 162). Non-manual workers most commonly used personal criterion (such as “good sorts” and “self respecting” people) while manual workers used class distinctions (such as “working class” and “manual workers”).

Runciman (1966) found difficulties in measuring status, as the class people belonged to and *believed* they belonged to was not necessarily the same. One question asked the participants to describe what class they considered themselves, and another question asked what they believed to be the middle class- non-manual or manual workers. Many people in manual labor jobs considered themselves as middle class (22%) while many people in the non-manual jobs considered themselves as working class (19%). This suggest that studies of status may be measuring something different than objective measures of class, as people do not necessarily ascribe themselves to the class their station in life suggests.

As a measure of RD, Runciman asked: “Do you think there are any other sorts of people doing noticeably better at the moment than you and your family?” (p. 192). For individuals at the bottom tier for income, roughly 1/5<sup>th</sup> stated “no” (21% of low income non-manual workers and 17% of low income manual laborers). Runciman summarized the finding as such: “It is still remarkable that so many of the poor should be unable (or unwilling-in either event, their reticence is remarkable) to think of others who are doing better.” (p. 193).

Perhaps they were aware of others doing better *financially* but not doing better overall. Just as how promotion rate was not indicative of overall quality of position, financial comparisons may not be used the same by everyone as a measure of success. Those in the high self-reported income group had similar responses as well, regardless of status as manual or non-manual work (Runciman, 1966). This, Runciman argued, was indicative of the selection of referents because although both groups reported doing better than anyone with similar frequency (31% of non-manual laborers reported that there are no other people doing noticeably better at the moment compared to 39% of the manual high wage earners), the wage of a high earning manual worker was objectively less than the wage of a high earning non-manual worker. Thus, though earning less, manual workers experienced less RD than non-manual workers.

Beyond illustrating the difficulties in measuring status and the importance of reference (comparison) groups in perceived advantage and disadvantage, Runciman developed a widely used definition of RD:

A is relatively deprived of X when (i) he does not have X, (ii) he sees some other person or persons which may include himself at some previous or expected time, as having X

(whether or not this is in fact the case), (iii) he wants X, and (iv) he sees it as feasible that he should have X. (pg. 10)

Many of the terms used by Runciman have parallels to SCT. The “some other person or persons” having “X” were referred to by Runciman as reference groups. The term *reference group* is derived from Hyman (1942), who used the term to describe those groups one uses for status comparison. An individual uses reference groups for comparison to determine their own successes and shortcomings, and people can have many reference groups at any given time. Hyman stated that while someone may have a large number of reference groups, the number of habitually used groups is small and likely to relate to specific problems. This is conceptually identical to the Festinger’s (1954a) comparison groups.

Runciman (1966) distinguished between *comparative* and *normative* reference groups, those that people contrast against their situation and those from which people take their standards, respectively. These two groups can and do overlap. The membership reference group, the one used for normative comparisons, can vary as people may consider themselves as belong to many different groups. A middle-class Caucasian father in the United States may draw normative comparisons to other members of his income-level, other Caucasians, other Americans, and other humans.

Though not stated by Runciman (1966), the concept of normative and comparative references is similar to Festinger’s (1954a) SCT of abilities and opinions. Festinger (1954a) contended that we desire conformity with opinion but to be better in abilities. Runciman’s (1966) argument can be reconstructed under SCT to be that opinions are normative, and abilities are comparative. As further support for this

merging, Runciman argued that RD was considered a consequence of comparative references. He stated: “If relative deprivation is to be precisely described, all inequalities which give rise to feelings of relative deprivation must be treated as inequalities between and only between the membership reference group and comparative reference group” (p. 14).

Runciman (1966) also expanded the concept of RD to include two subtypes: *egoistic* RD and *fraternalistic* RD. Egoistic RD is a function of comparing oneself to an in-group, while fraternalistic RD is derived from comparing one’s in-group to an out-group. Such distinctions are analogous to Festinger’s (1954a) notion of similarity, and Tesser’s (1991) emphasis on closeness.

Runciman (1966) considered one’s sense of justice as the psychological variable that can facilitate or impede the creation of RD from objective conditions, and a theory of justice is necessary for creating an appropriate assessment of RD: “Whatever meaning is given to ‘justice’, the appeal to justice will distinguish those feelings of relative deprivation which can and which cannot be properly described as a sense of envy rather than the perception of an unfulfilled right.” (p. 252). This is similar to Gurr’s (1966) notion of entitlement. Justice would regulate a sense of entitlement. The two however, appear to disagree on the causes and measures of RD. While Runciman (1966) considered one’s choice of membership groups as influencing comparison choice (and therefore RD), Gurr (1968) did not emphasize the importance of social information, only a sense of what one feels entitled to. Davis (1966), similar to Runciman, considered RD as a shorthand expression to describe the influence of one’s immediate surroundings on gauging ability.

Gurr (1970) made the case for objective measures to be used for RD.

Expectations, he argued, could be measured through “status quo” indices such as wage and salary data as well as national inflation rates. If, for example, these statistics demonstrated a decline in economic prosperity, they could be considered equivalent to expectations not being met by a declining ability- decremental deprivation. Rising expectations could be measured through increased rates in schooling and literacy as well as promises of reform from political leaders. Gurr also suggested that interpersonal RD may be measured through migration and religious affiliation. This measurement approach of aggregate objective condition data allows for comparisons across groups (such as across countries) and can be performed with existing data. However, as Merton and Kitt (1950) described, RD is a conclusion derived from comparisons to referents, and one has no guarantee that objective conditions are indeed creating RD. For example, if one experiences wage reduction, RD may be felt or one may feel RG because at least they still have a job. Without a consideration for individual interpretation of objective conditions, RD may be inferred but cannot be measured.

Stouffer et al.'s (1949) analysis of attitudes on promotion rates found that one's perception sometimes disagreed with objective conditions. Furthermore, the rate of mobility within a field does not necessarily correlate to a satisfaction with status. Though MPs were happier with promotion rate, they were less satisfied because in the “big picture”, they had one of the least prestigious positions in the army. This influence of prestige is similar to Davis' (1966) finding that GPA was a better predictor of future career aspirations than actual academic caliber. Runciman (1966) found a similar distinction as people tended to compare themselves along the similarity of their ascribed

class. Gurr (1970) identified that the consequences of frustration of perceived inequity could be civil strife and violence, though one needed a sense that they deserved better, similar to Runciman's (1966) emphasis on one's sense of justice influencing RD.

By 1970 RD has been applied towards a variety of topics such as urban violence (Gurr, 1968) and career choices (Davis, 1966). However, this broad scope of application may be due to RD being a concept with profound effects or because RD is poorly defined and therefore malleable as an explanatory device. Cook, Crosby and Hennigan (1977) argued for the validation of RD as a construct and outlined efforts to create construct validation. They reviewed previous research to determine if RD is capable of validation. They reviewed previous definitions, such as those described above, even going so far as to contend that "(there is) no single definition common to the major theorists" (p. 308). In reviewing previous research, six components of RD were identified: not having X, wanting X, social comparison, feasibility of having X, entitlement to have X, and no personal responsibility for not having X.

These are similar to Runciman's (1966) conditions of RD, but are expanded to include concepts of feasibility and an absence of personal responsibility for not having "X". These conditions provided by Cook et al. (1977) were not agreed upon in the literature, rather included an "all of the above" collection of the conceptions of RD from research by prominent RD researchers. For example, a lack of personal responsibility for not having X was indicated by Crosby (1976) as a necessary component, though none of the other prominent RD researchers of the time included such a condition.

Because of the variations of definition, RD could be considered either a hypothetical construct or an intervening variable. The former can be directly measured

while the latter is a label to describe a set of processes. For example, Cook et al. (1977) used *thirst* as a hypothetical construct and *cognitive dissonance* as an intervening variable. RD has been described as both, even by the same researcher and in the same publication. Davis (1959) referred to RD as a “subjective feeling” and a “belief that there is differential treatment” (p. 23). Here, the “subjective feeling” description implies it cannot be directly measured (an intervening variable) while the consideration of RD as a belief of inequity could be considered a hypothetical construct. Cook et al. (1977) argued that there is also difficulty in identifying if RD is a cognitive or emotional process, though many researchers used both. For example, Runciman (1966) considered the magnitude or perceived size of RD to be cognitive while the degree as emotional. Gurr (1970) suggested that magnitude and intensity co-vary (Cook et al., 1977).

The differing definitions and conceptions likely contributed to another problem with RD theory- lack of specificity of measurement. This impeded the theory’s development. Gurr (1970) considered RD as measurable from objective economic conditions such as inflation rates and average salaries. Runciman (1966) used yes/no questions such as “Do you think there are other sorts of people doing noticeably better at the moment than you and your family?”(p. 302). Cook et al. (1977) regarded both efforts as measures of magnitude (which they considered cognitive), but not intensity (considered emotional). There is no measure of the tension that develops from a disparity. Also, there were no efforts to identify if the comparisons mattered to the individual (which could be considered similar to a measure of intensity, as the first would likely dictate the second), no measure of expectancy, or any concern that such

questions had primed a response by making such differences more salient simply by asking.

Cook et al. (1977) explored previous RD research in an effort to identify if the construct was an entity separate from other psychological constructs (such as equity and frustration). They restricted their review to egoistic RD (Davis, 1966; Runciman, 1966; Gurr, 1970). They stated that since most authors assumed RD is directly measurable, then RD is considered in the research as a hypothetical construct, though not specifically stated. They found that researchers often do not distinguish between cognitive magnitude measures and emotional intensity measures. Cook et al advocated for RD to be considered as an intervening variable. The benefits of treating it as an intervening variable include: no longer requiring a presumption of tension, not requiring a specific statements of the independent and dependent variables involved and statements of conditions, and serves as a label of convenience. To justify RD as an intervening variable, the authors revised the concept of RD to the following conditions: one does not have or have enough of X, one wants X, one compares with similar others who have X or more of X, and one feels it is feasibly to have X. These conditions are not independent and can happen in any order. For example, feasibility may influence social comparison. One may consider it feasible to become famous and therefore use famous people for comparisons. Furthermore, the two previously described conditions of entitlement and lack of personal responsibility for not having “X” are moderators of how RD will be expressed (e.g., anger, achievement).

Furthermore, Cook et al. (1977) argued for treating RD as a hypothetical concept partially because of the difficulty in previous RD research to verify an emotional pre-

condition: “This position has the advantage of not requiring direct measures of an elusive and presumptive tension state” (p. 310). No major support is given for this declaration other than what seems to be a desire for validation of the theory. Wheeler and Zuckerman (1977) criticized this distinction of RD from equity theory, stating: “(I)t seems questionable whether feasibility is missing from the equity model” (p. 354). They contend that RD is a state experienced by those that are a “victim” of inequity while relative gratification is experienced by one who comes out ahead in perceived inequity.

Finally, shifts in measurement approach may also contribute to measurement issues. Lerner (2003) criticized the measurement of all justice theories (equity, just world, and relative deprivation). Specifically, Lerner noted the increasingly common practice of measuring one’s sense of justice from role-play scenarios. He argued that such hypothetical situations may not hold interest to the participant, thus creating what he called “low impact” scenarios where individuals may respond in socially normative ways instead of justice-motivated interests.

It seems that the work of Gurr (1966) and Runciman (1966) were prototypes for what would become two primary approaches to measuring RD: assessing deprivation through either objective standards or perceptive ratings of people’s attitudes of “doing better”. Yitzhaki (1979) argued that RD for an individual could be calculated through the computation of the weighted sum of the difference between an individual’s income and the income of people in the same reference group with greater incomes (an objective standard). Conversely, Crosby (1982) measured RD in a more perceptive approach by asking how “bitter or resentful” (p. 80) participants felt regarding seven aspects of women’s employment situations on a six-point Likert scale.

The review of RD by Cook et al. (1977) pertained more to the individual application of RD. This form of RD was in the vein of Davis (1966) work on perceptions of career choice. Gurr's (1966) application of RD towards social movement and civil strife was different direction for the theory, as it accounted for group behavior rather than individual differences. This is similar to Runciman's (1966) notion of fraternalist deprivation. The application of RD to explaining the causes of civil strife and social movements was very popular in the 1960's and 70's. Davies (1962) believed RD was a necessary but not sufficient condition for social movement, Vanneman and Pettigrew (1972) assessed race relations in the United States and considered RD as existing in Black communities because of the relative advantage of Whites, and Morrison (1973) believed RD acted as a catalyst for social movements because RD was cognitive dissonance- people were upset that their expectations were not fulfilled.

The ability of RD to account for social movement, however, was criticized by Gurney and Tierney (1982), primarily because of inconsistencies in the conceptual, theoretical, and empirical design of the construct. They argued that types and patterns of RD are often offered, but a relationship between objective and perception conditions is never articulated. For example, Gurr (1970) offered three patterns of RD: detrimental, aspirational, and progressive while also suggesting there were four sources of one's expectations: previous experience, abstract ideals, standards of a leader, and one's choice in reference groups. Gurney and Tierney (1982) also argued that such distinctions further complicate the theory of RD because they obfuscate a fundamental issue of RD- the relationship between objective and perceptive conditions. For example, Gurr's (1970) three patterns of RD may describe objective conditions that all result in the same

RD (an objective condition) or three unique forms of RD (a perceptive condition). The failure to distinguish between objective or perceptive conditions has allowed for what Gurney and Tierney (1982) considered a misstep in RD theory- the use of macroeconomic indices to assess RD. A failure of the literature to clearly outline how objective conditions relate to perceived conditions resulted in multiple definitions and varying measurement approaches. This lack of researcher agreement in the theory and application created results that could not be compared to one another. This was stifling to the development of the theory. Because of the lack of researcher consensus, RD never became a coherent theory. Unfortunately, this criticism continues to plague RD research. There is still no agreed-upon mode of measurement of RD and researchers continue to attempt alternate methods (Webber, 2007; Pettigrew, Christ, Wagner, Meertens, van Dick, & Zick, 2008).

Another deficit in subsequent exploration of the consequence of perceived inequity is the lack of research exploring RG. RG was recognized early in the RD literature (Stouffer et al., 1949) but not explored. Runciman (1966) recognized the concept with: “(T)he proverbial way to make oneself conscious of one’s advantages is to contrast one’s situation with that of others worse off than oneself.” (p. 9). RG was considered a therapeutic technique employed when one was experiencing RD. Earnest research on the effect of RG is relatively recent (Guimond & Dambrun, 2002).

Recent researchers have attempted to investigate the effects of perceived benefit rather than deprivation. Two phenomena have both been labeled as this condition. Yitzhaki (1979) introduced the first label, Relative Satisfaction (RS), defining it as the mathematical inverse of the YRD. Eibner and Evans (2005) referred to Relative

Satisfaction (RS) as “being prosperous relative to one’s peers” (p. 10). They did not directly measure RS but assumed that an effect found from an increase in RD was equivalent to a decrease in RS.

Other researchers have referred to this condition as relative gratification (RG). For example, Guimond & Dambrun (2002) attempted to create RD by having participants read a fictitious report stating that their job market was dwindling. The authors attempted to create RG by informing a second group that the job market would improve considerably. Interestingly, RD and RG were measured via just one question that asked participants to rate on a 7-point scale how satisfied they were when thinking about the future, comparing their current situation with what they were going to do in a couple of years.

The most recent and comprehensive attempt to measure RD and RG built upon the findings of Guimond & Dambrun (2002) by assessing RD and RG among a representative sample of South Africans (Dambrun, Taylor, McDonald, Crush, and Mèot, 2006). The researchers tested the hypothesis that RD and RG both predict prejudice in a bilinear model of predicting prejudice. Participants rated on a 5-item Likert scale their degree of satisfaction with the comparison such that higher scores were indicative of RD while low scores indicated RG. Their instrument included two components of RD and RG: economic conditions and general conditions. Both components appear to be further broken down into questions that assessed egoist and fraternalist comparisons. The economic RD/RG measures assess the individual (egoist) or the individual’s ingroup (fraternalist) current economic conditions and future outlook of economic condition. For example, the egoist economic current condition was assessed

by: “At the moment are you (satisfied/dissatisfied) with your personal economic conditions?” (p. 1035).

The general RD/RG measure also had two questions for each the egoist and fraternalist category, but instead of using the comparisons of current rating and future outlook, “outgroups” were chosen for each. The general egoist condition questions assessed personal conditions compared to other South Africans and other members of the participants’ ingroup while the fraternal questions assessed the conditions of the participants’ ingroup compared to other groups in South Africa and South Africa compared to other South African countries. For example, the fraternalist general question that asked the participant to compare their ingroup to other groups in South Africa was, “Would you say that the overall conditions of people from your ingroups are (better/same/worse) than those of other groups in South Africa?” (p. 1035). This is our understanding of the survey framework used by Dambrum et al. (2006), not one explicitly stated in the published research paper.

By including an RG dimension, the researchers were able to detect an effect not anticipated in the literature, that RG predicts prejudice as well as RD. Should the researchers have chosen to measure only RD, they would have simply found that RD did indeed influence prejudice. Therefore, the inverse of RD should not be assumed to have the inverse of the effects of RD.

Unlike SCT, which can be considered to have expanded over the last sixty years, RD has had a less organized development due to a series of unresolved contentions in the literature. One area of contention has been how RD is distinct from other social theories. Davis (1959) stated that wanting, not-having, and social comparison were

necessary for RD to occur. These conditions, as Wheeler and Zuckerman (1977) argued, are the same as those of inequity described by Cook et al. (1977). This author does not find this problematic. Rather, I consider perceived inequity as a possible consequence of social comparisons, and if distress is felt then this condition is RD (and if satisfaction is derived then RG is felt). Along the course of this review of RD I have interjected the SCT parallels and correlates to suggest that RD is inequity perceived from social comparisons.

A second area of contention is the ambiguity of whether RD is an intervening variable or a hypothetical construct. I argue that RD is a hypothetical construct because distress derived from perceived inequity could be directly measurable through participant ratings or other methods. Furthermore, the support for RD as an intervening variable seems to be the difficulty in measuring the construct (Wheeler & Zuckerman, 1977). This is not a valid reason for RD to be considered an intervening variable. A third contention is the definition and nature of RD (Crosby, 1979).

Of the multiple definitions of RD, I consider Runciman's (1966) interpretation of RD being dependent upon social comparisons and Gurr's (1966) explanation of the relationship of RD in which the magnitude and intensity of RD co-vary. However, I add the caveat that *perceived* magnitude and intensity co-vary. Perceived deprivation is the condition where RD may occur, therefore they co-vary in that one may exaggerate the disparity because of the felt RD (intensity) or the intensity may be greater because one feels the difference, though perhaps objectively small, is considered large (magnitude). Gurney and Tierney (1982) criticized RD researchers for often failing to clearly outline how objective and perceptive conditions interact. In my interpretation and application of

RD, the phenomenon is perceptible. RD is a consequence of social and temporal comparisons.

What one has had in the past and abilities observed in others dictates the RD. One assesses their standing in these comparisons (though they may not be accurate) and any distress derived from these comparisons is RD (while any satisfaction derived would be RG). Furthermore, RD is directly regulated by SCT. As SCT explored self-evaluation and self-assessment, one finding was that there were defenses people engaged in to avoid seeing themselves unfavorably (Brickman & Bulman, 1977). RD is the “leftover” negative self-evaluation derived from comparisons not effectively rationalized. Objective conditions are likely indirectly relate to RD in that comparison are likely selected based on objective conditions, but the two are independent and may move in the same or different direction.

This is not a novel interpretation of the phenomenon. Rather, this interpretation of RD was likely the original understanding. As described by Merton and Kitt (1950): “Relative deprivation can provisionally be regarded as a special concept in reference group theory” (p. 52). RD was originally considered a *social* theory- a unique condition from reference groups. Reference group theory is similar to SCT in that comparisons are made to others as a frame of reference. Therefore, RD is a specific conclusion of social comparisons, and under SCT one would expect RD to emerge from dissatisfaction from comparisons and efforts to discount this dissatisfaction (such as shifting the importance of the comparison or the perceived degree of similarity with the comparison) are not entirely effective. This interpretation of RD was difficult to measure in social movement research, which required large data sets and analysis often with historical data.

Adjustments were made to the theory that allowed objective measures to be considered as measures of RD, a technique that contributed towards the muddling of the theory (Gurney & Tierney, 1982).

Similar to Wood's (1996) core definition of social comparison, a similar definition of RD is that it is distress derived from comparisons. The degree of this distress will vary by the degree of closeness or similarity with the comparison group and relevance of the object of comparison. The distress may lead to either stress or social upheaval; depending on if the RD is egoistical or fraternal, which likely dictated by whether the comparisons are SEM or SCT. SEM creates closeness and would be fraternal while SCT predicts similarity-based comparisons and would therefore be egoistical. The fraternal predictions of RD to motivate collective response have been difficult to verify, leading to the conclusion that RD is not useful (Gurney & Tierney, 1982).

While the difficulty of fraternalist RD to predict collective action does warrant criticism of the theory, RD is a viable construct. RD is a hypothetical construct that is a link in a chain reaction stemming from social comparisons. Social comparisons may create emotional reactions. One's emotional response to social comparison may create distress (RD) or satisfaction (RG). From here the chain of effects can move in several different directions as RD and RG can influence one's satisfaction with their current environment, desire for social movement, and career aspirations. One area to test this model would be status-related health disparities. The relationship between status and health could be mediated through the stressfulness of undesirable conclusions following

social comparisons. Status is likely evaluated through social comparison and as such SCT can predict status-health effects through the measure of RD and RG.

### Stress and Health

In the introduction I suggested that social information affects health and I provided examples such as the lifespan of academy award winners and British civil servants. Thus far, this text has reviewed how expectations can emerge from social information and a proposed model is offered by which SCT can create emotional states such as RD and RG. A person may develop a sense of unhappiness derived from social information (relative deprivation) and/or may feel satisfied (relative gratification). These “consequences of social comparisons” may influence behaviors and feelings (such as career aspirations, happiness with current status, and desire to revolt). These comparisons may also negatively affect health due to the distress caused by social comparisons. These states would increase in unfavorable as status decreased, accounting for status-related health effects. Even gratification can be stressing (such as when one feels guilty) or can be beneficial (such as with coping). While the notion that perceptible states affect health is generally accepted, reliably demonstrating this effect has been difficult (Marmot, 2005). This has been due, in part, to difficulties in defining stress and health. The purpose of this section is to outline the mechanisms of stress and health. The review of stress and stress research is presented in the following order:

- The biology of stress
- The concept of stress
- The measurement of stress
- Stress and health

## *The Biology of Stress*

The physiological response to a threat, known as the *stress response*, is adaptive for survival in times of crisis (Clow, 2001). The types and frequency of modern *stressors*- variables causing stress- are often different than the stressors our ancestors experienced. Biologically, the stress response is designed for infrequent life threatening situations. Today's humans face frequent but rarely life threatening events, though the body cannot discriminate between types of stressors and therefore responds to all stressors as if they were life threatening. Frequent activation of the stress response results in dysregulation and can damage the body, often more so than the stressor (especially since most modern stressors are not life threatening). The stress response falls under the domain of the autonomic nervous system and is comprised of the sympathetic and parasympathetic nervous systems (SNS and PNS, respectively).

Stress activates the SNS while relaxing activates the PNS. The process starts when the brain identifies a stimulus as a stressor (Clow, 2001). Then, an emotional response to the stressor is regulated in the limbic system. The limbic system functions primarily for survival behaviors- sexual reproduction, fear and aggression. In particular, the hippocampus and amygdala function to regulate the stress response in this stage. The limbic system can then activate the hypothalamus, which can control the two stress response systems: sympathetic adrenal medulla (SAM) and hypothalamic-pituitary-adrenal (HPA). Together these two systems regulate the cardiovascular and immune system. Either one can be activated by psychological stressors.

SAM is a general stress response that is activated by arousal, pleasure, and strain (Clow, 2001). When activated by a perceived threat, the sympathetic nervous system

will activate the release of nonadrenaline to activate internal organs and adrenaline into the bloodstream, released by the SAM system. The SAM system also changes cardiovascular systems. For example, the heart beats faster (upwards of five times the resting state amount) and the blood vessels restrict to increase blood flow. Blood is also routed away from arteries that supply blood to the digestive system, the skin, and the kidneys. Vasopressin, a hormone that inhibits urine formation is released to further inhibit processes that are ancillary during a life and death situation. The continued and repeated activation of this system, however, can cause damage. For example, the continued increase in blood pressure can cause damage to the blood vessel lining, allowing fatty deposits to accumulate in micro-tears. This plaque lining is referred to as atherosclerosis and can cause a heart attack. This example illustrates how frequent stress can lead to health problems.

The HPA is a system only occurs as a stress response and is activated in extreme situations, though what is considered “extreme” varies across individuals (Clow, 2001). If during the SAM response the stressor is deemed to be a threat, a chemical messenger, corticotrophin releasing factor (CRF), is released from the hypothalamus. When CRF reaches the pituitary gland, adrenocorticotrophic hormone (ACTH) is then released into the circulating blood system and eventually reaches the adrenal glands. Here, the outer cortex of the adrenal glands responds to ACTH by releasing glucocorticoid cortisol. Cortisol and ACTH circulate in periphery blood normally, regardless of exposure to stress. In stress encounters, however, the cortisol level becomes elevated and works to release energy reserves. This, when coupled with increased blood flow from the SAM system, results in greatly increased energy expenditure in stressful situations. Glycogen,

converted starch, sugars, and carbohydrates are released from the muscle and liver and are broken down into glucose in the blood. This is intended to give the body energy during the crisis period. However, too much glucose creates the fatty substance that adheres to the tiny tears that occur in the development of atherosclerosis and underscores the potentially unhealthy consequence of over activation of the stress response. Another mode by which chronic stress can affect health is through abundance of cortisol, a steroid that can pass to every part of the body and affect the brain by reducing tryptophan rates. Tryptophan is a precursor of serotonin (5Ht), a neurotransmitter associated with mood. Cortisol in the brain can therefore contribute to depressed mood.

Stressors can also activate the immune system and also sometimes create health problems. The immune system has two strategies to ward off infection (Clow, 2001). The first is humoral immunity, which is non-specific (i.e. it is a non-specialized general response to threat). In humoral immunity, antibodies are secreted. Antibodies are proteins that bind to and deactivate invading pathogens. This non-specific immune system works for pathogens that are outside a body's cells. Once a pathogen has invaded cells, the second system is valuable- cell-mediated immunity. This system actively destroys invading pathogens and is used primarily for viruses. The immune system is in charge of protecting the body from infectious agents and can fail in two ways. First, it can fail due to *under vigilance*, whereby agents gain entry and infection occurs. Anytime one becomes sick, this would be a case of under vigilance. Second, *over vigilance* can occur in which the immune system itself causes illness. Asthma is an example of over vigilance.

The cells of the immune system are sensitive to the stress hormone cortisol (Clow, 2001). Cortisol down-regulates cell-mediated immunity but can enhance humoral immunity. Acute stress may up-regulate cell mediated immunity as a spike in cortisol seems to initiate this response. Continued cortisol secretion, however, shifts the focus of the body away from cell-mediated immunity to humoral immunity. A chronically stressed person may have an attenuated cell-mediated immune system (under vigilance) while an over activity of humoral cells (over vigilance).

Stress is not the only variable that influences cortisol levels. Both cortisol and immune cells demonstrate circadian changes and individual differences (Clow, 2001). Cortisol is always present at some level in the blood, as the hormone is involved in other processes such as controlling the biological clock. During situations without marked stressor-related cortisol activity, the hormone is lowest during the night. The act of waking up causes a shock that produces a three-fold increase within the first 30 minutes. Cortisol levels then gradually decline as the day progresses. The stress response of the HPA takes 20-30 minutes and as such, cortisol spikes 30 minutes after a stressor and then slowly falls until returning to baseline. The down regulation of cortisol is conducted by cortisol detection cells in the pituitary and hypothalamus. These receptors know the appropriate levels (baseline) of cortisol, which fluctuate across the day. When levels of cortisol are high for that time of day, down-regulation is initiated.

The fluctuations in humoral and cell mediated immunities are referred to as *counter regulatory* (Clow, 2001). In a 24-hour period these cycles will alternate, with humoral immunity being most active in the day and cell-mediated immunity occurring largely at night. Individuals may be predisposed to one strategy over another and stress

can influence both. If stress suppresses one system (down-regulated), the other system then increases in activity (increasing the likelihood the system is in over-regulation).

This description is a broad review, not intended to serve as a formal representation of the sequela of processes that are influenced by the stress response. There are many other processes and components of immunity that can affect health. For example, sIGA is an antibody of the mucosa, which provides primary defense and protection of the mucous lining (Clow, 2001). Acute stress raises sIGA while chronic stress lowers it. Not every immune process that can be influenced by stress is outlined here. The purpose of this summary of immune functioning is to provide an explanatory mechanism for the relationship between stress and health: both share at least one common hormone, cortisol. This review illustrates the multidimensional influences of stress and health. Stress does not simply equate to impaired health through attenuated systems. Rather, chronic stress can create dysregulation, interfering with biological homeostasis. Generally, cortisol is a product of stress and a regulator of immunity and this dual role may not be coincidence. The immune system requires significant amounts of energy to produce and maintain- resources that become more precious in an emergency situation.

Cannon (1932) described the physiological responses of the sympathetic nervous system as a *fight-or-flight response*. The term now refers to the supposed purpose of the sympathetic nervous system in response to threat- to create an increase in short-term energy intended to help the organism escape or combat the threat (Taylor, Klein, Lewis, Guenwald, Gurun, and Updegraff, 2000). The term, however, may be inappropriate due to the implication that the response to a threat creates only two possible reactions and is

equally possible between the sexes. Women may respond differently to threats than men. The popularity of the fight-or-flight conception of stress may have been partly due to the utilization of male participants in the majority of stress research. The disproportionate use of males, as opposed to females in such research has been attributed to the high percentage of males in colleges during the early part of the 20<sup>th</sup> century as well as females having been excluded from studies due to their greater cyclical variation in neuroendocrine responses. The existence of gender differences in response to threats was posited by the field of evolutionary psychology, which is a discipline of psychology that seeks to understand the evolutionary value of psychological processes.

Taylor et al. (2000) cited evolutionary psychology as the conception of research that facilitated the consideration of revising the flight-or-flight response to stress after considering multiple endocrine and behavioral studies in affiliation. A central tenet of the theory is the importance of producing offspring and the development of behavioral mechanisms that facilitate successful reproduction. Due to a greater parental investment, (ie., the large amount of energy required for producing offspring) fighting or fleeing from a threat does not appear to be as adaptive for women as men. For example, in the likely environment in which humans evolved (the African plains), an attack by a predator would be a threat. Pregnant women (or a new mother with a child) would not be adept at fending off a predator by attacking or running away. Because women are the biological caregivers for infants, a fight response would put them and their offspring at risk while a flight response may not be feasible for women that are pregnant or caring for young children. Instead, strategies may have evolved among women to affiliate with other women in times of stress. Therefore, behaviors may have evolved that involved

removing offspring from a threat and creating support systems that preemptively provided protection in case a threat arose. Such a system would likely involve cooperation with other women, as cooperation with males creates a risk of rape and other violence. Taylor and colleagues (2000) referred to this response as *tend-and-befriend*. This system would likely have emerged from the attachment and care giving that mothers often feel with their offspring. This is not to suggest that fight-or-flight does not occur in women, only that the response may be more descriptive of the male stress response (nor is this to suggest tend-and-befriend only applies to females, as discussed later).

Taylor et al. (2000) identified hormonal differences between men and women that supported the notion that the fight-or-flight response is less descriptive of the female stress response than the male response to threat. Hormonally, men have greater concentrations of testosterone while women have greater amounts of oxytocin. Testosterone increases with acute stress and is associated with higher levels of hostility (Girdler, Jammer, and Shapiro, 1997). Additionally, testosterone is an androgen associated with aggression and rough-and-tumble play is largely absent in women. Female hostility is not associated with the arousal of the sympathetic nervous system. This, Taylor et al. (2000) argued, is supportive of the notion that while women are also aggressive, it is not tied to testosterone and the arousal of the sympathetic nervous system. While aggression levels may not manifest differences across gender, the conditions that create aggression may differ. Men may feel aggression in response to stress due to the higher amounts of testosterone. Therefore, the stress response of

aggression would not be as common in women. This reasoning challenges the likelihood of a “fight” stress response among females.

The presence of increased levels of oxytocin in women would reduce the likelihood of a “flight” response as well as provide a mechanism for tend-and-befriending as a stress response. Oxytocin is a hormone released by the posterior pituitary as part of the PNS. The hormone is found to have a sedative and anxiety-reducing property found in virtually all mammals. Because of this property, oxytocin is utilized by the body in down-regulation of the stress response. While oxytocin is in females and males, Taylor et al. (2000) identified three reasons that there may be sex differences regarding the role of oxytocin in the stress response. First, females have more oxytocin released in a stress response than males. Second, androgens inhibit oxytocin and therefore the higher levels of testosterone found in males during the stress response would inhibit the oxytocin released. Third, estrogen influences the effect of oxytocin and would be more influential in women as they have higher levels of estrogen than males.

Affiliation behavior also seems related to oxytocin (Taylor et al., 2000). Oxytocin may promote maternal bonding with an infant based upon data showing that pregnant and breastfeeding women were found to be calmer and more social than matched-age women, and this mood was correlated strongly with oxytocin levels. The calming effect of oxytocin may have initially evolved to promote bonding between mother and offspring, but it also may increase females’ tendencies to create affiliation. This has the evolutionary advantage of additional protection from predators. Female affiliation has been seen among many primates, and is beneficial for allowing young

females to gain experience in tending infants. The affiliation behavior can take place without the hormonal influence as well (Keverne, Nevison, & Martel, 1999). Oxytocin can prime affiliation behavior among females but affiliation may be maintained among primates because of higher order brain functioning that allows the primates to recognize the benefits of affiliation.

Taylor et al. (2000) contended that oxytocin was related to relaxed mood, maternal behavior, and social affiliation. The benefits of social affiliation can be recognized by organisms with higher cognitive abilities. Sex differences then exist in the stress response, as females would be more likely to engage in social behavior when stressed because of lower testosterone and greater amounts of oxytocin than males. Taylor and colleagues argued there is ample evidence supporting this, such as in rodent studies in which crowding is found to stress male rodents but calm female rodents (Brown & Grunberg, 1995).

The affiliation choice is not random, as Schacter (1959) found that women preferred to affiliate with other women who are perceived as being similar to themselves. Women engage in more social networks, which Taylor et al. (2000) view as being a part of the stress response. Women also engage in more close relationships with same-sex peers and are more likely to refer to them for support (Belle, 1987). Recall that Wheeler & Miyake (1992) found that women were statistically significantly more likely to compare themselves with a family member and men were more likely to compare themselves with a famous person, suggestive of the greater importance of affiliation among females. This does not mean that women are less stressed. Taylor et al. (2000) reviewed research that found female affiliation could be stressful. For example, women

rank interpersonal stressors as the most frequent stressor and causing the most stress (Davis, Mathews, & Twamley, 1999). While social affiliation may be a stress response found in females, this response may itself cause stress.

The differences in hormone levels between men and women appears to prime them to adopt different strategies in response to stress (Taylor et al., 2000). This conclusion is a generalization, as women will engage in fight-or-flight and men can engage in tend-and-befriend, but the point is that there is a tendency towards sex differences in response to stress. The greater down-regulation of the stress response and the enhanced social support generally found in females over males is a possible explanation for the longer life span of women because of a reduced vulnerability to stress-caused health deficits. Women may then be less likely to engage in stress-coping risk behavior (such as smoking) and less vulnerable to conditions associated with chronic stress (such as cardiovascular problems).

Stress is a deemed threat that creates a cascade of hormonal changes in the body. Sex differences found in the stress response (Taylor et. al, 2000) imply biological variation. Once a threat has been identified, men and women may respond to a threat differently due to differences in their amounts of testosterone and oxytocin. Differences occur not only between biologies but also between philosophies. The HPA stress response is regulated by the hypothalamus, which is first activated when the amygdala detects a threat. The amygdala may be influenced by psychological processes. People with high psychological resources (perceived personal control and self-esteem) experience greater de-regulation of amygdala activity when presented with a threat (Taylor, Burklund, Eisenberger, Lehman, Hilmert & Lieberman, 2008). Corresponding

lower cortisol activity was also found with lower amygdala activity. The lower amygdala reactivity in those with higher psychological resources likely results in decreased cortisol activity. Psychological processes are affected by social comparisons. For example, self-esteem has been argued in the previous sections as being influenced by social comparisons. One's perception of stress influences one's biological response to stress, and social comparisons play a role in perceptions.

### *The Concept of Stress*

The previous review of the stress-response underscores our considerable progress in understanding the biological relationship between stress and the body. Identifying a strong relationship between stress and health, however, has been surprisingly difficult. This is due in part to the difficulty in measuring the concepts of stress and health as well as how the two interact. Social comparison seems to be involved, even in the biological process of stress, as the tend-and-befriend stress response apparently causes an organism to seek out "similar others". Stress, like social comparison and RD, is an indirect variable that cannot be directly measured. As such, there is controversy surrounding the definition and measurement of stress (Jones & Bright, 2001). There is also disagreement about whether or not the construct of stress even exists. The history of the construct of stress and the relationship between stress and health will be briefly described to underscore current issues in the measurement of these constructs.

Seyle (1956) formally proposed the current conception of *stress*. His notion of stress only referred to the physiological response to challenge. Despite a wide range of types of stress, Seyle contended that the body has one non-specific response- the general

adaption syndrome (GAS). The GAS consisted of the three stages of alarm, resistance, and exhaustion. Alarm is the initial indication of the presence of a stressor (a stressor is anything inducing stress, a rather circular definition). Resistance is the body's maintenance of the physical changes. Exhaustion occurs when the exertion of the body's physiological response can no longer be maintained. These physiological changes include increased heart rate and a redirection of biological processes (discussed previously). This conceptualization of stress has been criticized, most notably for being overly simplistic (Hinkle, 1973) and for being circular in defining stressors as anything that causes stress (Lazarus & Folkman, 1984). Seyle (1976) admitted that at the time his GAS model was first put forward that wasn't very proficient in English and that he would have preferred to use the word *strain* to refer to responses to stress (Jones & Bright, 2001).

The stress research community has not agreed upon a definition of stress but subsequent definitions incorporated the influence of perception into the stress response.” (p. 20). Lazarus & Folkman (1984) defined stress as a “particular relationship between the person and the environment that is appraised by the person as taxing”. McGrath (1976) referred to stress as “A (perceived) substantial imbalance between demand and response capability under conditions where failure to meet demand has important (perceived) consequences his or her resources and endangering his or her well being” (p. 19). Other researchers however have recommended abandoning the use of the construct of stress entirely (Hinkle, 1973; Pollack, 1988) while others contend there is no disagreement about the definitions (Cox, 1993). Jones & Bright (2001) concur with Lazarus & Folkman (1984) that the term “stress” should be used to describe a process of

the relationship between person and environment, accommodating the role of perception and stress.

As stress research grew in popularity, the way in which stress is perceived has changed (Jones & Bright, 2001). The current conception of stress is dependent on perception and therefore changes *as a result* of study. Stress has shifted in mainstream culture from being considered an unavoidable facet of life to a man-made creation, capable of being avoided (but they contend that stress in some form or fashion probably cannot be completely avoided). Jones and Bright primarily blame social scientists for propagating this incorrect conception and suggest a rather entertaining posit that people may be more stressed today because they feel we can avoid it, but cannot avoid it because stress is unavoidable. The conceptualization of stress seems to be as important to experiencing stress as it is to understanding the concept.

Three approaches to conceptualizing stress are currently used- traditional, interactional, and transactional (Jones & Bright, 2001). The traditional approach assesses stimulus-response situations. For example, a researcher may compare the health of one country to another, conjecturing that one country's inhabitants may have more stressors than the other country and therefore greater health problems may occur in that country. This approach is useful for identifying overall trends but ignores individual differences in the conceptualization of stress.

The interactional approach recognizes three conditions in the stress response: environment, intervening variables, and strain outcomes (Jones & Bright, 2001). Environmental events are the stressors (such as workload). Intervening variables are individual differences (such as personality type). Strains are the effects experienced

(such as anxiety). The intervening variables are a set of conditions that will either increase or decrease the strain of the stressor. The interactional approach considers individual differences, unlike the traditional approach, but not the interaction between the environment and the person's response.

The transactional approach shifts the emphasis of stress research to the interaction of variables (such as the environment and personality) on the appraisal processes of the individual in determining if a situation is stressful (Lazarus, DeLongis, Folkman, & Gruen, 1985). Appraisals are combined from *personal agendas* and *objective inferences*. Personal agendas include such things as one's beliefs about the self and the world. When a transaction is deemed as stressful, coping processes are then employed to manage the relationship between the environment and person (this determines the type and intensity of the stress reaction). For example, previous experience with a given stressor will influence future appraisals. Under a transactional approach, a stress relationship is constantly changing because appraisal and coping are tied to environmental conditions and previous experience. What taxes a person's resources once may be insignificant the second time around and the event may have never bothered another individual because of personality, environment, and/or biology.

The relationship by which social comparison affects health is considered here as being transactional. Social comparison is a process whereby one evaluates one's abilities in relation to others (and therefore evaluates one's social standing). Furthermore, social comparison is likely employed in appraisal and coping. Appraisal processes include a personal agenda, which would utilize social comparisons in the creation of beliefs about the self and the world. Social comparisons would also be used to gauge if one has the

resources to contend with a situation by comparing oneself to others with similar resources. Coping process may then be employed to blunt any distress created, and may also use social comparisons, such as adjusting the relevance of comparison groups and/or dimensions. When a person is evaluating whether they have the resources to meet a demand, comparisons to similar others would be made. As an example of social comparison in appraisal and coping, a person about to take their first driver's test might think of peers who have recently taken the test and recall what they said about the test to determine if the experience might be stressful (appraisal). If the act of taking a driving test was evaluated to be stressful, then social comparisons may be used to manage the stress, such as thinking of less competent peers that passed the test (coping).

Frustration may emerge as a consequence of these comparisons when people believe that they should have something that they currently do not have (but others do have). Social comparison may also be used as a source of comfort in the face of stress (coping). Some stressors arise from social comparisons and some are managed by social comparisons. This raises the issue of just how many of our stressors are created by social comparisons. In a modern society in which lethal physical threats are rare, just about all stressors may be derived from social comparisons. In modern societies, stress may be perceived during public speaking, difficult tests, job interviews, and blind dates because people are concerned that their wishes or expectations will not be fulfilled. Social comparison may play a large role in the stress response because most stressors are derived from expectations, which are themselves derived from comparisons to others.

### *The Measurement of Stress*

Since there are multiple interpretations of stress, there are multiple methods of measurement. The interactional and transactional theories differ in their measurement approaches. Interactional approaches are supposed to independently measure environmental and intervening variables (Kasl, 1978), while transactional approaches recommend measuring environment and intervening variables together (Lazarus, 1990). Transactional theorists take issue with separating these variables from environmental conditions. Under a transactional approach, stress is a process and not a condition, yet stress measures are typically *output measures*- measures of a stress condition. As evidence of such output measurements, Cohen, Kamarck, & Mermelstein, (1983) created the perceived stress scale (PSS) which uses 14 statements that describe varying stress conditions and is scored along a 4-point scale. The final score is a measure of how stressed the individual has been in the last month. Lazarus (1990) took issue with the PSS and most output scales because they ignore the context of the stress and the resulting unidimensional stress score does not illustrate the qualitative variations in stress response (for example, is one threatened or challenged by the stress felt?).

Lazarus (1990) instead recommended two approaches towards a transactional measurement of stress. The first is a continuing approach in which *points of transition* are measured. Points of transition are the times when the relationship between person and environment has changed. These points would be difficult to establish but should be theory based. Perhaps realizing that this approach would be difficult for most researchers to implement, Lazarus' second approach uses a single score that is assumed to be a summary of the stress response. This is an aggregate appraisal of the stress

process by the individual. Unlike output measures, which are presumed by the researcher in terms of quality and quantity of stress, the individual rates their own stress. This approach marginalizes the role of fluctuations and memory errors and risks distorting the relationship of the person and environment with regards to the stress response. However, such an approach does avoid the pitfalls of traditional output measures because the qualitative condition of the stress is measured.

In either interactional or transactional approach, measurement issues emerge (Jones & Bright, 2001). For example, both the interactional and transactional approaches may utilize self-reports. Intervening variables such as personality are also typically measured through self-reports. Self-reports have reliability and validity issues because recall and social facilitation may bias responses to the items. The use of self-report measures of stress is problematic particularly because stress may not be a conscious experience (Jones & Kinman, 1990). Breznitz (1990) criticized Lazarus' (1990) approach to measuring stress, stating that his measures rely only on stress components that are accessible to awareness. By doing so, these measures inadvertently focus on sociological content (such as work, family, and health) and shift the focus away from the psychological factors (such as the need to be liked, anxiety, and loss of perceived control). In this regard, objective measures would then be superior, such as a report of major life events. Such a scale is objective because though self-report, it does not have a perceptive stress rating. The ratings of the stressfulness of such events are pre-set and consistent across participants. However, such a measure is not necessarily a comprehensive measure of stress because while major life events are stressful, so are daily hassles (Jones & Bright, 2001).

Additionally, objective measurements discount the perceptive component (Jones & Bright, 2001). Perceptive measures of stress have been found to be related to health. The PSS, despite being criticized by Lazarus (1990) for ignoring the context of the stress and the emphasis on a unidimensional stress score, has been found to be psychometrically fit and a superior measure of health than a life events scale (Cohen, Kamarck, & Mermelstein, 1983) and has been shown to be significantly correlated to anxiety and depression (Spada, Nikčević, Moneta, & Wells, 2008). As will be described next, the relationship between stress and health is far from understood, in no small part because of conceptual and measurement issues related to stress. The nature of the relationship of stress and health is hypothesized to be indirect and is another difficulty in stress-health research.

### *Stress and Health*

In the typical study that examines the relationship between stress and health, participants have been exposed to a common stressor and a specific health effect is measured (e.g., resistance to infection, cardiovascular trouble, mortality). For example, Romanian children that had been adopted into the UK before the age of four showed profound psychological dysfunction in attachment, increased inattention/overactivity, heightened autistic features and greater cognitive impairment than children of the same age but adopted within the UK (Rutter, Kreppner, & O'Conner, 2001). Romanian institutions were at the time regarded as very poor, and therefore stressing, while those in the UK were considered superior. A greater number of the Romanian children demonstrated an indiscriminating social approach, a seeming lack of awareness of social boundaries, and a difficulty in picking up social cues on what is socially

appropriate or acceptable to other people. The pattern of inattention/overactivity was also greater in the Romanian sample. The authors commented that these symptoms are typically considered predominately genetic with only minor psychosocial influences.

Such studies emphasize the continuing development of our understanding of the effects of stressors and deprivation. Stressors, however, do not cause equal amounts of stress and the health effects can be varied. Additionally, the relationship between stress and health is indirect and therefore makes it more difficult to identify the existence of this relationship. Cassell (1976) considered this relationship the imperative charge of modern epidemiologists. In traditional epidemiology, the causal agent is sought for a specific disease. Stress can be considered an agent but it does not cause a specific disease. This makes a “cause-effect” relationship difficult to identify. This is especially true in analyses with aggregate data, such as using large groups of people that have been exposed to a suspected stressor because there would likely be different and multiple health effects. Therefore, the agent-disease approach is not appropriate for the stress-health model. Because stressors are assumed to influence health through adjusting a person’s endocrine levels and changing susceptibility to disease, no specific disease will always result from stress.

Cassell (1976) believed that stress influences an organism’s *host immunity*- the overall ability of the body to respond to threats. Constant stress leads to high cortisol levels that over time reduce the effectiveness of the body’s response to a perceived threat. Cassell recommended that instead of viewing environmental conditions only as stressors and non-stressors, they should also be considered as sometimes being

protective or beneficial. Psychological processes may improve host immunity, similar to what stress-immune researchers now consider coping (Jones & Bright, 2001).

Cassell (1976) charged modern epidemiologists to identify the characteristics of the psychological phenomenon that evoke major endocrine changes in the brain. Instead of looking towards concrete events and situations, the constants in stress-health relationship are likely to be psychological. Marmot (2005) considered these constants to be one's sense of control and social participation. People that had a high sense of control could cope better with stressors, and those who participated in activities their culture deemed necessary had good protective health. Consequently, they would be more resilient to a stressor than someone who had less perceived control and could not engage in as much social participation.

The health effects of stress are likely opportunistic, resulting from increased susceptibility caused by stressors. This susceptibility can be due to compromised immune functioning from the endocrine effects (such as prolonged cortisol production) or from unhealthy coping behavior (such as smoking). An ideal study of stress and health would require hundreds of measures of health, and this is assuming that one has a valid measure of stress.

Cassell (1976) suggested that social processes may influence health indirectly. Similarly, the psychological processes suggested by Marmot (2005) to influence health may result from social comparison. One's amount of perceived control and social participation were the defenses against stress' adverse effect on health. Perceived control would be related to social comparison in that it is *perceived* control that influences the impact of stressors. People gauge their ability to control a situation based largely upon

their previous experiences and what is observed in others. Perceived control, then, is evaluated in part through social comparison. Social participation, the ability to engage in activities one considers necessary to their position, also would be derived from social comparisons to similar others. Furthermore, entitlement (the arguable required state for RD) can be considered to result from a perceived inability to engage in social participation. One identifies social participation- what one should have and do- from comparisons to those deemed we identify as similar. This is how social comparisons may affect health. By serving as the psychological process by which perceived control and social participation is derived, and the RD experienced from such comparisons may be chronically stressing, thus lowering host immunity.

Status has been tied to health, and this relationship is not explained away by the possibility that those in higher status have access to better resources (Marmot, 2005). This status syndrome is an explanation for the finding that disease and mortality follow a social gradient. This is not to suggest that one's boss will always outlive them. While status is related to health, so too is one's assessment of one's status (an overall sense of "how one is doing").

Social comparisons are a possible mechanism by which an assessment of status influences health because social comparisons have a bi-directional relationship with self-esteem. The threat to self-esteem derived from social comparisons in the evaluation of status may be chronically stressing. This would have two effects on health. The first would be impaired host immunity that would increase the rate and duration of illness. The second would be employing coping mechanisms, which may be detrimental to health (such as smoking or risk-taking behaviors).

In studying the nature of social comparison, the processes by which status influences health can be predicted. Marmot's status syndrome may therefore be predicted by the nature of the social comparisons. If the underlying nature of the comparison processes that lead to status-related health effects can be identified, then control may be achieved and possibly even management of this relationship. The purpose of this research is to create a social comparison-based measure of RD/RG and test the relationship between RD/RG and status and health. These constructs will be reviewed here, followed by the measures selected to measure them.

### *Review of the Measurement of Selected Constructs*

#### I. Status

The construct of *status* is based upon the hierarchy of individuals within a given system (Marmot, 2005). Social status is the hierarchy of individuals along culturally valued dimensions. Marmot identified two predominate mechanisms of status: perceived control and participation. The relationship between these two constructs and health was referred to as the "status syndrome". He challenged dominant alternate explanations such as health selection (i.e. poor health causes lower social status, not the reverse), and spurious variables (e.g. education and genetics). While these variables do contribute to the so called status syndrome, they do not sufficiently explain why social stratification of health occurs in industrialized countries.

Marmot (2005) contended that the difficulties in identifying a relationship between status and health are due to an inability to adequately measure status. His experience with the relationship between status and health began with the Whitehall study, now known as Whitehall I, which began in 1967 with 18,000 male British civil

servants classified by pay grade (Marmot, Adelstein, Robinson, & Rose, 1978). When the sample was culled to only those aged 40-64, those on the bottom of the hierarchy had four times the risk of death as those at the top of the hierarchy. Even those second from the top of the hierarchy had a higher mortality rate than those at the top of the hierarchy. A subsequent study, termed Whitehall II, began in 1985 with 10,308 men and women enrolled in a continuing health study to further explore the status syndrome (Marmot, 2005). The study indicated that the status syndrome occurs in women as well as in men. The social gradient has been implicated in the development of a wide range of diseases such as cancers, depression, suicide, lung disease, and self-reported health. All of these disease processes were inversely related to status. This supports a consideration of status level as a stressor that compromises host immunity.

One explanation for the status syndrome is simply that those at the top have better funding for medical screening and care (Marmot, et al. 1997). Those at the lower grades were more likely to be identified as having the diseases and being treated for diseases. Marmot (2005) argued that this discounted the explanation of the top status individuals having better medical care because if those in the top tier of status lived longer because of better medical care, incidence would be higher at the top because of increased detection. The increased treatment of disease at the lower levels was indicative of the lower levels *having* more disease.

Marmot (2005) contended that at first blush the samples seemed homogeneous- all participants were non-labor based employees in an industrialized country; none were living in abject poverty and none had ostensible wealth. However, within the sample were clear stratifications that he felt genuinely represented one's perception of the

world. Marmot noted that the stratification of pay grade in the civil service represented differences in income, education, and the level of control in the job- the factors he hypothesized to create the status syndrome. Rather than being three orthogonal variables, each variable relates to the other to contribute to one's sense of status and control. A scientist studying each variable would likely run into difficulty identifying a health effect. For example, if education was used as a variable, the college graduates in religion may earn significantly less money than accounting majors. Likewise, studying income separately would be confounded by variations in perceived control and social status across equally paying positions. The Whitehall studies identified an effect because pay grade in the British civil service was a very precise classification system of status. Marmot considered status as a combination of education, income, and perceived control. RD/RG is considered here to be influenced by status as well as by the two contributors to the status syndrome (social participation and perceived control) but also to influence perceived control and social participation.

### *Measuring Status*

Citing the interest of Stouffer et al. (1949) in education as a predictor variable of the relationship between promotion and attitude, Runciman (1966) believed that there was a parallel to this finding occurring in Britain, and considered education as generally equivalent to status. Runciman (1966) contended that education was both the mode of social mobility and a contributor to RD. When a controlled sample of participants were given an infectious dose of an upper respiratory infection (URI) causing agent, participants with high school diplomas or less education had the most colds, while participants with some college had the fewest colds (Cohen, Frank, Doyle, Skoner, &

Gwaltney, 1998). This effect was significant even after controlling for age and antibody resistance. Marmot (2005) identified a health gradient tied to status, and considered status as a combination of education, income, and perceived control.

Previous research has identified a relationship between self-reported perceived control and health. The relative contributions of material deprivation, education, national economic inequity, and perceived control to self-reported health were examined among seven post-communist countries (Bobak, Pikhart, Rose, Hertzman, & Marmot, 2000). The data was obtained from surveys administered during the 1990s, subsequent to the collapse of communism in Eastern Europe. The authors noted that during this time period inequity was rising in many of the countries and was at best sustained in the countries. The seven countries included in the study were: Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, and Russia. Material deprivation was determined through the sum of responses to three questions relating to the frequency with which participants did not have all the food, clothing, and electricity/heating needed on a four-point scale. National economic inequity was measured with the Gini coefficient.

The Gini coefficient is a mathematical calculation of income distribution in a region (Yitzhaki, 1979). It is an overall measure of inequality in a given region that can range from 0-1, with a score of “1” denoting that one person has all of the resources and has been argued to be a measure of social inequity. Perceived control was assessed through aggregated ratings of agreement to nine questions (such as “At home, I feel I have control over what happens in most situations” and “I often have the feeling that I am being treated unfairly” (Bobak et al., 2000). The Cronbach’s alpha reliability score

for this study was 0.65. Self-reported health was measured by one question, “How do you rate your health over the last 12 months?”

The Gini value was not found to be a statistically significant predictor of self-reported health once individual characteristics (age, sex, education, and marital status) were controlled (Bobak et al., 2000). Education and material deprivation were, however, related to self-reported health, and perceived control was found to statistically significantly mediate the effect of material deprivation. Put differently, the impact of material deprivation on health appeared to be influenced by the extent to which an individual felt that they could exert control (change) a situation.

One’s objective situation may be less important to health than one’s beliefs regarding one’s ability to change that situation. Research suggests that in a sociological context, self-reported health measures are influenced by perception of control (e.g., Bobak et al., 2000). Measures of objective inequity do not uniquely predict social health gradients once perceived control and individual differences are accounted for.

#### *Measuring Social Participation*

Marmot (2005) considered perceived control and social participation to be the two factors contributing to the status syndrome. Social participation, the second construct of the status syndrome, is the degree to which a person engages in societal activities. For example, many members of the United States consider home ownership an important activity. Perceived control is likely both a component of status and an influencing process while social participation is a characteristic of status. At each status level participatory activities exist, but the list of those activities one can engage in increases as one moves up in social status. Marmot stated: “All societies have rankings because

individuals are unequal in a variety of ways; but not all societies have the same gradients in health. What matters is the degree to which inequalities in rankings lead to inequalities in capabilities- being able to lead the lives they most want to lead” (p. 240). The construct is expansive and ambiguous, complicating measurement. Every individual has a unique set of formal and informal membership groups, and each has a collection of activities.

Social participation has been related to mental health and civil order. Gough (1952) stated “The thwarted needs for group appreciation and recognition found outlet in socially disapproved ways. Once this assessment was made, the counseling process was simple and straightforward, and a constructive solution could be achieved” (p. 227). In a longitudinal study, social participation was found to be moderately stable over a twenty-year period (Hyypä, Maki, Alanen, Impivaara, & Aromaa, 2008). Ellaway & Macintyre (2007) did not find any consistent pattern of social participation to be predictive of risk factors for CVD, though participation in groups was found to be negatively correlated to distress, and this effect was stronger in men than women.

Social participation may be a broader construct than perceived control. One’s sense of perceived control is a more concise concept, while social participation has greater ambiguity about what is considered regular participation in formal and informal membership groups. Chapin (1928) measured social participation by recording the quality of membership to provided organizations. A score of 1 indicated membership, 2 for attendance, 3 for financial contributions, 4 for committee memberships, and 5 for holding office. Gough (1952) tested behavioral and psychological problems related to social participation using the *Minnesota Multiphasic Personality Inventory* (MMPI) as a

template (32 of the total 42 items were taken directly from the MMPI). The scale included such items as “I like science” and “When in a group I usually do what the others want rather than make suggestions”. This measure was designed to assess tendency towards social participation, operating on the assumption that social participation was important for psychological health.

The longitudinal study of social participation described above by Hyyppä et al (2008) recorded the frequency of attendance in clubs (e.g. concerts, sports events), congregations (e.g. church attendance), outdoor and productivity activities (e.g. hiking, fishing), hobbies (e.g. drama, singing), and studying and cultural interests (e.g. reading books, listening to music). Though the measure was not provided in the publication, the scale was likely lengthy, tailored to the population of interest (urban dwelling Finnish individuals), and influenced by the researchers’ beliefs regarding social participation (twenty years ago). Ellaway & Macintyre (2007) used reports of frequently participating in an exhaustive list of groups divided into: political groups, trade unions, environmental groups, civic groups, church or other religious groups, charitable organizations, education, art or music group or evening class, social clubs, and sports clubs, gym or exercise classes. The Assessment of Life Habits (LIFE-H) measures social participation in those with disabilities across dimensions such as attaining adequate fitness, recreation, and interpersonal relationships (Fougeyrollas, Noreau, Bergeron, Cloutier, Dion, St-Michel, 1998). There are multiple measures of social participation, and large variation between them, and each measure appears to target a specific population (much more than perceived control). Similar to RD, social participation seems to have a contentious history of alternate interpretations and varied approaches to measurement.

### Measurement of Stress

Stressor effects occur when a situation is evaluated as threatening or demanding, and resources to cope with the situation are deemed to be insufficient (Lazarus, 1990). Stress is measured in the present study because stress is predicted to be the mechanism by which RD/RG affects health. Measures of stress are objective or perceptive (Jones & Bright, 2001). Objective measures of stress, such as measures of life change events or daily hassles, attempt to assess the occurrence of events predicted to be stress inducing. Both perceived stress and objective ratings have been found to correlate significantly to stress, and both likely contribute uniquely to the stress health relationship (Cohen, 1986). The health effect of stress, however, is believed to be from the emotional response to an objective event (Lazarus, 1977). The success of objective stress measures in predicting health is therefore probably due to the frequency with which objective stressors are perceived as stressing (Cohen, Kamarck, & Mermelstein, 1983). Objective stress measures miss situations that may be stressing to an individual but which are not considered to be stressing by researchers. Furthermore, from a practical consideration, no measure could be considered truly “objective” because researchers determine what is considered a stressful event from either theory or data (i.e. neither could be considered completely immune to human perception). As such, objective measures not only fail to measure the emotional response to an event but also are likely to never fully assess the entirety of stress inducing events.

The perceptive measures of stress may be more suitable for assessing the relationship between stress and health as they measure the perceived stress derived from

events, coping processes, and individual differences (Jones & Bright, 2001). The objective measures provide an assessment that controls for participant bias.

### *Measuring Stress*

A reasonably psychometrically fit measure of perceived stress is the Perceived Stress Scale (PSS), a 14-item measure of the perceived stressfulness (Cohen, Kamarck, & Mermelstein, 1983). The PSS was designed around three factors of the stress experience: unpredictable, uncontrollable, and overloading. These factors have been found to be central to the stress experience (Cohen, 1978; Lazarus, 1977). The scale is evenly divided between positively and negatively worded items. It was designed for samples with at least a junior high school education (Cohen et al., 1983). The original study of the PSS found no sex or age differences, and was adequately reliable- both in coefficient alpha (ranging from .84 to .86) and test-retest reliability (.85 at the two day mark and .55 for the six-week mark). This study included three samples, two from a college in the Western United States and the third consisted of a community sample derived from a smoking cessation study.

Previous validation of the PSS identified relationships between the scale and measures of life events, social anxiety, depression, and health (Cohen et al., 1983). For life event scores, the PSS had a small to moderate correlation in all three samples. Age effects were found in regards to the relationship between the PSS and life events when the sample was divided into “young” (16-25) and “old” (36-70). While no significant overall age differences had been found in PSS scores, the relationship between PSS and number of life events was strong in the young ( $r = .65, p < .05$ ), while in the old this relationship was not found ( $r = .19, p > .05$ ). The PSS was found to be a statistically

significantly better predictor of health outcomes than the life-events score (Hotelling t-test  $p < .05$ ). Increases in social anxiety were associated with increases in perceived stress (Hotelling t-test  $p < .001$ ), and the number of life events was unrelated to social anxiety. Other authors have cautioned, however, that self-rated stress scales are mired by anxiety (Jones & Bright, 2001; Marmot, 2005).

The perceived stress index (PSS) has been found to be psychometrically fit and a superior measure of health than a life events scale (Cohen, Kamarck, & Mermelstein, 1983) and correlated to anxiety and depression (Spada, Nikčević, Moneta, & Wells, 2008). The PSS was found to be significantly correlated to affective and physical symptoms after controlling for the variance of the life events measures. This suggests that the PSS measures a unique component of the stress-health relationship not attained by life events measures.

## Measurement of Health

Similar to the measurement of stress, identifying an appropriate measure of health is also difficult because there are many processes and indicators of health to consider (e.g. mortality, self-reported symptoms, blood pressure, hormone measures such as cortisol, and immune measures such as immunoglobulins). In spite of these theoretical barriers, establishing that stressful events have a health effect has been moderately successful (Marmot, 2005). For example, research studying hostility behavior has identified a relationship between this personality type and coronary heart disease (CHD; Williams, Haney, Lee, Kong, Blumenthal, & Whalen, 1980). A hostile personality was found to be a predictor of CHD independent of other contributing variables such as blood pressure, serum cholesterol, and cigarettes smoking. Identifying an effect, however, does not explain the reason for the effect. For example, in the study of CHD and hostility, the authors argued that rather than hostility being causative, being non-hostile was likely therapeutic (and certainly both are possible).

The difficulty in interpreting stress-health effects is due to the indirect effect of stress on health. An increased risk of CHD for those under constant stress does not occur because “stress germs” get in the heart. Stress mediates the susceptibility to infectious disease through changes in immune and endocrine activity (Glaser, 2005). The HPA is a prominent system of study for understanding stress and health because the system maintains host immunity. The effect of stress on endocrine levels seems to depend on the duration of the stressor. Initial exposure to a stressor increases HPA activity (such as a spike in cortisol activity) (Miller & Cohen, 2005). Chronic stressors, however, create the opposite effect- the de-regulation of the system results (such as reduced cortisol

production). Because RD and RG are considered stable psychological states, RD is predicted to be an acute and chronic stressor and RG may function as a coping behavior, also durable (acting as a stress buffer).

### *Measuring Health*

Marmot (2005) stated “(I)t is a good deal more useful to ask people if they are healthy than if they are under stress. People who respond to a question on how healthy they feel by saying ‘poor’ are correct. This single question is remarkably predictive of subsequent risk of dying.” Idler and Benyamini (1997) examined the ability of self-reported health to predict mortality in 27 studies across several countries. The authors selected studies that were published in English, utilized representative community samples, and provided estimates of the effect of self-ratings of health on mortality or survival time after covariates for health status and socio-demographic factors were included in the analysis. They found self-reported health was a better predictor of survival than variables such as medical records or self-reports of medical conditions. The authors noted that the health self-rating question across studies, with some studies asking respondents to compare their health to others their age while other studies asked for an overall general rating. Other studies emphasized the current situation (e.g. “at the present time...”). All showed high predictability. The authors also noted that the ability of self-rated health to predict actual health was maintained across languages; suggesting the effect was robust to semantic wording and cultural variations.

Idler and Benyamini (1997) identified sex differences in the relationship of self-reported health and mortality: the unique effect of self-rated health on the prediction of mortality was more apparent in men than in women. The univariate distributions of self-

ratings of health did not differ as a function of gender, and both men and women referred to the same criteria when self-reporting health. The authors posit that the reason for the sex difference may be that women evaluate their health over their entire life searching for prevalent health problems such as existing conditions or in comparison to other women. Because women may be more likely to use social comparison in health evaluation, the validity of self-rated health is decreased.

The self-rated health question assesses the overall array of illness and could also be incorporating symptoms of undiagnosed diseases (Idler & Benyamini, 1997). This can also be interpreted as indicating that self-ratings of health are representative of human judgments of the severity of the current illness. Objective measures of health will not capture the severity of symptoms as well as self-reports. Also, the effects of interactions between conditions are not reflected by the additive measure of chronic disease. The authors noted that they had found no study that had tested the association between self-rated health and interactions of conditions. The self-rated health question also reflects family history. The authors noted that three studies included measures of family history or longevity and that two of these studies showed an independent effect of self-ratings when this family history was taken into account. Another consideration made by the authors is that self-rated health functions as a dynamic evaluation capable of incorporating trajectory of health as well as current level. Still another benefit to self-rated health is the implication that reverse-causation is also being measured. Self-rated health may affect behaviors that affect health. Poor perceptions may lead to less willingness to engage in self-care, adhering to screenings, medications, and treatment, and preventative practices, while good perceptions would increase willingness to

participate in health promotion. Lastly, self-rated health may also reflect within-person resources. Overall it appears that self-rated health may be an efficient rating of health, incorporating both stress and coping ability.

Self-rated health is a perceptive rating, but ideal studies of stress and health should include perceptive and objective measures (Jones & Bright, 2001). In our study, HPA activity is the biological mechanism predicted to influence health through mediation by the stress of relative deprivation. Therefore, HPA activity measures are considered as objective measures of stress and health. Cortisol is considered the most prominent endocrine indicator of a stress effect on health (Miller, Chen, & Zhou, 2007). However, conflicting findings have emerged from stress studies with regard to the hormone. There are two prominent reasons for this. The first is due to the diversity and interaction of stressors. Stressors can be short-term (*episodic*) or long term (*chronic*), and both are believed to have different effects on health (Maier & Watkins, 1998). Episodic stressors include such events as having an argument or a car accident. Chronic stressors include enduring problems such as taking care of a loved one with dementia or being in a hostile work environment. Episodic and chronic stressors, however, can be tied to one another, such as when one loses a job (episodic) and therefore becomes unemployed (chronic). Episodic stressors have been studied primarily in laboratory settings while chronic stressors in humans require field research. An organism does not receive exposure to only one form of duration of stress- an organism experiencing chronic stress is *also* likely to experience acute stressors, creating a complex interaction that is difficult to interpret.

The second reason for cortisol's standing as a contentious measure of stress is that other variables besides stress have been shown to also influence levels of the hormone. Circadian rhythms and the act of awakening influence cortisol levels (Czeisler & Klerman, 1999). Sex differences have also been found. Kirschbaum, Kudielka, Gaab, Schommer, Untied, and Hellhammer (1999) compared salivary cortisol among women in varying levels of menstruation and men in response to a stressor. Men's cortisol response to social stressors tended to be larger than women's. The gonadal hormone estradiol may be responsible for this difference (Kirschbaum et al., 1999). Estradiol hormone fluctuates greatly in women over the course of the menstrual cycle, with low quantities in the early follicle stage, peaking around ovulation, and then decreasing in the leuteal phase. Kirschenbaum et al. found that such social stress tasks as public speaking and performing arithmetic in front of an audience induced conditions suggestive of increased HPA activity- increases in ACTH, cortisol, heart rate, and reduced report of calmness. This response was greater in men than women and was believed to be at least partially due to estradiol levels.

Salivary immunoglobulin A (sIgA) is found in mucosa and is regarded as the main immune defense against oral antigens (Bosch, Ring, de Geus, Veerman, & Amerongen, 2003). Due to this "front line" status and the ease with which sIgA can be collected through salivary collection; the protein has been extensively utilized as a measure of immune functioning (Pawlow & Jones, 2005). Individuals with sIgA deficiency often have reoccurring and severe infections (Braconier, Nilsson, Oxelius, & Karup-Pedersen, 1984), and sIgA levels are negatively correlated with increased susceptibility to urinary tract infections (Brandtzaeg, Baekkevold, Farstad, Jahnsen,

Hohansen, Nilson, & Yamanaka, 1999). Similar to cortisol, acute stress may raise sIgA levels while chronic stress may decrease levels of the protein. Chronic stress exposure has corresponded with reduced sIgA output (Deinzer & Schuller, 1998).

Levels of sIgA may follow a social gradient, supporting the utility of measuring the protein in stress research (Evans, Der, Ford, Hucklebridge, Hunt & Lamber, 2000). In a longitudinal health study in Scotland utilizing a sample of 1000 participants, variations in sIgA were obtained based upon participants' sex, age, and socioeconomic status (status was derived from census data). All were significant unique independent predictors of sIgA. Women, lower social class, older age, and smoking were associated with lower sIgA. The effect of social status, however, was reduced when the smoking variable was considered in the prediction model (social status was significant at  $p = .024$  but was  $p < .10$  when the smoking variable was added). This finding does not refute the use of sIgA in studies of the status syndrome. First, even after accounting for smoking there was still a substantial portion of variance in sIgA explained by status. Second, smoking and status interact with one another. People of lower socioeconomic status smoke more as coping behavior (as well as engage in other risk behaviors, see Marmot 2005). A study of sIgA should therefore include a measure of smoking status as well as other risk behaviors.

Cortisol and sIgA, however, have not been found to consistently respond to stressors in a predictable manner. Vokmann and Weekes (2006) examined the effect of academic examination on cortisol and sIgA levels. Students with three or more major exams, papers, or projects had increased perceived stress (as measured by the PSS) compared to students in the first week of summer when school was not in session. No

significant differences were found between the groups in terms of cortisol and sIgA. However, when participants who had increased scores in PSS before and after the stressor were isolated and assessed, an increase in cortisol, but not sIgA, was found.

Pawlow and Jones (2005) tested the effects of relaxation training on both cortisol and sIgA levels. Significant decreases of both measures occurred subsequent to the training, and were statistically significantly lower than control subjects. This study of the effect of relaxation therapy on cortisol and sIgA levels underscores the influence of social variables on these two measures of HPA activity, while also demonstrating the influence of acute stressors.

Bosch et al (2003) noted that of 13 studies of the effect of sIgA on academic stress, half found an increase whereas the other half found a decrease in the protein. The reasons suggested for this inconsistency are similar to those with cortisol. The utility of using the measure in relation to stress health research is mired by non-stress influences as well as differing effects of acute and chronic stressors. Studies of stress and health have continually found that acute and chronic stressors produce opposing results (Bosch, de Geus, Ring, and Amerongen, 2004). RD/RG may have an effect on cortisol and sIgA, but the interplay of chronic and acute stressors may complicate identifying such an effect. Many of the studies of endocrine and immune health find effects that are small and within the range of healthy functioning (Glaser, Rabin, Chesney, Cohen & Natelson, 1999).

### Relative Deprivation/Gratification

As discussed previously, different approaches to measuring RD have been used due to the varying interpretations of RD theory. Specifically, Sen (1981) stated that two major interpretations of RD have resulted in two different approaches to measuring RD- Runciman's (1966) feelings-centered approach and Townsend's (1974) conditions-centered approach. Runciman's (1966) feelings-centered approach is targeting the emotional consequences of perceived inequity while the Townsend-based approach focuses upon income-inequality and RD within income-inequality.

Although variations in measurement procedures are often necessary as researchers attempt to apply the RD construct to other constructs, the previously discussed approaches to the measurement of RD vary along three important dimensions that compromise the coherence of RD theory. First, there is little consensus on which reference groups should be used when making comparative evaluations. Second, there has been an overriding focus on income or economic factors as the object(s) included in such comparisons. Third, there have only been very limited efforts to assess RG and the interaction or combination of RD and RG.

The majority of RD research has been performed using comparisons groups that have been chosen on an a priori basis by the investigators. This may be useful in evaluating the RD experienced if or when comparisons are made to a designated comparison group, but does not determine if any of the participants actually utilize that comparison group in naturalistic settings. Essentially, researchers assume that their population sample is making the comparison to the selected group.

Research into the influence of RD on prejudice or health might benefit by determining whether the RD resulting from comparisons to the groups chosen by the researcher reflects the participants' RD, or if this is contrived condition that may be affected by other comparisons. For example, an individual may have a larger income than the majority of the people in the surrounding neighborhood, but still experience a high level of RD because that individual also focuses on comparisons to friends who have still higher incomes and may or may not reside in the immediate neighborhood. If researchers selected the surrounding neighborhood as the comparison group, the individual would likely report a lower level of RD than if friends were specified as a comparison group.

Smith & Leach (2004) are a notable exception to this critique. Through having participants record who they were comparing themselves at random intervals they found that participants were in the process of making a comparison 70% of the times they were paged, generally make more individual than group comparisons, and use their previous self as a comparison group. Researcher-selected comparison groups assess the level of RD experienced when comparisons are made to the selected group made but not *whether* the participants compare themselves to that group. Based on the example of Leach & Smith (2004), there is no reason to assume neighbors are a significant comparison group. Selection of comparison groups should be research driven, not researcher driven.

The next area of concern regarding the assessment of RD relates to what is being compared or the so-called "objects of comparisons". The most common objects of comparison have been participants' economic status, typically including their income and/or possessions. Income is a commonly used object of comparison by researchers but

this may be due to the ease of data procurement and analysis rather than an assumption that this is the only primary object on which people compare themselves to others. For example, people may compare themselves to others based upon on their accomplishments or the number of friends that they have. These comparisons could create RD. In support of this, Leach & Smith's random interval comparison group assessment research found that comparisons were made concerning money as well as relationships and family. RD consists of distress-inducing perceptions of inequity and a scale measuring RD should assess an individual's thoughts and beliefs while making as few assumptions as possible about the participant's objects of comparison or comparison groups.

The final limitation of extant RD measures involves the failure to measure the related condition of RG. The feelings produced by perceived advantage may, in part, offset the feelings caused by perceived disadvantage. For example, if an individual felt RD based upon comparisons to a particular ethnic group, but felt RG due to comparisons with other ethnic groups, the cumulative effect might be different than if RD was felt relative to all of these groups. RD and RG may be opposite ends of a single continuum or related but not orthogonal conditions but by focusing on only one, researchers may to have only been measuring half of a construct. For example, recent research assessing the influence of RG on prejudice suggests that while RG may be the inverse of RD, the impact of RG is not necessarily the opposite of RD (Dambrun et al., 2006).

By measuring RG and RD together, direct comparisons can be made that assess if RG has the inverse effect of RD as well as any cumulative conditions the two may create.

From the literature reviewed above, inconsistencies are apparent in both the theoretical and the implementation in RD theory. Crosby (1982) critiqued the implementation of RD theory stating that there is no clear focus of RD theory as social scientists focused on what created RD while political scientists focused on the consequences of RD. Though written twenty-five years ago, the critique that RD research suffers from a lack of operational focus appears to still be a legitimate concern. The above-discussed research review demonstrates a lack of a coherent framework of RD. Rather, the majority of RD research has attempted to identify situations where RD *may* occur; identifying set comparisons from which RD can emerge rather than the situations under which RD *does* emerge. Given the lack of coherence in the application of RD theory, the future research of RD should remain grounded in the work of the original researchers of the field. At the very least, current efforts to measure RD should adhere to the lessons learned by previous efforts, such as the multiple definitions and the importance of directly measuring dissatisfaction. Therefore, researchers developing a measure of RD should make as few assumptions about the comparison groups and objects of comparison as possible, and rely upon the population of interest to supply this information. From economics (Yitzhaki, 1979) to mental health policy development (Eibner, Sturm, & Gresenz, 2004), RD is a construct with profound implications for social scientists in viewing human interaction and therefore requires careful deliberation about the appropriate method of measurement.

Given the interest in measuring RD as opposed to income inequality, and our posit that RD and RG can occur regardless of income distribution, the current study will

follow Runciman's conception of RD and his approach for measuring feelings of inequity.

Under Runciman's (1966) definition, a person is relatively deprived when four conditions are met 1) when the individual does not have "X", 2) when the individual sees other person or persons which may include himself at some previous or expected time, as having X, 3) when the individual wants "X", and 4) when the individual sees it as feasible that they should have "X". Few RG scales were found in the literature but the general trend was to measure it as the inverse of RD (Yitzhaki, 1979; Guimond & Dambrun, 2002; Eibner & Evans, 2005).

A measure of RD/RG should assess the comparison groups an individual uses (Runciman, 1966, Sen, 1981), the "who" of social comparison. No two individuals, however, have identical reference groups. Therefore, likely comparison groups that most individuals compare to on some level would provide the best approach for an overall measure of RD/RG. Additionally, the objects of comparison, what is being compared, should also be identified. Finally, an RD/RG measure should assess the perceived inequity to the comparison groups. The approach for identifying the extent to which an individual derives distress from comparisons to prevalent groups along prevalent objects will not perfectly measure an individual's overall level of RD/RG. This is because no individual will likely have the exact groups/objects identified (with no other groups/objects beyond those identified). By including likely comparison groups and objects for comparison the authors hope not to perfectly capture RD, but gain a reasonable assessment that can be used in many situations and many different groups of

people, a “global” measure of RD. The RG component of the scale was similar in approach to the steps taken to develop the RD measure.

In a global measure of RD/RG, feelings of deprivation/pleasure would be assessed for a variety of likely comparison groups. For every comparison group and every object of comparison that is assessed, a measure of displeasure this comparison creates is necessary. The validity of such a design depends upon the appropriate selection of comparison groups and objects of comparison that are used by the population of interest. Measuring RD/RG through personal assessment may be difficult given that someone will quite likely have several comparison groups with which they identify. However, Hyman (1942) stated the number of reference groups a person uses is small and likely to relate to specific problems. This suggests that reference groups can vary in location but are relatively constant and stable across time. However, it is unlikely that the comparison groups will be the same across participants. The selection of comparison groups should be specific enough to discriminate individuals but not so much so as to exclude participants. For example, selecting “brothers” as a comparison group would be too specific and cumbersome, as a “sisters” counterpart should be included, followed with possibly a “mother”, “father”, perhaps also “grandmother” and “grandmother”. In the above example, these potential comparison groups were collapsed into “family members”. The participant would, in theory, read into the comparison group whichever individuals they define as their family members, negating the need for questions about specific family members. This was the approach used for the current measure of identifying comparison groups- attempting to be specific enough to separate

groups but vague enough to allow individual variation in whom specifically belongs in these groups.

Besides comparison groups and objects of comparison, there are other concerns to be considered for scale development. Given that the comparison groups and the perceived distress felt from comparisons to these groups will fluctuate with the individual, a global measure of egoist RD/RG should be elastic in the way in which it assesses perceived deprivation/satisfaction across comparison groups.

In the review above, the argument is made that social comparisons are unavoidable and when one compares oneself to others and is distressed from them they will suffer stress. This stress is likely enduring (chronic) and significant health problems may emerge from compromised host immunity. Furthermore, RG may be therapeutic. If these predictions are accurate, then understanding one's selection of comparison groups and the dimensions of comparison are necessary for predicting the impact of social comparison and health. The validation of a measure of RD/RG should be achieved through identification of significant paths between the construct and status, stress, and health. First, the operationalization of the constructs of RD/RG, status, stress, and health are provided here. Then, a proposal is provided for a two-part study to identify prominent comparison groups and objects of comparison and then build a measure of RD/RG from these responses and test the construct in relation to status, stress, and health.

#### *Measuring Prominent Comparison Groups and Objects*

To identify comparison groups and objects, the researchers created the Prominent Comparison Groups and Objects (PCGO) scale (Holland & Kern, in prep).

Expected comparison groups were selected based upon previous research of social comparison and temporal comparison theory. The only group specifically stated under Runciman's (1966) conditions for RD was previous self. Therefore this group was included. Schor (1998) also identified the following groups of comparison: coworkers, family members, friends, and members of the same religion. Schor (1998) considered media figures as possible comparison groups. This can include characters displayed on television even if these characters do not actually exist. Therefore, "fictional characters" were included as possible reference group to include those individuals the participant watches on television, in league with Schor's theory of a media forming a reference group. In summary, the comparison groups identified as prevalent from previous research were: previous self, coworkers, family members, friends, members of the same religion, and fictional characters.

Similar to comparison groups, prominent objects of comparison were also identified. Previous literature has primarily used financial-based comparisons (Runciman, 1966; Townsend, 1974; Yitzhaki, 1979). The use of financial objects makes sense using Runciman's (1966) conditions. Whatever "X" is, it can probably be bought. Therefore, one object of comparison in an RD/RG scale should be a "financial" object. Not everything one might use for comparison, however, can be purchased. As such, there are likely other prominent objects of comparison in addition to financial objects.

A second prominent object of comparison for consideration is interpersonal relationships. Wilkinson's (2005) argument that the quantity and quality of interpersonal relationships affects health and is used as a standard of comparison suggests the possibility that interpersonal relationship comparisons may occur and result in RD. A

third potential object of comparison is a general, non-specified object of comparison. Recall the first condition of RD specified by Runciman (1966): When an individual does not have “X”. The “X” is not defined but assumed by much of the previous research to be assessed through what the current authors refer to as the “financial” object of comparison. Additionally the current authors posit that interpersonal relationships are an object of comparison. These are two objects of comparison but there are likely other prominent objects of comparison. However, if there is a good or service the individual desires, this can be measured without specifying what “X” truly is. The definition of “X” need not be defined by the test makers but by the participants. If asked “I feel my coworker has it better than me” the respondent determines the degree of agreement with this statement based on what “Xs” they find important. The term “general” describes overall comparison- one that is not limited to financial gains or interpersonal relationships but functions as a wide-ranging assessment. Questions measuring comparisons for an RD/RG scale should consider “X” as being possibly financial, interpersonal, or general to reduce the possibility of excluding possible sources of comparisons.

To identify these prominent comparison groups and objects, the researchers created the Prominent Comparison Groups and Objects (PCGO) scale. Expected comparison groups were selected based upon previous social comparison research, temporal comparison theory, and the previous testing of the PCGO (Holland & Kern, in prep). Participants were given a set of comparison groups to rate for the “general” object of comparison, then the “financial”, followed by the “interpersonal” object of comparison.

Participants were provided with the following instructions at the beginning of the questionnaire:

Think about how well you are doing generally. Now that you have the idea about how well you are doing generally, how did you reach that conclusion? To what degree did you compare yourself to the following groups?

Comparison groups were then provided and participants rated the degree to which they believed that they compare themselves to each group. Participants were asked to rate their degree of comparison to each comparison group on a scale of 1-5 with 0 indicating “no comparison” and 5 indicating “heavy comparison”. For example, the item for the comparison group “coworkers” and the “general” object of comparison stated “To what degree did you compare yourself to your **coworkers?**”

To identify additional comparison groups not listed, the following question was included after the ratings for comparison groups for the “general” object of comparison:

Are there other groups you compare yourself to that we did not include? If so, please write them in below along with how important this comparison is to you. Three blanks are given, but you can write more if you need to.

This item is repeated for the “financial” and “interpersonal” comparison objects. The “general” object set of ratings was given to participants first to prevent financial and interpersonal-based comparisons from influencing participants’ assessment. To identify additional objects of comparison, the following item was included following the section instructing participants to identify additional comparison groups:

We just asked you to compare yourself to others. We asked you to compare yourself on how well you were doing generally. What was it that you compared yourself on?

The PCGO was created with the intention of identifying prominent comparison groups and objects that are used for the development of the RD/RG questionnaire.

### *Measuring RD and RG*

The findings of the PCGO would dictate the items for a global measure of RD/RG. The currently proposed RD/RG measure is based upon self-reported inequity of a specified reference group in a specified condition (e.g. comparisons to coworkers on emotional experiences). This is intended as a measure of magnitude. This alone may be sufficient as a measure of RD/RG because although it is an objectively phrased item, participants would likely not rate their comparison objectively. If participants believed they were doing worse than someone and were distressed by the disparity, they might answer this item with a lower value than another participant who felt the same disparity but was not distressed about the situation. This item measures perceived disparity similar to the local area index measures of RD and the condition-based approach, but is rated by the participant instead of being derived from data (such as census data).

Previous literature has placed a heavy emphasis on financial-based comparisons. Researchers developing a global RD measure will should identify other objects used in addition to financial-based comparisons. For example, an individual may compare with one group in terms of physical prowess, another in intelligence, another in employment and still another in philosophy. One possible additional object of comparison is interpersonal relationships. Smith & Leach (2004) found that many comparisons participants made were about relationships and family. Wilkinson (2005) stated that the quantity and quality of interpersonal relationships affects health. Eibner and Evans' (2005) research found a relationship between their measure of RD and health. The impact

of interpersonal relationships on health suggests they may be used as an object of comparison. In order to identify prominent objects of comparison, an approach similar to the assessment of comparison groups should be utilized in which a list of potential objects are provided for participants to rate as well as open-ended questions to find additional objects of comparison. Wheeler and Misayke (1992) used self reports of comparisons to identify prominent comparison groups. Common referents (>5%) were close friends, ordinary friends, acquaintances, strangers, family members, imaginary persons, significant partner, and famous persons. Prominent dimensions included academics, personality, physical appearance, lifestyle, abilities, social skills, wealth, and opinion. Sex differences were also identified. Women were statistically significantly more likely to compare with a family member and men were more likely to compare with a famous person. Women were more likely to compare themselves on appearance and men were more likely to compare on opinion. Leach & Smith (2006) followed up their previous study (Smith & Leach, 2004) by evaluating if an ethnic minority would report differences in both the type of comparisons (individual to group, group to individual, etc) made and results from such comparisons. The researchers used open-ended questions to identify what we call objects of comparison (Smith & Leach referred to them as “domains of comparison”). The objects identified were academic skill, wealth, social status, personality, life experience, and relationships. Measurement of these objects is more likely to validly assess RD than a list of objects derived solely from the researchers.

## Hypothesis

Using the above described measurement approach; the following predictions will be tested:

- Self-report data will yield prominent comparison groups
- Self-report data will yield prominent comparison objects
- The RD/RG measure will have adequate psychometric properties within the populations of interest
- Status will significantly predict health
- Status will significantly predict RD/RG
- RD/RG will significantly predict health
- RD/RG will significantly predict health over and above the affects of stress, anxiety, and status

## CHAPTER 3

### STUDY 1: IDENTIFICATION OF PROMINENT COMPARISON GROUPS AND OBJECTS WITH A COLLEGE SAMPLE

The purpose of the present study was to determine the self-rated perceived importance of comparisons made towards likely comparison groups and objects of comparison, as well as to identify additional prominent comparison groups and objects of comparison. Unlike previous measures of prominent comparison groups (e.g. Smith & Leach, 2004), the measurement approach employed is a single instance self-report measure. This study is exploratory in that it seeks to identify prevalent comparison groups and objects of comparison. However, the research also attempts to validate groups/objects identified in previous research. Specifically, the *a priori* comparison groups tested were: coworkers, family members, friends, members of the same religion, previous-self, and fictional characters. The *a priori* objects of comparison tested were: general, financial, and interpersonal.

Hypotheses:

- The proposed comparison groups will be statistically highly rated as tested through one sample t-tests
- Open-ended items will identify additional prominent comparison groups as will be assessed through two tests: the proposed group being reported in 2% of the total sample and the group must be theoretically likely to create both RD and RG.
- The proposed comparison objects will be statistically highly rated as tested through one sample t-tests

- Open-ended items will identify additional prominent comparison objects as will be assessed through two tests: the proposed group being reported in 2% of the total sample and the group must be theoretically likely to create both RD and RG.

## Method

### *Participants*

Two hundred and thirteen participants participated (148 female) for course credit. Incomplete questionnaires were removed, resulting in 199 participants (141 female) with a mean and median age of 19. Prominent ethnic groups of the sample included Caucasian (53%), Asian (17%), Hispanic/Latino (11%), and African American (9.4%), with the remainder self-reported “other”.

### *Procedure*

Participants were given the Prominent Comparison Groups and Objects scale (PCGO). The PCGO contained a list of comparison groups in which participants indicated the importance they ascribed to each group. The following instructions were provided at the beginning of the questionnaire:

Think about how well you are doing generally. Now that you have the idea about how well you are doing generally, how did you reach that conclusion? To what degree did you compare yourself to the following groups?

Comparison groups were then provided and participants rated the degree to which they believed that they compare themselves to each group on a 0-5 scale (with 0 indicating “no comparison” and 5 indicating “heavy comparison”). For example, the item

for the comparison group “coworkers” and the “general” object of comparison stated: *To what degree did you compare yourself to your coworkers?*

The comparison groups were coworkers, family members, friends, members of the same religion, previous-self, and fictional characters (described in the questionnaire as “TV, movie, or book characters”). The scale was designed in such a way so as to present each comparison group within each object of comparison. Participants were given the set of comparison groups to rate for the “general” object of comparison, then the “financial” object, followed by the “interpersonal” object of comparison.

Since the sample consisted of college students, the participants were instructed to consider fellow students as coworkers. To identify additional comparison groups not provided, the following question was administered:

Are there other groups you compare yourself to that we did not include? If so, please write them in below along with how important this comparison is to you. Three blanks are given, but you can write more if you need to.

This question was repeated subsequent to the financial and interpersonal objects of comparison. The general object set of ratings was given to participants first to prevent financial and interpersonal-based comparisons from influencing participants’ assessments of the broad general area. To identify additional objects of comparison, the following question was written below the section instructing participants to identify additional comparison groups:

We just asked you to compare yourself to others. We asked you to compare yourself on how well you were doing generally. What was it that you compared yourself to others to?

## Results

### *Analyses*

To test the above-described hypotheses, the data from the PCGO fell into two categories: values derived from the close-ended items evaluating the provided comparison groups and values derived from the open-ended items requesting additional comparisons. Evaluation of the provided comparison groups and objects of comparison was by statistical criteria. The ratings scale ranged from 0-5. A one-sample t-test was performed to identify if mean ratings were greater than a rating of 2. The cutoff mean of 2 was selected because this value is less than half of the scale and corresponds to the response of “I care a little about my comparison to this group”. Any comparison group with a mean less than half would not be a popular enough candidate for an overall measure of prominent groups. To avoid inflating type-1 error by conducting this analysis for every comparison group by every object of comparison, comparison groups were collapsed by comparison object. Any non-significant comparison group was analyzed individually to evaluate if the comparison group was significant by any of the objects of comparison.

Evaluation of the open-ended responses was conducted using both quantitative and qualitative analyses. First, the answers were categorized. For example, “my wife” and “my life partner” were categorized as “significant other”. Then, in terms of quantitative evaluation, any new object or group that was added by less than 2% of the participants was excluded from further analysis. Besides meeting mathematical criteria, the objects and comparison groups provided by the open-ended responses were required to be theoretically plausible groups or objects that could create status identification.

*Selected Comparison Groups:*

Hypothesis: *The proposed comparison groups will be statistically highly rated as tested through one sample t-tests*

Analyses were performed to test the significance of each comparison group within each object of comparison using one-sample *t*-tests, resulting in three tests for each comparison group. Two comparison groups were not statistically significantly greater than the value of “2”; members of the same religion and fictional characters both had means below a value of 2. Effect size calculations identified that the object/group combinations of financial family and financial coworkers both had effect sizes that were between  $\eta^2 = .2$  and  $\eta^2 .5$ , placing the effect size as greater than small but not meeting the cutoff for medium ( $\eta^2 = .31$  and  $.34$  respectively). All other significant effects had sizes above the “medium” cutoff of  $.5$ .

Gender differences were not statistically significant except for groups of the same religion and fictional characters (see table 2). However, the means for both females and males was below 2 for both members of the same religion (1.46 and .98 respectively) and for fictional characters (1.70 and 1.19). The hypothesis was supported for the comparison groups: friends, family, coworkers, and previous self, but not for members of the same religion or fictional characters.

*Selected Objects of Comparison:*

Hypothesis: *The proposed comparison objects will be statistically highly rated as tested through one sample t-tests*

The objects of comparison mean scores were derived from calculating the means of each comparison group within each object of comparison. For example, the mean score

for the general object of comparison was derived from the mean of all ratings of the six comparison groups within the “general” object. One-sample t-tests comparing the ratings of the objects of comparison to the cutoff of 2 indicated that the means of all three objects (general, financial, and interpersonal) were significantly greater than the cutoff. The effect sizes for general and interpersonal comparisons were large (above .80). The effect size for the financial object of comparison was  $\eta^2 = .23$ , a small effect.

Gender differences were examined by splitting the analysis based upon the self-reported gender of the participant. For females, the mean rating of all three objects was statistically significantly greater than 2. The mean ratings for males, however, were statistically greater than 2 for the interpersonal and general objects of comparison, but not the financial object ( $t(58) = .633, p = .53$ ). The hypothesis was supported for all three selected objects of comparison except when the sample was split by gender- then males only rated the interpersonal and general comparisons statistically significant.

*Identification of Additional Comparison Groups:*

*Hypothesis: Open-ended items will identify additional prominent comparison groups as will be assessed through two tests: the proposed group being reported in 2% of the total sample and the group must be theoretically likely to create both RD and RG.*

Sixty-one participants identified additional comparison groups beyond those given in the questionnaire. Any response that was given by more than two percent of the entire sample ( $n = 4$ ) was a candidate for inclusion. After coding the responses, six groups emerged: athletes ( $n = 5$ ), celebrities/TV characters ( $n = 4$ ), members of the same gender ( $n = 4$ ), significant partners ( $n = 5$ ), strangers, people the participant sees in social settings such as clubs and bars ( $n = 7$ ), and teammates ( $n = 4$ ). After theoretical consideration,

significant partners was removed for consideration as this group was not considered to cause RD and RG independent of the existing groups of friends and family members and that unlike the other groups, not all participants have a significant other (while participants are more likely to have coworkers, friends, and family members they do not necessarily have these groups. However, the likelihood of a participant not having a significant other was too great to consider the group for inclusion). The hypothesis was supported in that additional comparison groups were identified from open-ended questioning.

*Identification of Additional Objects of Comparison:*

*Hypothesis: Open-ended items will identify additional prominent comparison objects as will be assessed through two tests: the proposed group being reported in 2% of the total sample and the group must be theoretically likely to create both RD and RG.*

Open-ended questioning was used to evaluate the identified objects (general, financial, and interpersonal) and identify additional objects of comparison. Responses were categorized by similarity and evaluated using the 2% frequency cutoff. The item used in the PCGO for identifying objects of comparison was not optional, unlike the open-ended item identifying additional comparison groups. Therefore more participants responded to this question (n = 171). Responses were categorized based upon perceived similarity. For example, the statements “what I own” and “finances” were coded as “financial”. Participants frequently provided multiple responses, yielding 199 coded responses. The responses that were similar to those previously provided to the participants were also recorded and all three were identified by more than 2% of the participants (general, n = 11; financial, n = 23; interpersonal, n = 40). However, three

additional objects of comparison provided by the participants included academics (n = 14), accomplishments (n = 15), and appearances (n = 16). In theoretical consideration, accomplishments and academics may be similar. Additionally, if this scale were applied to a non-academic population, the academic object would likely not be appropriate. As such, the object of academics was considered to be removed from an RD/RG scale. The hypothesis was supported in that additional objects of comparison were identified: accomplishments, and appearances.

### Conclusion

Recent research in social comparison has made considerable progress in identifying individual differences in social comparisons but research investigating the *kind* of social comparisons is still deficient. This study was conducted to test the strength of hypothesized comparison groups and objects of comparison while also identifying other groups and objects that have not been previously identified. A single-instance self-report measure was used to identify prominent comparison groups and objects of comparison. This research is important not only for the prominent comparisons that emerged but from the technique used to identify them. Both close-ended and open-ended items were used. Close-ended items were evaluated based on the strength of the comparison while the open-ended items were evaluated by the popularity of the response in the overall sample. The close-ended items were participants' ratings of comparison groups and objects that were identified in previous research. The significance of the comparison groups identified in prior studies and the participants' supplied comparisons were tested. Open-ended items allowed participants to provide additional comparison groups/objects that had not been previously identified by research but that the participants found to be meaningful. These

were evaluated through three stages: 1) by popularity- all comparisons made by 2% or more of the sample were considered, 2) novelty- the answer did not duplicate a previous answer, and 3) theoretical plausibility - likely to be used in comparisons that can create emotional states (such as RD).

In terms of comparison groups, the prominent comparison groups provided by the researchers were coworkers, family members, friends, previous-self, members of the same religion, and fictional characters. Evaluation of participant ratings identified that two of the previously identified comparison groups, the members of the same religion and fictional characters groups, were not significant. For the objects of comparison, all three of the groups identified by prior research were supported through participant ratings: general, financial, and interpersonal. This was true regardless participants' sex. Effect size calculations identified that all comparisons were at least of "medium" strength with the exception of financial comparisons to family members and coworkers. This suggested that participants generally did not consider financial comparisons to these groups, or that such comparisons were not as important as financial comparisons to other groups (such as friends).

Open-ended questioning was used to identify any additional comparisons not specified by prior studies selected by the researchers. The quantitative criterion for selection was that each group had, at a minimum, at least two percent of the participants supplying the new comparison group. This resulted in the identification of seven possible comparison groups for inclusion: significant others, teammates, athletes, celebrities/TV characters, strangers, people seen in social settings such as clubs and bars, and members of the same gender. Additionally, a theoretical criterion for inclusion was that the

comparison groups not duplicate a previously provided comparison group and would potentially elicit comparison derived emotional states.

Two of these participant supplied groups were similar- celebrities/TV stars and athletes. These are both groups that the individual watches on television or in films but does not physically interact with. Therefore, these two groups could be considered one comparison group: celebrities/athletes. Celebrities/TV characters are possibly similar to fictional characters, which was included in the survey but was found to be a statistically non-significant comparison group. Rather than drop this group and add the celebrities/TV characters, these groups were merged into “fictional characters, athletes, or celebrities”. The reported groups “strangers” and “people the participant sees in social settings such as clubs and bars” are similar to the “general” object of comparison because they represent a broad group of individuals that the individuals interacts with but are not familiar with. Because of this, this comparison group was considered similar to the general comparison group.

In addition to duplication of other comparison groups, some of the provided groups were similar to comparison objects. For example, “significant others” are unlikely to be used as a comparison group. They are more likely to be used as an interpersonal object of comparison (e.g. “I have a wife, therefore I’m doing better off than most”). Furthermore, significant others can be considered family and/or friends, two comparison groups already included. Lastly, while most participants are likely to have members of the selected groups such as family members (or members they *consider* to be like family) not everyone has significant others and inclusion of such a group could be exclusionary. Another contentious comparison group, teammates, was considered to be too exclusive of

a comparison group to be considered as a relevant comparison group. Additionally, teammates are often regarded as friends and would fall into an already existing comparison group.

Therefore, after theoretical consideration, the only comparison group added to the researcher-designed list of prominent comparison groups was members of the same gender. The resulting list of prevalent comparison groups contained the following groups: friends, family members, coworkers, previous self, celebrities/athletes, members of the same gender, and general.

The selected prevalent objects of comparison (general, financial, interpersonal) were significantly relevant to participants. Males, however, did not give statistically significant ratings for the financial object of comparison. This may be due to the majority of males not making significant financial comparisons or not making significant financial comparisons to the selected comparison groups. Effect size calculations identified the interpersonal and general objects as strong magnitude effects while financial was a low magnitude effect. This finding suggested that participants generally did not conduct financial-based comparisons or that participants were not actively aware and/or willing to admit that they do so.

The low rating of the financial object suggests problems with using the word “financial” to communicate that the term includes not only physical money, but also material means. For example, participants may not have understood that cars, houses, vacations, clothes, and virtually all material objects are included under the concept of “financial”. The low reported importance of financial-based comparisons is an important consideration for a scale that requires self-assessment because much of the previous

research on consequences of social comparison (such as RD) suggests that financial-based comparisons do indeed occur. The current finding regarding the low self-rated importance of financial comparisons among some males should be studied further to determine if using the term “financial” adequately represents the construct of material means.

Open-ended questioning was used to provide an additional evaluation of the selected objects of comparison and to identify additional objects. The quantitative criteria for inclusion were that the object be listed by at least 2% of participants. As a check of the selected objects, the number of participants who reported each of these three selected objects was above this value. Additionally, based on open-ended questioning, three new objects of comparison were considered: academics, appearances, and accomplishments. Appearances and accomplishments were both added because there did not appear to be any sufficient arguments against their inclusion. The object “academics”, however, is unique to the academic setting and as such is not recommended for general use, but should be considered with student populations. In summary, all three researcher-supplied objects of comparison were supported by the data and two additional objects were added, resulting in five objects of comparison: general, financial, interpersonal, appearances, and accomplishments.

One limitation of the current research was the population sampled. The participants were college students at a southwestern U.S. university. Some responses from participants may be unique to college students. For example, the finding that academic accomplishment is an important object of comparison probably will not generalize to a non-student population. Similarly, the low rating of financial objects of comparison in

males may also be an artifact of a college-based sample. The low ratings of importance for members of the same religion and fictional characters may also be due to the college sample and may prove to be important to a non-college sample. Therefore, we are currently conducting this study with a sample from the general population of adults. Another limitation of the current research is that the questionnaire did not include items that directly measured the object of comparison. This was instead inferred through the collapsing of the comparison groups within each object. While participants identified other objects of comparison through open-ended questioning, the lack of questions dedicated to objects of comparison is a limitation of this study. This procedure was adopted because the present study was focused on comparison groups. A subsequent study focusing on objects of comparison may yield additional objects of comparison. A third limitation of the current research was the self-report technique employed, which may suffer from participants' lack of awareness of their own engagement in conducting social comparisons and their propensity to answer according to social norms rather than earnest feelings of comparison. For example, several participants answered the open-ended questioning with such statements as "I do not compare myself to others" but they nonetheless completed the survey, giving non-zero values to the supplied comparison groups. The survey's introductory statements attempted to control for this potential problem following the recommendations supplied by Wood (1996) for the use of self-report measures of social comparison theory.

This research identified prevalent comparison groups and objects of comparison, resulting in the identification of seven comparison groups: friends, family members, coworkers, previous self, celebrities/athletes, members of the same gender, and general,

as well as five objects of comparison: general, financial, interpersonal, achievements, and appearance. Additional research should explore the variation of importance of select comparisons across demographic variables, such as sex. The present research established prominent comparison groups and objects useful for the development of a general scale to measure relative disparity (RD) and relative gratification (RG).

## CHAPTER 4

### STUDY 2: PERCEIVED INEQUITY SCALE DEVELOPMENT AND VALIDATION WITH A COLLEGE SAMPLE

The purpose of this study was to develop a measure of RD and RG derived from the prominent comparison groups and objects identified in the previous study. These comparison groups were: general, friends, coworkers, family members, previous self, fictional characters, athletes, or celebrities, and members of the same gender; as well as five objects of comparison: general, financial, interpersonal, achievements, and appearances.

The current RD/RG measure is based upon self-reported inequity of a specified reference group in a specified condition (e.g. comparisons to coworkers on emotional experiences). This is intended as a measure of magnitude. This alone may be sufficient as a measure of RD/RG because though it is an objectively phrased item, participants would likely not rate their comparison objectively. If participants believed they were doing worse than someone and were distressed by the disparity, they may answer this item with a lower value than if another participant felt the same disparity was not distressed about the situation. This item measures perceived disparity similar to the local area index measures of RD and the condition-based approach, but is rated by the participant instead of derived from data. If this measurement approach is found to measure RD and RG, health indices will be explored to identify the relationship between the RD/RG measure and health.

Hypotheses:

- Magnitude and intensity items will emerge as unique dimensions as tested through exploratory factor analysis (EFA), independent t tests, and correlation.
- Magnitude and intensity will have sufficient reliability as tested by Cronbach's alpha and test-retest reliability.
- Magnitude and intensity dimensions of both RD and RG will not have statistically significant sex or ethnicity differences as tested through analysis of variance (ANOVA)
- Magnitude and intensity dimensions of both RD and RG will meet construct validity as tested by correlation with related variables as well as constructs predicted not to be related to the RD and RG.
- Magnitude and intensity dimensions of both RD and RG will differently predict self-reported health measures as tested through linear multiple regression.

## Method

### *Participants*

A total of 246 participants (150 female) participated in the study. Ages ranged from 18 to 42 with a mean of 20.1 and median of 19. The ethnicity breakdown was 24% Asian, 9 % Black/African American, 12% Hispanic/Latino, 5% Pacific Islander, 47% White/Caucasian, and 2% Other. The study was performed at a southwest educational institution. The participants were selected from an introductory psychology course, given informed consent, and earned course credit for participation. A subset of participants returned two weeks later and were re-tested with the same questionnaire (n = 36).

Missing data was accounted for through case deletion as the authors identified no theoretical reasons for data to be systematically missing.

### *Measures*

#### *RD/RG Scale*

The current RD/RG measure was designed to measure participants' perception of their comparisons. Items were created intended to assess participant perception in reference to the selected comparison groups along a given object of comparison. This was intended as a measure of the magnitude of the comparisons.

The format of the questions was “How do you view yourself compared to \_\_\_\_\_ on \_\_\_\_\_?” The range of answers was on a 7-point ranging from -3 (*doing worse*) and 3 (*doing better*). Sum scores can be created for RD by rescoring all positive responses to zero and RG by rescoring all negative responses to zero. A sum score, representing overall consequence of comparison can be derived from a total sum score.

Scale questions were designed around a two-part format for every comparison group by every object of comparison. The first is a measure of magnitude of disparity (e.g. *How are you doing in terms of money compared to friends?*). The second is a measure of intensity of the disparity (e.g. *How do you feel about this?*). The item “How do you feel about this” followed each comparison item, also on a 7-point scale with -3 equating to *satisfied* and 3 equating to *dissatisfied*. The format measured the perceived distress from comparisons consistent with the theories of Runciman (1966) and Cook et al (1977) notion of magnitude and intensity.

To assess concurrent validity, the RD/RG questionnaire was administered with measures of related phenomenon hypothesized to be related. One such relationship is that greater inequality negatively affects health (Eibner & Evans, 2005, Wilkinson, 1996). Given that RD is predicted to correlate with impaired health, the measure that better correlates to health measures is hypothesized to be the better measure of RD. RG is also measured here, yet there are few studies assessing RG for comparison and there is as yet no hypothesized relationship between RG and health, though RG may be employed to offset the distress created by RD (Runciman, 1966). RG may act like the opposite of RD and therefore the relationship between RG and health would be positive. However, previous research has found that the RG does not necessarily cause the opposite effect of RD (Dambrum et al., 2006).

### *General Health*

Single item self-report measures of health have consistently been found to predict disease and mortality (Idler & Benyamini, 1997). Therefore, participants were asked: “On a scale of one to ten please rate your overall health” with answers ranging from 1 (*Very poor health*) to 10 (*Excellent health*). The self-rated health question assesses the overall array of illness and could also be incorporating symptoms of undiagnosed diseases (Idler & Benyamini, 1997). This can also be interpreted as self-ratings of health are representative of human judgments of the severity of the current illness. Objective measures of health will not capture the severity of symptoms as well as self-reports. Also, the effects of interactions between conditions are not reflected by the additive measure of chronic disease. The authors noted that they had found no study that had tested the association between self-rated health and interactions of conditions. The self-rated health

question also reflects family history. The authors noted that three studies did include measures of family history or longevity and that two of these studies showed an independent effect of self-ratings when this family history was taken into account. Another consideration made by the authors is self-rated health functions as a dynamic evaluation capable of incorporating trajectory of health as well as current level. Still another benefit to self-rated health is the implication that reverse-causation is also being measured. Self-rated health may affect behaviors that affect health. Poor perceptions may lead to less willingness to engage in self-care, adhering to screenings, medications, and treatment, and preventative practices while good perceptions would increase willingness to participate in health promotion. Lastly, self-rated health may also reflect within-person resources. Overall it appears that self-rated health may be an efficient rating of one's health, incorporating stress and coping ability.

### *Mental Health*

The Hopkins Symptom Checklist (HSCL) is a self-report symptom rating, which has been used since the 1950's when it was titled the Discomfort Scale (Parloff et al., 1954). This scale was an adaptation itself from the Cornell Medical Index (Wider, 1948). The scale measured symptoms along five dimensions- Anxiety, Depression, Interpersonal Sensitivity, Obsessive-Compulsive, and Somatization. Coefficient alphas for each dimension were high, ranging from .84 to .87. All item-total correlations were calculated and found to contribute substantially to each dimension (all were above .50, most were near .70). Test-retest reliability was evaluated over a one-week period and was found to be high (values ranged from .75 to .82 for the five dimensions). Correlations between dimensions ranged from .10 to .44

The HSCL has been used to measure a variety of mental phenomenon including PTSD (Glaesser et al, 2004), panic attacks (Norton et al, 1985), and racial identity (Kibour, 2001). The HSCL has different versions that differ by the number of items. Versions range from the 90-item form to the 15-item form depending on how many dimensions the experimenter wishes to assess. The HSCL-30, for example, has thirty items and measures somatization, depression, and anxiety. The 30-item version of the scale is designed to assess symptoms of stress using a 5-point Likert scale.

### *Belief in a Just World*

RD is a justice theory in that the construct assumes an underlying motivation for justice in the world (Lerner, 2003). A similar justice theory is *belief in a just world* (BJW). BJW research follows people's desire to believe that people get what they deserve (Lerner, 1965). BJW theory holds that people develop general schemata that create general theories about reality. If one's theories suggest that the world is a just place then one has a higher BJW. This belief allows individuals to approach the world as if it is stable and orderly. Belief in a just world has been tied to mental health. Dalbert (1999) argued that BJW had a positive impact on mental health by creating feelings of competence and control, fostering the notion that good deeds will go rewarded (thereby allowing people to invest in their own good future), and fostering investment in long-term goals. By creating a stable and positive way to view the world, BJW enables people to feel that they have control over their lives.

Like RD, BJW has undergone refinement and has been divided into a general belief in a just world and a personal one (Lerner & Miller, 1978). Personal BJW centers on the individual and is the belief that one's own fate is just, that one's own local

environment is fair and just, that life is fair for oneself and one's own family and friends. General BJW is focused on the world at large and is the belief that the world in general is fair and just for all people. Lipkus et al. (1996) compared personal BJW and general BJW and found that personal BJW was a better predictor of life satisfaction, stress, and depression than general BJW. Based on these findings, personal BJW is likely a better correlate to RD.

Dalbert (1999) tested if his five-item General BJW scale and seven-item Personal BJW were fundamentally different by administering both scales and running exploratory factor analysis (EFA) and finding a two-factor structure emerging that distinguished between items designed to measure general BJW and personal BJW. Dalbert also found that mental health was more influenced by personal BJW than general BJW. Personal BJW correlated highly with mood level, life satisfaction, and self-esteem. No sex differences were found for BJW scores. Personal BJW was correlated statistically significantly stronger with life satisfaction and slightly stronger with self-esteem than was general BJW. Given that the current research seeks to develop a scale that measures egoist deprivation and is predicted to relate to health, a measure of RD is predicted to have a statistically significant negative correlation with the personal BJW scale.

### *Income*

Absolute income is hypothesized to be distinctly different than RD. Therefore; feelings of deprivation should not be highly related to income. Income was assessed through participant reports of estimated yearly income before taxes. Given that our sample will draw from subject pool- typically a college age group- the question was designed to ask the participant to also specify parents estimated income if they still rely

on their parents for financial support. Responses were predicted to have a low non-significant positive correlation with a measure of RD/RG.

## Results

The findings are presented here in the order of the hypotheses. First an EFA was conducted to identify an underlying factor structure. Second, magnitude and intensity were evaluated to identify if the constructs, as they were measured in the present study, were statistically different. Then RD and RG were calculated from the data and validation was tested through correlation analysis to selected measures that are hypothesized to correlate strongly to a measure of RD and RG (BJW, HSCL-30, self-reported health) as well as a measure that is not hypothesized to correlate with the proposed measures (income). Lastly, the capacity of RD and RG to predict health was tested through regression analysis.

### *Reliability*

*Hypothesis: Magnitude and intensity will have sufficient reliability as tested by Cronbach's alpha and test-retest reliability.*

Reliability scores in the current study were calculated for the magnitude and intensity items of the perceived inequity scale to assess the psychometric fitness of our scale. The magnitude and intensity inequity reliabilities were both  $\alpha = .94$ . A subset of the sample completed the survey again three weeks later ( $n = 36$ ). These test-retest reliabilities were  $r = .67$  for magnitude ratings and  $r = .74$  for intensity ratings, all statistically significant ( $p < .001$ ). The hypothesis was supported in that both magnitude and intensity ratings had sufficient reliability.

### *Exploratory Factor Analysis*

*Hypothesis: Magnitude and intensity items will emerge as unique dimensions as tested through exploratory factor analysis (EFA), independent t tests, and correlation.*

EFA was performed to identify the factor structure of the measure (n = 246, 154 female). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was .91 indicating that a sufficient sample for performing an EFA was obtained. Bartlett's test of sphericity was statistically significant ( $p < .001$ ), suggestive of high inter-item correlation.

The extracted communalities of the model ranged from .53 (intensity response to accomplishment comparisons to members of the same gender) to .87 (magnitude response to appearance comparisons with previous self). Fifteen factors emerged as having eigenvalues over 1, while scree plot inspection suggested an 11 factor model. The initial eigenvalue cumulative percent was 71% for the 11-factor model and 77% for the 15-factor model. The 15-factor model was selected for inspection. Assessment of the rotated factor matrix revealed consistent loadings on object of comparison. Both magnitude and intensity for a given object of comparison items would load on a single factor that was unique from other objects.

The comparison group, it seemed, did not load into factors with two notable exceptions. The items measuring comparisons on a general overall object did not load into a single factor but loaded across three, with no discernable pattern therein. Additionally, all items measuring comparisons for one's previous self loaded in unique factors independent of the other comparisons for the given object. These loadings were unique to the comparison; previous-self comparisons on finances loaded onto a different factor than previous-self comparisons on appearances. Magnitude and intensity loadings,

however, generally loaded into the same factor including the loadings for previous self. The factor loading suggest that the items loaded into groups by comparison object, except in those comparisons concerning one's previous self, which loaded independently. One exception to this was items measuring the general object of comparison. While intensity ratings loaded within a single factor for the general object, the magnitude ratings were dispersed across three factors. No items loaded below .3; the loadings ranged from .48 (intensity rating of financial comparisons with the same gender and also magnitude rating of looks with celebrities/athletes) to .86 (intensity rating of financial comparisons with previous-self). The factor transformation matrix off-diagonal elements ranged from .007 to .62, indicating that large rotation was applied.

No sex differences were identified when EFAs were performed separately for males than females. The test of sampling adequacy for males was low (KMO = .63, versus for females KMO = .81), but otherwise the trends in the data were identical for males and females.

The hypothesis was not supported because intensity and magnitude ratings did not load separately into different factors, suggesting high inter-correlation between magnitude and intensity.

### *Magnitude and Intensity*

Mean and variance scores for all scales and mathematically derived subscales are presented in Table 3. There was a statistically significant difference between magnitude and intensity scores ( $t(238) = 2.02, p = .045, d = .10$ ) with intensity scores statistically significantly higher than magnitude scores. Magnitude and intensity had a correlation of  $r = .80, p < .01$ . Multicollinearity was estimated through tolerance and variance inflation

factor (VIF). The tolerance was .36 and VIF was 2.7. Based on these values, analysis proceeded with the assumption that the magnitude and intensity measures were highly related but different constructs.

### *RD and RG*

Hypothesis: *Magnitude and intensity dimensions of both RD and RG will not have statistically significant sex or ethnicity differences as tested through analysis of variance (ANOVA)*

RD and RG values were derived from the magnitude and intensity of inequity scores to evaluate the association between these constructs and the selected health measures. The scale design allowed for three separate constructs to be compared: RD, RG, and the cumulative of these two constructs (termed “consequence of comparisons”). RD scores were calculated by first setting all scores above “0” to zero (these values are indicative of RG), then the sum of participants’ responses to the magnitude items and the sum of the responses to the intensity items were calculated. A similar procedure was carried out for RG. The total score designed to assess overall social consequence was derived through subtracting the total RD value from the total RG for each participant. Because the consequence of comparison is comprised of RD and RG, these three scales are not independent. The RD mean for magnitude was -10.5 ( $SD = 10.5$ ) and for intensity was -9.3 ( $SD = 11.1$ ). The RG magnitude rating had a mean of 44.0 ( $SD = 22.3$ ) and the intensity rating had a mean of 46.6 ( $SD = 25.8$ ). Table 3 displays the descriptive statistics for the inequity scales, B JW, income, self-reported health and HSCL-30.

Among all scales and subscales no significant differences were found between females and males with the exception of the HSCL-30 scale, in which the females

reported statistically significantly higher scores ( $F(1,244) = 5.53, p = .02, \eta^2 = .05$ ). No significant age differences were found between the RD/RG magnitude/intensity subscales.

Differences in ethnicity, however, were found for all of the social comparison subscales: RG Magnitude ( $F(5,228) = 5.37, p < .001, \eta^2 = .11$ ), RD Magnitude ( $F(5,228) = 3.78, p = .003, \eta^2 = .08$ ), RG Intensity ( $F(5,220) = 3.05, p = .01, \eta^2 = .06$ ), and RD Intensity ( $F(5,220) = 2.38, p = .04, \eta^2 = .05$ ). A Tukey HSD post hoc was conducted finding the greatest and only significant mean difference concerned self-reported Asians to other ethnic groups. For RG Magnitude, the lowest scores were in the Asian group ( $M = 34.29, SD = 21.76$ ). This was significantly lower than African Americans ( $M = 50.42, SD = 20.33; p = .034$ ) and Caucasians ( $M = 48.68, SD = 20.35, p = .001$ ). For RD Magnitude, the lowest scores were in the Asian group ( $M = -13.59, SD = 11.36$ ). This was significantly lower ( $p = .033$ ) than the African Americans ( $M = -5.90, SD = 8.09$ ). For RG Intensity the lowest scores were in the Asian group ( $M = 39.13, SD = 23.19$ ). This was significantly lower ( $p = .043$ ) than Caucasians ( $M = 51.28, SD = 24.98$ ). The RD Intensity had no statistically significant differences found in Tukey HSD. A more liberal post hoc test (LSD) was performed to identify where mean differences were (as the ANOVA was statistically significant). The LSD post hoc test identified a mean difference ( $p = .037$ ) between Asian ( $M = -11.20, SD = 11.52$ ) and Caucasian ( $M = -7.50, SD = 8.95$ ). The self-reported Asians consistently reported greater RD and less RG than other races. The hypothesis was not supported in that while age differences were not found, ethnicity differences were found in self-reported perceived inequity.

#### *Validation*

Hypothesis: *Magnitude and intensity dimensions of both RD and RG will meet construct validity as tested by correlation with related variables as well as constructs predicted not to be related to the RD and RG.*

As an evaluation of the validity of the current measure, correlational analyses were performed (see Table 4). Magnitude and intensity ratings were statistically significantly positively correlated to BJW, the general health question, and the HSCL-30. As ratings of inequity moved from feeling deprived to feeling gratified, so too did belief in a just world, overall self-reported health, and mental health. Furthermore, RD and RG for both magnitude and intensity were statistically significantly correlated ( $p < .01$ ) to BJW and both health measures. No measure was statistically significantly correlated to income except for magnitude RG ( $r = .14, p < .05$ ). The hypothesis was supported as the scales assumed to measure RD and RG for both magnitude and intensity were correlated to the measures predicted to relate to RD and RG (BJW, self-reported health, HSCL-30) and did not correlate to income.

Hypothesis: *Magnitude and intensity dimensions of both RD and RG will differently predict self-reported health measures as tested through linear multiple regression.*

Regression analysis was performed to evaluate the degree to which RD and RG in both magnitude and intensity ratings were predictive of the health measures of self-reported health and mental health (measured through HSCL-30). Four variables were entered for prediction for the two health models: 1) RD magnitude, 2) RD intensity, 3) RG magnitude, 4) RG intensity. For self-reported health, only the RG magnitude and intensity scores emerged as being significantly predictive. The RD variance overlapped with the RG such that only RG intensity emerged as a significant predictor. The final

predictive model of general health emerged using the variables RG magnitude and RG intensity ( $R^2 = .19$ ; RG magnitude  $b = .013$ ,  $SE = .006$ ,  $t = 2.32$ ,  $p = .02$ ; RG intensity  $B = .013$ ,  $SE = .005$ ,  $t = 2.694$ ,  $p = .08$ ). In the prediction of the HSCL-30 measure of mental health, only RD intensity emerged as a significant contributor ( $R^2 = .17$ ; RD intensity  $B = -.15$ ,  $SE = .02$ ,  $t = -7.03$ ,  $p < .001$ ). The hypothesis was partially supported in that at least one component of the inequity measures: RD or RG of magnitude and intensity were predictive of health.

### Conclusion

This study was performed to test a social comparison-based measure of RD and RG and test the relationship of these constructs to related phenomenon. From economics (Yitzhaki, 1979) to mental health policy development (Eibner, Sturm, & Gresenz, 2004), RD is a construct with profound implications for social scientists in viewing human interaction and therefore requires careful deliberation about the appropriate method of measurement. A review of RD measures suggested that RD is hypothesized to be stressing- even if not explicitly stated. Measures of RD are used to study behaviors ranging from anger over income to racism and mortality but the construct has also shifted in interpretation. For example, Stouffer's (1949) original study involved participants' dissatisfaction while Crosby's (1982) measure attempted to assess the respondents' bitterness. Pedersen (2004) used relative deprivation synonymously with envy. Despite the varying interpretations, each has an assumption that RD is believed to create a level of distress. Runciman's (1966) definition does not specifically state "distress" but can be viewed as a consequence of the conflict between the individual wanting X, feeling that it is feasible that they have X and not actually having X. Though not specifically outlined

by Runciman's (1966) definition of RD, a perceived level of distress is inherent to the existence of RD. The importance of assessing personal distress was made years ago and continues into current critiques of RD research (Crosby, 1979; Smith & Leach, 2004).

The current measure was found to be reliable in terms of inter-item correlation but low in test-retest reliability. The low test-retest reliability may be due to the relatively small sample size. Another explanation for the low reliability is that the value is that perceived inequity fluctuates among college students. Because the proposed measure had not been previously tested, an EFA was first performed to identify the underlying factor structure. The KMO and Bartlett's test both indicated that the data was sufficient in sample size and inter-item correlation to perform an EFA. Principal axis factoring was performed since the purpose of the EFA was to identify an underlying structure. Promax rotation was selected because the factors were predicted to be oblique to one another. No consistent factor structure emerged that corresponded to comparison groups, objects, magnitude, or intensity. The results indicated a relative uni-dimensionality of the items, suggestive that a single construct is being measured. However, the data also supported the possibility that objects of comparison were unique factors while comparison groups were not. This may be a testing artifact as the items were grouped by comparison object. Randomizing items in subsequent analysis should remove this effect if it was an artifact of the item groupings.

The second hypothesis tested was the proposed difference between magnitude and intensity ratings. The EFA did not identify distinct factors corresponding to magnitude and intensity items and although a statistically significant difference was found between the two means, the low effect size suggested this difference was due to a large sample

size. Furthermore, a high positive correlation was found between magnitude and intensity. Due to the measurement approach used in this study, the two dimensions of intensity and magnitude are assumed to correlate but this correlation was so strong they may be the same phenomenon. To explore the multicollinearity of the hypothesized dimensions the tolerance and VIF tests were performed and though intensity and magnitude were highly correlated, they were below the cutoffs for multicollinearity. Therefore, the magnitude and intensity were considered as highly related but independent.

The data were transformed to reflect RD and RG dimensions of both magnitude and intensity. Magnitude and intensity scores were similar and virtually indistinguishable. This is not to say that magnitude and intensity as constructs are not different but the method of measurement developed here, by nature of the self-report design, may combine both aspects. When the participants rate how they are doing in relation to another person along a given dimension they are already adjusting for intensity when they give their ratings. In other words, people are not giving a truly objective non-emotional rating of their position in each item but appear to be producing a conclusion based upon these comparisons. As such, the intensity dimension of this scale may be superfluous.

Ethnicity differences were found for the RD/RG scores. Post hoc analyses revealed the significant differences were from scores between self-reported Asian ethnicity against African Americans and/or Caucasians. Asians scores of RD Magnitude and RG Magnitude were significantly lower compared to African Americans and Asians scores of RG Magnitude and RG Intensity were significantly lower compared to Caucasians. This suggests that there may be cultural variations in perceived inequity

and/or cultural differences in the *reporting* of inequity. Further research that does not employ a self-report design would illustrate this effect further.

The self-report measure of inequity derived from prominent groups and objects of comparison yielded a measure that was evaluated as high in construct validity due to significant correlations with theoretically related constructs (BJW, general self-reported health, and mental health) while also showing a low correlation with a theoretically unrelated construct (income). Furthermore, components of the scale expected to measure the magnitude and intensity of the perceived inequity were independent from one another. EFA suggested a unidimensional construct.

RD and RG seem to account for the variations in health differently. In terms of self-reported health, only one's RG- overall amount of perceived superiority was contributive. Conversely, RD seemed predictive of mental health. A person's perceived gratifications- one's successes compared to others- was related to positive health. Perceived disparity seemed related towards mental health. While causation conclusions should be discussed with caution, this finding suggests that RG may be therapeutic to overall health while RD may be contributive to mental disorders.

Effect size estimation was  $d = .10$  (a dependent t-test, independent means and standard deviations were used for effect size estimation, consistent with Dunlop, Cortina, Vaslow, and Burke, 1996). This low effect size suggested that the statistical significance was due more to the large sample size. As such, the means were not considered to be statistically significantly different.

A limitation of this study was the method of measuring prominent comparison groups. Self-report was used to identify such groups and as such at least two

measurement issues arise. The first is the dependence on conscious comparisons. All of the groups and objects selected for the RD/RG measure were those highly rated by participants as commonly occurring. Comparisons that the participant is not aware are not captured. The second concern is the assumption that participants had at least a cursory understanding of social comparison processes. This was not always the case, such as one participant's response to the open-ended comparison group question that "I don't compare myself to others, I'm my own person". Being independent and looking to others to get a sense of standing are not mutually exclusive. Other researchers have explored the importance of social comparison to the individual (e.g. Buunk & Gibbons, 1999). The impact of these concerns is difficult to assess from the present data. One could argue that these are limitations create a shallow assessment of RD/RG. Alternately, perhaps RD and RG are only derived from conscious and conspicuous comparisons. A third concern was the sample, which consisted of college students at a southwestern U.S. university. The sample may provide a limited range of variation of perceived inequity. As the purpose of the present study was the validation of this measurement approach and to explore possible health effects, future research should expand on the health measures as well as increase the demographic variability of the sample.

## CHAPTER 5

### STUDY 3: IDENTIFICATION OF PROMINENT COMPARISON GROUPS AND OBJECTS FROM AN EXPANDED POPULATION

Both studies in experiment one were exploratory studies with the aim of identifying and validating a measure of perceived inequity. Study one attempted to identify prominent comparison groups and objects in the population of interest. This was accomplished using the Prominent Comparison Groups and Objects (PCGO), a measure by which participants rate their perceived importance of supplied groups along given objects of comparison found in previous research to be prominent. Additionally, the PCGO includes sections for participants to add groups or objects that they find meaningful. After the data from administering the PCGO were analyzed, items intended to measure perceived inequity from these groups/objects were created. The scale was validated through strong correlations with related phenomenon (e.g., belief in a just world) as well as low correlations with phenomenon predicted to be unrelated (e.g., income). The scores of perceived inequity were also calculated to estimate relative deprivation (RD) and relative gratification (RG) in both magnitude and intensity. With a reasonable support for a valid measure, the scale was compared to indicators of health. A strong relationship was found between perceived relative social position and health ratings.

The sample, however, consisted of psychology students enrolled in a southwestern U.S. university and they participated in the study for course credit. Furthermore, the health ratings were all self-reported. The perceived inequity scale measured both magnitude and inequity. This measure was lengthy and test fatigue may

have influenced responses. The overarching purposes of experiment two were to increase the demographic variability of the sample and to replicate the previously identified relationship between the developed scale and health.

Study one and two was conducted to evaluate the feasibility of developing a reliable and valid self-reported measure of inequity that can predict health. Both studies sampled from a state university and as such generalizing the findings of social comparison effects to a population beyond college students may not be accurate. The purpose of this study and the subsequent study are to evaluate the measurement technique and effects identified in studies one and two with a more diverse population.

The same scale that was used in study one was used here. The *a priori* comparison groups tested were: coworkers, family members, friends, members of the same religion, previous-self, and fictional characters. The *a priori* objects of comparison tested were: general, financial, and interpersonal.

#### Hypotheses:

- The proposed comparison groups will be statistically highly rated as tested through one sample t-tests
- Open-ended items will identify additional prominent comparison groups. These groups will be identified through two tests: the proposed group being reported in 2% of the total sample and the group must be theoretically likely to create both RD and RG.
- The proposed comparison objects will be statistically highly rated as tested through one sample t-tests

- Open-ended items will identify additional prominent comparison objects as will be assessed through two tests: the proposed group being reported in 2% of the total sample and the group must be theoretically likely to create both RD and RG.

## Method

### *Participants*

One hundred and eighty participants participated (76 female). Incomplete questionnaires were removed, resulting in 154 participants (68 female) with a mean age of 34 and median age of 31. Prominent ethnic groups of the sample included Caucasian (51%), Asian (4%), Hispanic/Latino (26%), and African American (16 %) with the remainder self-reported “other” (3%).

### *Procedure*

Participants were recruited from two Department of Motor Vehicles (DMV) locations in Las Vegas, Nevada. Local DMVs were selected because the general population of U.S. citizens visits the DMV; even those without a driver’s license get state IDs at a DMV. Additionally, patrons of the DMV often wait long periods of time in an indoor environment and that facilitates test taking. Participants were recruited at a booth set up at the entrance with a sign soliciting participation (the sign stated “Bored? Take a short study. Return it when you leave. For UNLV. Free water”). For participating, participants received their choice of either bottled water or hot chocolate, depending on the weather.

### *Measure*

Participants were given the Prominent Comparison Groups and Objects scale (PCGO). The PCGO contained a list of comparison groups in which participants indicated the importance they ascribed to each group. As per the findings in Study one, the instructions for the PCGO were revised to increase facilitation of admitting to social comparison. The following instructions were provided at the beginning of the questionnaire:

Sometimes we compare ourselves to other people. We compare ourselves to others to get an idea of how well we're doing. Sometimes we also compare ourselves to how we had it previously- our previous selves. Below is a list of groups of people you may compare yourself to. The purpose of this test is to measure which groups you compare yourself to and how important each comparison is to you. Please read each question and determine for each group of people how strongly you compare yourself to them. Thank you.

Also as per the findings of study one, open-ended questioning identifying prominent comparison groups and objects was immediately provided before any of the researcher-provided groups were offered. For the open-ended items, the following instructions were provided: "Think about how well you are doing generally. Remember, we compare ourselves to others to get an idea of how well we're doing." This was followed by the item asking "We asked you to compare yourself on how well you were doing generally. What was it that you compared yourself to others to?" along with space for the participant's response. To measure comparison groups, the next item was phrased "Who was it that you compared yourself to?"

Comparison groups were then provided and participants rated the degree to which they believed that they compare themselves to each group on a 0-5 scale (with 0 indicating “no comparison” and 5 indicating “heavy comparison”). For example, the item for the comparison group “coworkers” and the “general” object of comparison stated, “*To what degree did you compare yourself to your **coworkers?***”

The comparison groups were coworkers, family members, friends, members of the same religion, previous-self, and famous people (the previous implementation of the PCGO in study one had used a slightly different term, “fictional characters”, described in the questionnaire as “TV, movie, or book characters”). The scale was designed in such a way so as to present each comparison group within each object of comparison. Participants were given the set of comparison groups to rate for the “general” object of comparison, then the “financial” object, followed by the “interpersonal” object of comparison.

## Results

### *Analyses*

As with study one, analysis was performed by analyzing the values derived from the close-ended items measuring the provided comparison groups and interpreting the answers supplied for the open-ended items requesting additional comparisons. For the first set of data, evaluation of the provided comparison groups and objects of comparison were based upon statistical criteria. The objects of comparison mean scores were derived from calculating the means of each comparison group within each object of comparison. For example, the mean score for the general object of comparison was derived from the mean of all ratings of the six comparison groups within the “general” object. The ratings

scale ranged from 0-5. A one-sample t-test was performed to identify if mean ratings were greater than a rating of 2. The cutoff mean of 2 was selected because this value is less than half of the scale and corresponds to the response of “I care a little about my comparison to this group”. Any comparison group with a mean less than half would not be a popular enough candidate for an overall measure of prominent groups. To avoid inflating type-1 error by conducting this analysis for every comparison group by every object of comparison, comparison groups were collapsed by comparison object to avoid excess analyses. Any non-significant comparison group was analyzed individually to evaluate if the comparison group was significant by any of the objects of comparison. Gender differences were also investigated in the importance of selected groups and objects of comparison with analysis of variance (ANOVA). Age effects were also assessed through ANOVA (age was categorized as 18-24, 25-34, 35-49, and 50 and over).

Evaluation of the open-ended responses was conducted using both quantitative and qualitative analyses. As with study one, the answers were categorized. Then, in terms of quantitative evaluation, any new object or group that was supplied by less than 2% of the participants was excluded from further analysis. Besides meeting mathematical criteria, the objects and comparison groups provided by the open-ended responses were required to be theoretically plausible groups or objects that could create status identification. As with the quantitative data, gender and age differences were also investigated in the importance of selected groups and objects of comparison.

*Selected Comparison Groups:*

Hypothesis: *The proposed comparison groups will be statistically highly rated as tested through one sample t-tests*

Similar to study one, two comparison groups were not statistically significantly greater than 2 (see Table 5). Members of the same religion and fictional characters had means below 2 for all of the objects of comparison; general, financial, and interpersonal. There was no statistically significant difference between males and females between comparison group ratings (see Table 6). Age effects were found for the rating of one comparison group, members of the same religion (see Table 8). Those over 49 gave a statistically significantly higher rating for members of the same religion than those aged 18-24 and 25-34. The hypothesis was supported for all of the comparison groups except for members of the same religion and fictional characters.

*Selected Objects of Comparison:*

Hypothesis: *The proposed comparison objects will be statistically highly rated as tested through one sample t-tests*

One-sample t-tests comparing the ratings of the objects of comparison to the cutoff of 2 indicated that the means of all three objects (general, financial, and interpersonal) were significantly greater than the cutoff. The effect sizes for general and interpersonal comparisons were medium (above .30). The effect size for the financial object of comparison was .22, making the cutoff qualify as a small effect.

Gender differences were examined by ANOVA. None of the objects of comparison were statistically significantly different (see Table 6). Age similarly did not show a statistically significant difference by object of comparison. The hypothesis was

supported in that all three predicted objects of comparison (general, financial, and interpersonal) were statistically significantly highly rated.

*Identification of Additional Comparison Groups:*

*Hypothesis: Open-ended items will identify additional prominent comparison groups.*

*These groups will be identified through two tests: the proposed group being reported in 2% of the total sample and the group must be theoretically likely to create both RD and RG.*

One hundred and fifty-four participants (97%) identified additional comparison groups beyond those given in the questionnaire. Any response that was given by more than two percent of the entire sample ( $n = 4$ ) was a candidate for inclusion. Despite the large portion of the sample that chose to provide additional comparison groups, after coding the responses, only one additional group emerged, the general comparison group ( $n = 13$ ). Many other potential comparison groups were provided, however, there were few repetitions in the responses (such as “Barry White, which was re-coded as “famous people” and unrelated answers such as “the economy <is not good>”) or responses reproducing an existing group (such as “friends”, which was listed by twenty participants) or previous-self ( $n = 11$ ). Several groups that were popular in Study one were near but below the cutoff and were not eligible for consideration, such as the groups “members of the same sex” ( $n = 3$ ) and “famous/fictional characters” ( $n = 2$ ). The hypothesis was supported in that new groups were identified for consideration such as the “general” comparison group.

*Identification of Additional Objects of Comparison:*

Hypothesis: *Open-ended items will identify additional prominent comparison objects as will be assessed through two tests: the proposed group being reported in 2% of the total sample and the group must be theoretically likely to create both RD and RG.*

One-hundred and forty-one participants (89%) identified additional objects of comparison. Open-ended questioning was used to evaluate the identified objects (general, financial, and interpersonal) and identify additional objects of comparison. Responses were categorized by similarity and evaluated using the 2% frequency cutoff. For example, the statements “wealth” and “possessions” were coded as “financial”. The responses that were similar to those previously provided to the participants were also recorded and all three were identified by more than 2% of the participants (general, n = 11; financial, n = 45; interpersonal, n = 10). However, two additional object of comparison provided by the participants emerged as prominent, “health/looks” (n = 12), and “knowledge” (n = 13). The hypothesis was supported in that several new comparison objects were identified such as knowledge, health and looks.

### Conclusion

Identification of prominent comparison groups and objects is the first step of the two-step approach outlined above for the measurement of perceived deprivation. While meaningful groups and objects were previously identified in study one, several problems with the prior research were corrected: The instructions were revised to increase acknowledgement of engaging in social comparison, the open-ended items were presented before the close-ended items to reduce bias, and the sample was broadened to more accurately reflect the U.S. population. The revised questionnaire was presented to patrons of two DMVs in Southwestern United States, which increased the variance of

the sample. Results were similar to those found in study one but with fewer new comparison groups or objects provided by the participants.

The primary purpose of this study was to verify the findings of study one, with a sample with greater age and ethnic distribution. Close-ended and open-ended questioning was used to identify prominent comparison groups and objects. For comparison groups, the prominent comparison groups provided by the researchers were coworkers, family members, friends, previous-self, members of the same religion, and famous people (changed from “fictional characters” in the previous iteration of the PCGO). All of the provided comparison groups were supported as being prominent as tested through one-sample t-tests, except for members of the same religion and fictional characters. No gender differences were found in that both females and males agreed on which groups were meaningful. Because of the expanded age range, analysis was conducted to identify if the prominent comparison group ratings differed by age category. An age effect was identified for members of the same religion. Specifically, those over the age of forty-nine had statistically significantly greater means than those that were younger (18-24 and 25-34 but not 35-49). This may have been a statistical artifact caused by the somewhat small sample size of those over the age of forty-nine (20). However, it is also reasonable to assume that such comparisons are more meaningful to that group in the sample. As the overall mean did not meet the cutoff for consideration, this comparison group did not appear to be prominent enough to include in the measure.

In terms of additional comparison groups, only the “general other people” comparison group again emerged as prominent from the open-ended items seeking

additional groups. Unlike the previous study, however, no other groups emerged as prominent. Assessment of the open-ended responses indicated that a possible reason for this was that participants reported comparison groups that had already been selected by the researchers.

All of the provided comparison objects were found to be statistically significant (general, financial, and interpersonal). Like the identification of additional comparison groups, fewer additional objects emerged in this study compared to study one. Participants included “health” and “knowledge” as additional meaningful objects of comparison. This is interesting because the “health” object is similar to the “looks/appearances” object found with the predominately young sample in study one, but the use of the word “health” suggests that with increased age comes the increased emphasis on functioning of the body, not just appearances. Additionally, the “knowledge” object is similar to the “accomplishments” group that emerged in study one. Also, while in study one the “financial” object of comparison was not statistically significant among males; the males in the expanded age pool did report financial comparisons as being important.

While this study sought to overcome some of the weaknesses identified in study one, there were some limitations. The self-report nature of this measurement approach produces comparisons the participant *thinks* are important and are willing to disclose. Participants may not be aware of social comparisons (Wood, 1996). However, the purpose of this study was to identify if a self-report measure of social comparison can predict health. As such, a self-report approach to the identification of prominent comparisons is practical for the current test question.

There were considerably fewer comparison groups and objects provided by this sample than the sample of college students in study one. This may be indicative of test fatigue or not understanding the question. Additionally, this may be influenced by exposure to college courses that may have presented concepts such as social comparison. All of the participants in study one had exposure to at least one psychology course (by virtue of their participation for course credit). The present sample, however, may not have had any exposure to higher education. The participants' education was not assessed in the present study. This should be investigated in subsequent research because as education may relate to one's ability to identify with abstract concepts such as social comparison. Another reason for the reduction in participant-provided comparisons may be the greater demographic variability in the current sample. By expanding the population, the variability in comparisons may have been greater than the statistical cutoffs. For example, while celebrities emerged as a prominent comparison group in study one with college students, this group was not found to be prominent in the sample of the general population patrons of the DMV. . Perhaps celebrities are an important comparison group to young college students. As such, the final measure for study two may be too broad to validly measure participants' prominent comparison groups and may reduce the strength between prominent comparisons and health.

This research identified prevalent comparison groups and objects of comparison, resulting in the identification of five comparison groups: friends, family members, coworkers, previous self, and general, and four objects of comparison: general, financial, interpersonal, and health. The present research established prominent comparison groups

and objects useful for the development of a general scale to measure relative disparity (RD) and relative gratification (RG).

## CHAPTER 6

### STUDY 4: MEASUREMENT OF RD/RG IN THE GENERAL POPULATION AND THE ROLE OF SELF-REPORTED INEQUITY IN HEALTH PREDICTION

The purpose of this study was to develop a measure of RD and RG derived from the prominent comparison groups and objects identified in the previous study. These comparison groups were: general, friends, coworkers, family members, previous self; as well as five objects of comparison: general, financial, interpersonal, health, and knowledge. In study two, a relationship was found between perceived inequity and health, but may have been minimized by the youthfulness of that sample.

The current RD/RG measure was based upon self-reported inequity of a specified reference group in a specified condition (e.g. comparisons to coworkers on emotional experiences). This was intended as a measure of magnitude. This alone may be sufficient as a measure of RD/RG because though it is an objectively phrased item, participants would likely not rate their comparison objectively (supported by the finding in study two that the magnitude and intensity self-report measures were statistically near identical). If participants believed they were doing worse than someone and were distressed by the disparity, they may answer this item with a lower value than if another participant felt the same disparity was not distressed about the situation. This item measures perceived disparity similar to the local area index measures of RD and the condition-based approach, but is rated by the participant instead of derived from data.

Hypotheses:

- The magnitude ratings of perceived inequity will have sufficient reliability as tested by Cronbach's alpha

- The site types will differ significantly by personal income but not perceived inequity
- The magnitude dimensions of perceived inequity will meet construct validity as tested by correlation with related variables as well as constructs that are expected to be unrelated. Specifically, perceived inequity will be statistically significantly correlated to perceived control, perceived stress, and self-reported status but not statistically significantly correlated to income.
- The magnitude dimension of inequity will not have statistically significant sex or ethnicity differences as tested through analysis of variance (ANOVA).
- The magnitude dimension of inequity will have statistically significant differences in age and education as tested through ANOVA.
- The magnitude dimensions of both RD and RG will differently predict self-reported health measures as tested through linear multiple regression and contribute unique variance to the prediction of health in addition to perceived stress and perceived control.

## Method

### *Participants*

Two-hundred and ninety nine patrons of two DMVs of a city in Southwestern United States participated in the study. Incomplete surveys were addressed by case deletion. Also, a “check” question was included to verify participation interest (the item was “As a friendly check to make sure you’re reading every question, please circle “2”

for this question). Incomplete surveys and those participants not correctly answering the check item were removed from the sample, resulting in 267 participants (104 female). Ages ranged from 18 to 77 (mean of 31 years). Prominent ethnic backgrounds included White/Caucasian (40%), Hispanic/Latino (24%), Black/African American (20%), Asian/Pacific Islander (11%), and Native American (2%). For analysis purposes, the Native American group was collapsed into “Other” (5%).

### *Procedure*

Participants were recruited from two Department of Motor Vehicles (DMV) locations in Las Vegas, Nevada. The DMV was selected because the general population of U.S. citizens visits the DMV; even those without a driver’s license get state IDs at a DMV. Additionally, patrons of the DMV often wait long periods of time in an indoor environment, thereby facilitating test taking. These two specific DMVs were selected because one DMV (known as the American Pacific DMV) is in a relatively affluent area while the other (known as the Sahara DMV) is in a lower-income area of town.

Because this study also sought saliva samples from participants, only male participants were asked to provide saliva samples for assessment of cortisol and sIgA due to the biological variation in females previously described. Salivary collection was performed only between the hours of 11 a.m. and 3 p.m. to control for the influence of circadian rhythm. The only exclusionary criterion for participating in the survey portion was a failure to speak and read in English. For males, exclusion for submitting a saliva sample was acknowledgement of an autoimmune medical condition such as allergies, autoimmune disease, endocrine metabolic disorders, or acquired immunodeficiency syndrome.

Participants were recruited at a booth set up at the entrance with a sign soliciting for participation (the sign stated “Bored? Take a short study. Return it when you leave. For UNLV. Free water”). To assist recruitment and address the self-recruitment limitation identified in study one, a student also approached patrons to seek participation. The student stated, “Hello, I’m with UNLV conducting a research study. Please help us out; take this study in with you and fill it out while you wait. Drop the survey off when you leave.” If participants agreed, the student verified that they spoke English and understood the questions. The response rate to this recruitment was 65%. Participants received either bottled water or hot chocolate depending upon weather conditions.

#### *Measures*

*Income.* Absolute income is hypothesized to be distinctly different than RD. Therefore, feelings of deprivation should not be highly related to income. Self-reported personal income, household income, and education items were therefore measured through participant reports of estimated yearly income before taxes.

*Education.* An item of self-reported education was included to address a limitation identified in study three. Participants reported their education as one of the following choices: *Some high school, Completed high school/GED, Some college, Completed college, Some graduate school, Completed graduate school, or I prefer not to say.*

*Perceived Control.* Perceived control was assessed through aggregated ratings of agreement to nine questions (e.g. “At home, I feel I have control over what happens in most situations” and “I often have the feeling that I am being treated unfairly” (Bobak et al., 2000). Previous research has identified a Cronbach’s alpha reliability score for this

scale as 0.65 Scores along this measure mediate the affects of material deprivation, suggesting the measure is a valid assessment of perceived control.

*Social Participation/Status.* There is no consistently used measure or format to assess the construct of social participation. Marmot (2005) asserted that social participation was related to the status syndrome. This conclusion was based upon studies similar to those that are described above: exhaustive lists of organizations and abilities in which participants rate their ability to have or engage in them. Each available measure appears to have been tailored for the population of interest. None of the current measures of social participation are appropriate for the current study. The purpose of the current research is to test a measure of RD, and as such, an untested measure of social participation would not be useful for construct validation. Given the unclear boundaries and wide diversity that may occur in social participation, a generalizable and psychometrically fit measure would need to be developed for the current population of interest (U.S. adults). Developing and testing a new measure of social participation is beyond the scope of the current research. Marmot (2005) considered social participation the process by which status affects health (along with perceived control), though from a research perspective social participation was not necessary for the identification of status.

Due to the conceptual and measurement issues outlined above, social participation was measured through a self-reported status rating. Previous research has identified a discrepancy between one's reported status and actual standing (Runciman, 1966). A single question was included that measures status, following the structure provided by Thompson and Hickey (2005) of five class systems (lower class, working class, lower

middle class, upper middle class, and upper class). This class structure was selected among others (Gilber, 2002, Beeghley, 2004) because of the clear boundaries between classes (e.g. upper middle class is argued to be household income between \$100,000 and 500,000). Participants were asked: “What class do you consider yourself? (please circle the appropriate number). A scale was provided of five categories (*Lower, Working, Lower middle, Upper middle, and Upper*). Each ranking had four points, allowing for a response range of “1” (corresponding to the lowest ranking of *Lower* class) to “20” (corresponding to the highest ranking of *Upper* class).

*Perceived Stress.* The Perceived Stress Scale (PSS) (Cohen, Kamarck, & Mermelstein, 1983) was used to measure perceived stress. This 14-item inventory measures the frequency of feelings of anxiety regarding potentially stressful events over the last month. Responses range from zero (*never*) to four (*very much so*). The PSS was designed around three factors of the stress experience: unpredictable, uncontrollable, and overloading. These factors have been found to be central to the stress experience (Cohen, 1978; Lazarus, 1977). The scale is evenly divided in positively and negatively worded items and designed for samples with at least a junior high school education (Cohen et al., 1983). The original study of the PSS found no sex or age differences, and was adequately reliable- both in alpha (coefficient alpha ranged from .84 to .86) and test-retest (.85 at the two day mark and .55 for the six-week mark). Validation of the PSS identified relationships of the scale with measures of life events, social anxiety, depression, and health measures.

*Biological Parameters.* Cortisol and sIgA were collected through voluntary saliva donation. Saliva samples were collected and assayed with Salivary Cortisol Enzyme

Immunoassay Kit 1-3002 and Secretory IgA Enzyme Immunoassay Kit 1-1602, both from Salimetrics, LLC. Collection and handling procedures will be consistent with the recommendations of Shirtcliff, Granger, Schwartz, & Curran (2000) and Salimetrics. Only males were recruited for participation in the saliva donation for two reasons. The exclusion of females from salivary donation was because of hormonal variations that may confound with a stress-immune relationship and the nature of the study design. A relationship between the current measure of RD/RG and endocrine or immune activity has not previously been established. Sex differences may exist in regard to HPA activity (Kirschbaum et al., 1999), and an effect found may not be outside the range of health functioning (Glaser et al., 1999). Furthermore, this study utilized a between-subjects design without any within-subjects component. The majority of endocrine and immune research regarding stress establishes a within subject baseline, exposes participants to a stressor or control, and collects biological samples before and after the hypothesized stressor at varying time intervals (Coste, Strauch, Letrait, & Bertagna, 1994). The naturalistic setting of the present research does not facilitate within-subject testing. Any effect that exists may be moderated by sex and/or barely detected. Because of the influence of the menstrual cycle on the cortisol response to stress, and the previous research findings of greater cortisol response in men, the health measures of cortisol and sIgA were only attained from male participants.

*Self-reported health.* Single item self-report measures of health have consistently been found to predict disease and mortality (Idler & Benyamini, 1997). Self-reported health functions as a better predictor of survival than variables such as medical records or self-reports of medical conditions. The authors noted that the wording of the self-

rated health question differed across studies, with some asking respondents to compare their health to others their age while others asked an overall general rating. All showed high predictability of health, suggesting the concept of self-rated health status is robust to semantic differences in wording. In the current study the participants were asked: “On a scale of one to ten please rate your overall health” with answers ranging from 1 (*Very poor health*) to 10 (*Excellent health*). Additionally, they were asked “On a scale of one to ten please rate your overall health compared to others the same age” with answers ranging from 1 (*Very poor health*) to 10 (*Excellent health*).

*Health Behaviors.* Items assessing health behaviors were included to assess the frequency that participants smoke, drink, exercise, and wear a seatbelt.

*RD/RG.* The magnitude version of the Perceived Inequity measure format designed in study three was used to measure RD and RG. However, the items were revised to represent the prominent comparison groups/objects identified in study two. The five comparison groups were friends, family members, coworkers, previous self, and general; the four objects of comparison: general, financial, interpersonal, and health.

The intensity dimension of the scale was not included to reduce survey length, as seventy items may be too long and monotonous to adequately measure the consequence of social comparison (and then asking participants to complete the PSS, perceived control, and demographic items). The magnitude items were selected over intensity because the intensity items were tethered to the magnitude; the intensity items were “How do you feel about this?” after each magnitude item.

The RD/RG measure developed in study 2 was designed to measure participants’ perception of their comparisons. Items were created to assess participants’ perception in

reference to the selected comparison groups along a given object of comparison. This was intended to be a measure of the magnitude of the comparisons.

The format of the questions was “How do you view yourself compared to \_\_\_\_\_ on \_\_\_\_\_?” The range of answers was on a 7-point scale ranging from -3 (*doing worse*) and 3 (*doing better*). Sum scores can be created for RD by rescaling all positive responses to zero and for RG by rescaling all negative responses to zero. A sum score, representing the overall consequence of comparison can be derived from a total sum score.

## Results

### *Analyses*

In the descriptive analyses, the measure and the test locations were evaluated. The reliability of the perceived inequity measure was tested through Cronbach’s alpha. The test sites, the two DMVS, were contrasted by differences in age, income, and perceived inequity.

In the construct validity analyses, the perceived inequity measure was tested through correlation analysis with the measures of perceived control, perceived stress, and income. Analyses of variance (ANOVA) were used to determine whether perceived inequity scores differed by sex, ethnicity, and age. RD and RG were derived from inequity scores and the influence of age and education to the prediction of RD and RG were evaluated using linear regression.

After the reliability and validity of the measure were tested, the predictive analyses were evaluated. The ability of RD/RG to predict health was tested as well as the constructs’ relationship to related phenomenon. Linear regression was used to test the

ability of the RD/RG measure to predict health outcomes such as the self-rated health items, BMI, and immune measures in concert with other constructs that may predict health such as perceived control and perceived stress.

### *Descriptive Analyses*

The Consequence of Social Comparison inequity index item scores ranged from “-3” (indicating the maximum perceived sense of deprivation compared to others) to “3” (indicating the maximum perceived sense of gratification compared to others). The mean for the perceived inequity index was 30.99 ( $SD = 23.00$ ). Relative deprivation and gratification scores were derived through a similar procedure described in study 2. The RD mean was -4.70 ( $SD = 8.11$ ) and the RG mean was 35.85 ( $SD = 17.80$ ). Mean and variance scores for all scales and mathematically derived subscales are included in Table 9.

*Hypothesis: The magnitude ratings of perceived inequity will have sufficient reliability as tested by Cronbach's alpha; the site types will differ significantly by personal income but not perceived inequity.*

Cronbach's alpha for the perceived inequity was  $\alpha = .94$ . The average age of the American Pacific DMV was 27 ( $SD = 11.24$ ) and the Sahara DMV was 34 ( $SD = 13$ ). The personal incomes of the average American Pacific DMV participant was statistically significantly greater than Sahara,  $F(1, 233) = 5.79, p = .02, \eta^2 = .02$ . Inequity scores, however, were not statistically significantly different by site ( $F(1, 265) = 1.4, p = .24$ ).

*Hypothesis: The magnitude dimensions of inequity will not have statistically significant ethnicity differences as tested through analysis of variance (ANOVA)*

Scores of inequity were not statistically significantly different by ethnicity,  $F(4, 262) = 1.30, p = .26$ . Similarly, the inequity scores were not statistically significantly different by ethnicity when transformed to represent RD ( $F(4, 262) = 2.05, p = .11$ ) and RG ( $F(4, 262) = 1.17, p = .32$ ).

*Hypothesis: The magnitude dimension of inequity will not have statistically significant sex differences as tested through analysis of variance (ANOVA)*

The mean inequity for females was 33.50 ( $SD = 23.60$ ) and for males it was 29.60 ( $SD = 22.50$ ). Scores of inequity were not statistically significantly different by sex, ( $F(1, 265) = 1.85, p = .18$ ), Transformed inequity scores produced RD mean score for females of -4.58 ( $SD = 8.72$ ) and of males a mean of -4.78 ( $SD = 7.71$ ).

*Hypothesis: The magnitude dimension of inequity will have statistically significant differences in age as tested through ANOVA.*

Age was categorized into “18 – 24” ( $n = 113$ ), “25 – 34” ( $n = 61$ ), “35 – 49” ( $n = 67$ ), “50 – 64” ( $n = 22$ ), “65 and over” ( $n = 4$ ). Due to the low sample size of the final group, the last two groups were collapsed into “50 and over”. ANOVA did not identify a statistically significant difference by age for inequity scores ( $F(3, 263) = .74, p = .53$ ) nor for the derived measures of RD ( $F(3, 263) = .20, p = .90$ ) or RG ( $F(3, 263) = .47, p = .70$ ). Therefore, the hypothesis was not supported.

*Hypothesis: The magnitude dimension of inequity will have statistically significant differences in education as tested through ANOVA.*

The self-reported education distribution was “some high school” ( $n = 15$ ), “Completed high school/GED” ( $n = 45$ ), “Some college” ( $n = 148$ ), “Completed college” ( $n = 32$ ), “Some graduate school” ( $n = 8$ ), “Completed Masters degree” ( $n = 8$ ),

“Completed Doctorate degree (n = 4). For ANOVA these groups were recoded into three education categories: “Some/completed high school” (n = 60), “Some/completed college” (n = 180), and “Some/completed graduate school” (n = 20). The ANOVA test revealed no statistically significant difference in education between inequity scores ( $F(2, 257) = .08, p = .92$ ). Therefore, this hypothesis was not supported.

### *Construct Validity*

*Hypothesis: The magnitude dimensions of perceived inequity will be significantly correlated with related variables but will not be significantly correlated with income.*

Correlation analyses were performed to identify the relationship between the inequity index and measures predicted to be associated as well as those predicted to not correlate with the inequity measure (see Table 10). The inequity index did not correlate statistically significantly with PSS scores ( $r = -.12, p = .05$ ) or perceived control ( $r = -.03, p = .67$ ). The inequity index was statistically significantly correlated to personal income ( $r = .24, p < .01$ ) and was correlated statistically significantly with self-reported class ( $r = .30, p < .01$ ).

### *Health Prediction*

*Hypothesis: The magnitude dimensions of both RD and RG will be significantly correlated with health, as will the alternate hypothesized constructs argued to predict health (perceived control, perceived stress, status, and income).*

Five indicators of health were measured: BMI, self-reported health, self-reported health respective to participant age, salivary cortisol, and salivary immunoglobulin A (sIgA). BMI correlated statistically significantly with both self-reported general health ( $r = .36, p < .01$ ) and self-reported general health to others of the same age ( $r = .30, p <$

.01). The self-reported general health question was also statistically significantly correlated to self-reported general health to others of the same age ( $r = .76, p < .01$ ).

CORT and sIgA were not statistically significantly correlated to one another ( $r = -.07, p = .60$ ). CORT was not correlated statistically significantly to collection time ( $r = -.06, p = .62$ ) or length of time to collect sample ( $r = -.10, p = .42$ ). Similarly, sIgA was not significantly correlated to collection time ( $r = -.06, p = .86$ ) or time to collect sample ( $r = .09, p = .46$ ). Samples of sIgA were correlated to age ( $r = -.23, p = .05$ ) but was not statistically significantly correlated to any of the hypothesized constructs. As such, the CORT and sIgA are not included in subsequent analysis.

RD was not significantly correlated to BMI ( $r = -.08, p = .18$ ). RD was significantly correlated to self-reported general health ( $r = .36, p < .01$ ) and self-reported general health to others of the same age ( $r = .30, p < .01$ ). Similarly, RG was not significantly correlated to BMI ( $r = -.01, p = .82$ ). RG was significantly correlated to self-reported general health ( $r = .40, p < .01$ ) and self-reported general health to others of the same age ( $r = .36, p < .01$ ).

Table 11 includes the correlations of RD, RG, perceived stress, perceived control, status, and income to the health measures of BMI and the two self-rated health items. Income was not significantly correlated to any of the health measures. Similarly, neither PSS nor perceived control scores were correlated significantly to BMI or either self-reported health questions.

Sex differences were obtained in the correlation of self-reported class and health predictors. With males, self-reported class was not significantly correlated to any of the health predictors (see Table 11). For females, however, self-reported class was

significantly correlated to BMI ( $r = .26, p < .01$ ), self-reported health rating ( $r = .22, p = .02$ ), and self-reported health with others of the same age ( $r = .23, p = .02$ ). Therefore, this hypothesis was partially supported in the general health ratings were correlated to inequity measures.

*Hypothesis: The magnitude dimensions of both RD and RG will both significantly predict self-reported health measures as tested through linear multiple regression and contribute unique variance to the prediction of health in addition to perceived stress and perceived control.*

All status indicators were included in a regression model prediction of health. RD, RG, status, income, PSS, class, and perceived control scores were included as independent variables for the prediction of the general health question using standard linear regression (see Table 12). Only RD and RG were identified as statistically significant predictors. The model was then re-run with only RD and RG as predictors ( $R^2 = .19$ ; RD  $B = .04, SE = .01, t = 3.36, p < .01$ ; RG  $B = .03, SE = .01, t = 4.53, p < .01$ ).

The health prediction estimation model was performed again using RD and RG as predictors but splitting the sample by sex. The model was significantly predicted by RD but not RG for females ( $R^2 = .15$ ; RD  $B = .01, SE = .02, t = .53, p = .60$ ; RG  $B = .03, SE = .01, t = 3.26, p < .01$ ) and both RD and RG for males ( $R^2 = .26$ ; RD  $B = .07, SE = .02, t = 4.10, p < .01$ ; RG  $B = .02, SE = .01, t = 3.43, p < .01$ ). Therefore, this hypothesis was supported.

*Hypothesis: The magnitude dimensions of both RD and RG will significantly predict risk behavior.*

RD and RG were significantly correlated to measures of risk behavior assessing participants' ratings of exercise (see Table 13). RD and RG were not significantly correlated to alcohol consumption or cigarette use. Regression analysis was performed separating the sample by sex. For the prediction of exercise rates in females, RD was not a significant predictor while RG was a significant predictor ( $R^2 = .10$ ; RD  $B = .003$ ,  $SE = .01$ ,  $t = .23$ ,  $p = .83$ ; RG  $B = .02$ ,  $SE = .01$ ,  $t = 2.68$ ,  $p < .01$ ) while for males the inverse was found; RD was a statistically significant predictor but RG was not ( $R^2 = .05$ ; RD  $B = .03$ ,  $SE = .01$ ,  $t = 2.65$ ,  $p < .01$ ; RG  $B = -.001$ ,  $SE = .005$ ,  $t = -.20$ ,  $p = .84$ ).

### Conclusion

The purpose of this study was to employ the method of measuring RD and RG identified in study one and two but with a larger and more diverse sample. Additionally, other constructs that have been argued to also measure the relationship between inequity and health were administered to identify which measure best predicts health. To achieve greater age diversity, both studies three and four recruited participants from DMVs in Southwestern United States. The first study in this experiment identified the comparison groups and object that were prominent with the new population. These comparison groups were: friends, family members, coworkers, previous self, and general; the four objects of comparison were general, financial, interpersonal, and health. A measure of RD/RG was designed to measure the perceived inequity of these comparisons, the same method used in study one. Participants also completed scales measuring the constructs of perceived control, perceived stress, status (as a measure of social participation), and self-reported income.

Perceived inequity scores (as well as the RD/RG derivations) did not differ significantly by ethnicity, sex, age, or education level. Age and education were both hypothesized to be significantly different by inequity. The failure to obtain support for this hypothesis suggests that perceived inequity as measured through prominent comparison groups/objects may be durable to variations in such demographics. This may be because the measure allows the participant to identify who specifically is making up each group. For example, a person with an advanced degree may feel the same when comparing themselves to their friends as someone who did not complete high school because they both feel they are doing worse than their friends. Both people are likely considering different people in the comparison but the effect could be the same.

The perceived inequity measure was compared to other constructs identified as important in the relationship between inequity and health. Perceived stress and perceived control were not significantly correlated to the inequity measure while self-reported class and income were significantly correlated. In the transformation of RD and RG, RD was significantly correlated to perceived stress, while RG was not. This supports the notion that RD and RG are separate constructs and holds with theory that RD may be stressing (Wilkinson, 2005). Additionally, the correlation with RD/RG and self-reported class and income demonstrate that a link between perceived inequity and income does exist in an absolute sense as well.

The relationship between the selected inequity constructs and health were explored. Health is a challenging construct to measure (Maier & Watkins, 1998; Jones & Bright, 2001; Wilkinson, 2005). As such, three approaches to measuring health were considered: immune measures, self-ratings, and BMI. Two immune measures were

collected through salivary collection: cortisol and sIgA, two biological products related to both stress and health (Pawlow & Jones, 2005). Salivary collection only occurred in males and between the hours of 11am and 3pm to control for influences of circadian rhythm. The immune measures were not found to relate to any of the measured constructs or demographics (with the exception of age for cortisol). Most immune studies are performed within-subjects; participants have their immunity tested before and after an experimental condition (Jones & Bright, 2001). There is likely far too much individual variation in such immune measures to find a correlation between immune-measures and the selected social constructs. As such, while ambitious the measures of cortisol and salivary immunoglobulin A were poor measures of health for the current study.

There were, however, several significant correlations between the self-reported health items and the constructs predicted to relate to health. Self-reported health has consistently predicts mortality (Idler & Benyamini, 1997). Two items were used, self-reported health and self-reported health compared to others of the same age. Overall, these correlations were stronger for the “overall” health assessment than the “same age” rating. This may be due to the same age item being too narrow in scope. RD, RG, and self-rated class were all significantly correlated to self-reported health. Perceived control and perceived stress, however, were not correlated significantly to the self reported health measures, suggesting the links between stress and control to health, while meaningful and potentially significant, are not as linearly correlated as the effect of perceived inequity. Put differently, perceived inequity appears to be a better predictor of health than income, stress, and control.

BMI was not found to be significantly related to any of the inequity indexes. Sex differences were also identified with these correlations. Class was a statistically significant predictor of health in females for both self-rated health items and BMI (the only significant correlation found). As self-rated class increased, BMI decreased.

This effect of self-rated class in females is interesting. One explanation is that class is very meaningful to the health and happiness of females. Another explanation is that the worse a female's health is, the lower their perceived status. Again, the item was self-rated status, a measure of social participation (the ability to do the things one wants and feels they should be able to do). A more accurate measure of true class, income, was not correlated to any of the health measures. As such, there is a relationship between females' sense of status and health that suggests it relates to inequity.

Our measures of inequity, RD and RG, support this conclusion by demonstrating a strong relationship with the health indices. RG was strongly predictive of health assessments in both men and women. However, RD was more predictive of men's self-reported health than women's health. The increased strength of the relationship of RD and health for men than women is noteworthy. Perhaps males are more impacted by doing worse than their comparison groups. The identification of sex differences in the strength of the relationship between perceived inequity and health has not been previously documented and would need to be a consideration for future studies exploring this link.

To test the strength of the social constructs to predict health among one another, multiple regression was used to determine which predictors were useful after accounting for the effects of other predictors. RD and RG consistently emerged as superior predictors

of health. They were the only two predictors to account for a significant proportion of the variance of the general health question.

As with the previous study, sex differences were identified. RG did not significantly predict health in females. This may be due to social facilitation. This explanation is unlikely as the construct was RG, one's sense of gratifications. Perhaps females felt they would be "showing off" to flaunt their believed successes over others. Another possibility is that there is a distinct sex difference in the relationship of perceived inequity and health.

The ability of RD and RG to predict health behaviors was also tested. Health behaviors are often considered in the relationship between perceived inequity (believed to be stressing) and health (Bobak et al., 2000; Marmot, 2005; Wilkinson, 2005). Smoking and alcohol consumption rates were not significantly related nor predicted by RD or RG. A possible explanation for this lack of effect is that these are two items people may not feel comfortable disclosing or are not aware of how much they do smoke cigarettes and consume alcohol. Exercise rates, however, were significantly predicted by RD and RG. Again, sex differences emerged. Only RD was predictive of exercise rates in males while RG was only predictive of exercise rates in females.

There were several limitations to this study. The measurement model required testing the population at one point in time, determining prominent comparison groups and objects, and then retesting the sample with those prominent groups considered. The population may have fundamentally changed between testing sessions, reducing the reliability of the measure. Additionally, as the test is self-report in nature, social facilitation is a concern. Participants may not be aware of their comparisons or are not

willing to discuss them. This is a well discussed concern in social comparison research (Wood, 1996). The relationship between this measurement approach and social facilitation should be explored. Additionally, alternatives to such a direct model of measurement should also be explored now that this approach has been supported as a usable method.

This study established that the “global RD/RG” approach to measuring perceived inequity is a reasonably reliable and valid measure. The measures reliability over time was alpha level = .94. The internal consistency was high and indicated that the inter-item correlation was very strong. The scale was consistently measuring the same construct with each item. . The measure manifested strong construct validity. It was strongly correlated to class. Additionally, this model can be employed for measuring inequity related health disparities as determined by the strong correlation to health measures as well as the superior variance accounted for by RD and RG over other social inequity measures such as perceived stress, perceived control, perceived status, and personal income. Additionally, sex differences were obtained in the relationship between perceived inequity and health, a relatively unexplored phenomenon. Further research should identify alternative methods of measuring global RD and RG to identify if the self-report method is causing the sex differences or if these are real differences in how females and males are affected by inequity. A “global” measure of RD and RG derived from prominent comparisons predicts health and is a critical step towards understanding the role of perceived inequity in health disparities.

## CHAPTER 7

### DISCUSSION

The purpose of this research was to test the feasibility of a comparison-based self-report measure of RD and RG, identify the ability of such a measure to predict health, and to test this predictability against alternative constructs that have been argued to predict health. To accomplish these goals, a measure of RD and RG based on prominent social comparisons was developed. Measures of health were identified and the ability of the new measure to predict health was identified. Then this measurement approach was tested with a more demographically diverse sample along with other measures of constructs that have been argued by other researchers to predict health.

Health and inequity seem to be linked, and this goes beyond objective conditions. In our study, income and self-reported health had virtually no relationship. Income appears to be related to health but once basic needs are met this relationship diminishes (Wilkinson, 2005). Marmot contended that the perceived control and social participation were the cause of status-linked health disparities. Stress may also be the agent that affects health through influencing the body's immune response (Maier & Watkins, 1998).

Research has identified *a* effect relationship between inequality and health but not *how* it occurs. This effect would require a psychological process by which seemingly arbitrary possession disparities result in concrete health disparities. The psychological processes of relative deprivation (RD) and gratification (RG) may be such processes'. RD is the sense that one is feeling deprived, regardless of objective conditions. The converse is RG, in which one feels better about their condition. Both processes may be created through social comparisons with others. RD may create stress that over time can lower a

person's baseline immunity while RG increases one's ability to cope with stress (thereby acting as a buffer to the health effects of stress). RD and RG-distress and gratification (respectively) derived from inequity- belong among these constructs as essential to understanding the relationship between inequity and health.

Previous measures of RD and RG were not suitable for the current study. Multiple measures of RD have emerged due to the various definitions of the construct. Stoufer et al. (1949) vaguely described RD, leaving subsequent researchers to develop their own definitions and subsequent measures. Runciman's (1966) definition of RD was based from the following logic: "If relative deprivation is to be precisely described, all inequalities which give rise to feelings of relative deprivation must be treated as inequalities between and only between the membership reference group and comparative reference group" (p. 14). Gurr (1970), conversely, chose objective measures of RD. RD was derived from expectations, Gurr argued, and these expectations could be measured through "status quo" indexes such as income. Objective-based approaches to measuring RD, however, have been discounted. Merton and Kitt (1950) described RD as a conclusion derived from comparisons to referents, and one has no guarantee that objective conditions are indeed creating RD. Research has gravitated towards supporting a reference-based approach of measuring RD and the inverse, RG (Guimond & Dambrun, 2002; Dambrun, Taylor, McDonald, Crush, and Mèot, 2006).

A similar comparison-based approach was selected for the current research design, though none of the existing measures of RD were appropriate as all were designed for specific test questions and narrow situations. No "general" RD/RG measure was identified. As such, a measure of RD/RG was designed. In the planning stages of

developing the measure, the role of social comparison was explored. The construct of one's comparisons to others had been argued as necessary to RD by most of the RD researchers (Davis, 1959; Runciman, 1966; Cook, Crosby, and Hennigan, 1977). The comparison groups and objects people employ- those likely to elicit feelings such as RD and RG- are unique for each person (Wood, 1996).

A social comparison approach to measuring status assumes that the health effects of status are derived from perception of status instead of or beyond the absolute effects of status. However, there are two benefits to measuring social position with such a “perceptive” approach. The first is the flexibility of such a measure across diverse groups. In the Whitehall studies of British civil servants, social position was easily measured through civil service rankings (Marmot, 2003). Pay grade was a near mirror image of status. However, in groups with a variety of professions and interests, the estimation of social position becomes more difficult. In this regard, perceptive measures may be more precise than researcher derived calculations. Given the possible influence of social comparison on health, a clear understanding of the relationship between these two constructs is beneficial for understanding health disparities. Identifying prominent comparisons is an appropriate starting point as the effects of social comparisons are likely derived from the comparisons the individual finds most meaningful.

An overall scale of RD/RG, a so-called “global” measure that was based on social comparisons, would therefore require being able to identify prominent comparison groups and objects of comparison for each person. Rather, we took the approach of identifying the *prominent* comparison groups and objects within a population, and then developing an inequity measure around those groups and objects. The goal of such an

approach was to not fully measure RD and RG within each person but measure an overall sense of RD/RG.

Runciman (1966) contended that RD was made up on “magnitude” and “intensity”. Magnitude was the extent of the disparity while intensity was the degree of distress caused. These dimensions were included in our measure. Participants rated how they were doing compared to a selected comparison group by a given object. This was to measure the “magnitude” component of RD/RG. Such a method creates the possibility that a participant does not compare themselves to any of the groups or objects selected by the population. As such, a follow up item was included in which participants rated the degree to which such a comparison impacted them. This follow up item was designed to measure the “intensity” component of RD/RG.

The RD/RG measure was derived through two stages. First, the population of interest was surveyed for prominent comparison groups and objects. A self-rating approach was employed by which participants rated the degree to which they care about comparisons towards groups and objects deemed by previous research to be prominent. Additionally, participants were encouraged to volunteer groups and objects they use for comparison to identify additional comparisons. Second, the RD/RG measure was created in which items were created to measure the perceived inequity perceived compared to each group by each object. Participants rated each such comparison from “-3” to “3”. Negative scores denoted RD while positive scores were interpreted as RG.

This method of measuring RD and RG was first tested with college students from a southwestern university in the United States. Prominent comparison groups were friends, family members, coworkers, previous self, celebrities/athletes, members of the same

gender, and general. The five prominent objects of comparison were general, financial, interpersonal, achievements, and appearance.

These groups were used for the creation of a perceived inequity measure. Additionally, a measure of belief in a just world (BJW) and self-reported income was used for the construct validation of the scale. Perceived inequity was hypothesized to correlate to BJW and not to income, and that hypothesis was supported in the study. The RD and RG derivations correlated to measures of health. However, this study suffered from the limited age variation in the sample. Also, the measures of health were encouraging but limited. Therefore, we sought to duplicate the study with a sample containing a greater age range and using more sophisticated health measures.

Our method of measuring RD and RG was applied to a more diverse population through sampling at DMVs in a southwest city in the United States. As the population of study had changed, the prominent comparison groups and objects identified in study one using college students, were generally replicated with a the more representative general population. However, several differences emerged. may not be accurate. As such, the assessment of prominent comparison groups and objects was repeated. With this study the prominent comparison groups that emerged were friends, family members, coworkers, previous self, and general. The four prominent objects of comparison that emerged were general, financial, interpersonal, and health. Unlike the previous sample, members of the same gender and celebrities/athletes did not emerge as prominent groups. As for objects, appearances did not emerge as prominent but a related object, health, did. This illustrates the importance of the prominent group/object assessment of the RD/RG measurement as well as emphasizing a link between health and social comparison as

many people in the sample are consciously aware and admit to making comparisons to others regarding their health. The approach identified here was valid but laborious- each measure of RD/RG would require two studies- one to identify prominent comparison groups and objects and another to test the perceived inequity. However, between both experiments several groups and objects were present in both; comparisons such as those to friends, coworkers and family and along such dimensions as money, interpersonal, and an overall sense of doing well to others. These groups and objects could suffice as a global measure of RD/RG. Further research is needed to evaluate this possibility.

A scale to measure perceived inequity was derived from the above comparison groups and objects. Having previously established a relationship of our perceived inequity measure to health with a less demographically diverse sample, we sought to not only identify if this relationship held with the new sample, but also to test the strength of this relationship compared to other constructs that had been hypothesized to predict health. Income, perceived stress, perceived control, and social participation were measured. Income was measured through self-reported income. Perceived stress was measured (Cohen, Karmarck, & Memelstein, 1983) along with perceived control (Bobak, Pikhart, Rose, Hertzman, & Marmot, 2000). Social participation, the act of doing and affording the things one thinks they ought to be able to, does not have a pre-established scale but is similar to one's perceived status (Marmot, 2005). The self-reported perceived inequity scale and the RD/RG derivations differently correlated to the related measures. RD was strongly associated with stress scores. RD may be stressing but RG does not seem therapeutic as there was no strong relationship between RG and PSS scores. Income was strongly related to RD and very strongly related to RG, but there is potential for overlap

in variance as income was an object of comparison in the scale. However, income is tied to one's sense of inequity. Status, the measure of social participation, was strongly related to RD and RG.

The strong relationships identified between these constructs suggest they are important towards understanding health inequities. Measuring health, however, is difficult because there is no one universal measure of health (Jones & Bright, 2001). A strong predictor of mortality and health problems is self-reported health (Idler, & Benyamini, 1997). This item was included along with self-reported health compared to others the same age. Additionally, efforts were made to measure health through objective indicators. BMI was calculated for each participant from reported height and weight and salivary cortisol and immunoglobulin A (sIgA) were measured. While the BMI and neuro-endocrine measures were not tied to any of the social constructs, the strongest of the health measures to the inequity constructs was self-reported health. Using regression analyses to evaluate the ability of each of the proposed inequity measures to predict health, RD and RG emerged as stronger predictors of health than income, perceived stress, perceived control, and social participation.

The underlying mechanisms that support the relationship between inequity and health are profoundly influential but not understood. The current research was an effort to identify if RD and RG can predict health and how well this measure predicts health when considered among other constructs that are also theorized to be tied to health inequities. Our measures of RD and RG were derived from the targeted study of the prominent comparison groups and objects. The measure was an attempt to directly measure individual's perceived inequity rather than inferring perceived inequity as is measured

with constructs such as perceived stress and social participation. The strength of such an approach is the measurement of the *perception* of inequity. This was accomplished through evaluating social-comparison derived inequities and considering such conclusions the equivalent of RD and RG. Moving forward, the durability of this relationship should be tested. If these effects could be diminished through education about such an effect, then perception could also be the means to reduce health inequities.

APPENDIX 1

IRB APPROVALS

Table 1

*Descriptives and t-test values for groups and objects of comparison against a score of “2” (value equivalent of “I care a little about this group”)*

<b>Comparison</b>		<b>Mean</b>	<b>Std deviation</b>	<b>t</b>	<b>Cohen’s d</b>
Groups	Coworkers	2.70	1.15	8.64**	0.61
	Friends	3.43	1.10	18.38**	1.30
	Family members	3.25	1.18	15.05**	1.06
	Previous self	3.45	1.20	17.12**	1.21
	Members of same religion	1.32	1.40	-6.89**	-0.49
	Fictional characters	1.55	1.29	-4.91**	-0.35
Objects	General	2.79	.83	13.88**	0.95
	Financial	2.22	1.01	3.11**	0.22
	Interpersonal	2.79	.85	13.11**	0.93

\*\*  $p < .01$

Table 2

*ANOVA values for groups and objects of comparison by gender*

Comparison		Sum of Squares	df	Mean Square	F	
Groups	Coworkers	Between Groups	2.18	1	2.18	1.66
		Within Groups	260.62	199	1.31	
	Total		262.80	200		
	Friends	Between Groups	.24	1	.02	.02
Within Groups			243.23	199	1.22	
Total		243.25	200			
Family		Between Groups	.05	1	.05	.04
	Within Groups		275.17	199	1.39	
	Total		275.22	200		
	Previous Self	Between	<.001	1	<.001	<.001

---

		Groups				
		Within	289.56	199	1.45	
		Groups				
		Total	289.56	200		
Same		Between	9.21	1	9.21	4.80*
Religion		Groups				
		Within	378.40	199	1.92	
		Groups				
		Total	387.62	200		
Fictional		Between	10.83	1	10.83	6.68*
character		Groups				
		Within	323.09	199	1.62	
		Groups				
		Total	333.92	200		
Objects	General	Between	.13	1	.18	
		Groups				
		Within	145.69	210		
		Groups				
		Total	145.81	211		
	Financial	Between	1.70	1	1.69	

---

---

	Groups			
	Within	199.69	198	
	Groups			
	Total	201.40	199	
Interpersonal	Between	2.99	1	4.16*
	Groups			
	Within	143.17	199	
	Groups			
	Total	146.17	200	

---

\*  $p < .05$ , \*\*  $p < .01$

Table 3

*Means and descriptives for Consequences of Social Comparison scale and subscales,*

*HSCL-30, BJW, self-reported health, and income*

Scale	n	Mean	Std. Dev
Consequences of Social Comparison	238	71.90	59.76
Magnitude	244	33.98	30.26
RD Magnitude	246	-10.52	10.52
RG Magnitude	246	44.42	22.32
Intensity	238	37.24	32.85
RD Intensity	238	-9.40	11.09
RG Intensity	238	46.58	25.77
HSCL-30	246	4.88	4.07
BJW	246	29.86	5.51
Self-reported health	244	7.92	1.36
Income	213	19976.82	21207.53

HSCL-30: Hopkins Symptomology Checklist 30 Questions

BJW: Belief in a Just World

Table 4

*Correlations of Perceived Inequity and Validity Indices*

	BJW	Income	Self-reported health	HSCL-30
Consequence of Comparison	.35**	.13	.42**	-.38**
Magnitude	.36**	.13	.41**	-.35**
RD	.32**	.07	.30**	-.40**
RG	.34**	.14*	.41**	-.29**
Intensity	.32**	.12	.40**	-.37**
RD	.27**	.07	.24**	-.42**
RG	.29**	.12	.41**	-.29**

\*  $P < .05$ , \*\*  $P < .01$

HSCL-30: Hopkins Symptomology Checklist 30 Questions

BJW: Belief in a Just World

Table 5

*Descriptives and t-test values for groups and objects of comparison*

	Comparison	Mean	Std deviation	t	Cohen's d
Groups	Coworkers	2.36	1.48	3.08**	0.24
	Friends	3.03	1.48	8.81**	0.70
	Family members	3.18	1.50	9.89**	0.79
	Previous self	3.51	1.48	12.96**	1.02
	Members of same religion	1.67	1.67	-2.50*	-0.20
	Fictional characters	.90	1.29	-10.77**	-0.85
Objects	General	2.53	1.02	6.62**	.52
	Financial	2.27	1.21	2.84**	.22
	Interpersonal	2.51	1.14	5.76**	.45

\*p &lt; .05, \*\*p &lt; .01

Table 6

*Analysis of Variance of Comparison Groups by Gender*

Comparison		Sum of Squares	df	Mean Square	F	
Groups	Coworkers	Between	4.70	1	4.70	2.17
		Groups				
		Within	340.63	157	2.17	
		Groups				
	Total	345.33	158			
	Friends	Between	4.31	1	4.21	1.95
		Groups				
		Within	339.74	157	2.16	
		Groups				
	Total	343.95	158			
	Family	Between	.86	1	.86	.38
		Groups				
Within		355.76	157	2.26		
Groups						
Total	356.62	158				
Previous Self	Between	1.70	1	1.70	.78	

---

		Groups				
		Within	343.12	157	2.19	
		Groups				
		Total	344.81	158		
Same		Between	5.64	1	5.64	2.04
Religion		Groups				
		Within	434.03	157	2.77	
		Groups				
		Total	439.67	158		
Famous		Between	7.86	1	7.86	4.85**
People		Groups				
		Within	254.31	157	1.62	
		Groups				
		Total	262.17	158		
Objects	General	Between	3.37	1	3.37	3.31
		Groups				
		Within	159.95	157	1.02	
		Groups				
		Total	163.32	158		
	Financial	Between	3.66	1	3.66	2.51

---

---

	Groups				
	Within	229.19	157	1.46	
	Groups				
	Total	232.85	158		
Interpersonal	Between	.58	1	.58	.45
	Groups				
	Within	205.02	157	1.31	
	Groups				
	Total	205.60	158		

---

\*\* p < .01

Table 7

*Analysis of Variance of Comparison Groups by Age*

Comparison		Sum of Squares	df	Mean Square	F		
Groups	Coworkers	Between Groups	5.71	3	1.90	.87	
		Within Groups	339.62	155	2.19		
		Total	345.33	158			
		Friends	Between Groups	4.00	3	1.34	.61
		Within Groups	339.95	155	2.19		
		Total	343.95	158			
		Family	Between Groups	3.31	3	1.10	.69
			Within Groups	353.31	155	2.28	
			Total	356.62	158		
		Previous Self	Between	6.54	3	2.18	1.00

---

		Groups				
		Within	338.27	155	2.18	
		Groups				
		Total	344.81	158		
Same		Between	33.13	3	11.04	4.21**
Religion		Groups				
		Within	406.53	155	2.62	
		Groups				
		Total	439.67	158		
Famous		Between	6.86	3	2.29	1.39
People		Groups				
		Within	255.30	155	1.65	
		Groups				
		Total	262.17	158		
Objects	General	Between	2.57	3	.86	.83
		Groups				
		Within	160.76	155	1.04	
		Groups				
		Total	163.32	158		
	Financial	Between	4.39	3	1.46	.99

---

---

	Groups				
	Within	228.46	155	1.47	
	Groups				
	Total	232.85	158		
Interpersonal	Between	4.28	3	1.43	1.10
	Groups				
	Within	201.32	155	1.30	
	Groups				
	Total	205.60	158		

---

\*\* p < .01

Table 8

*Differences by Age Group on Rating of “Members of the Same Religion”*

Group	n	M	SD	18-24	25-34	35-49	Over 49
18-24	43	1.30	1.50	NS	NS	NS	*
25-34	45	1.44	1.50	NS	NS	NS	*
35-49	51	1.74	1.80	NS	NS	NS	NS
Over 49	20	2.78	1.73	*	*	NS	NS

Note: NS = nonsignificant differences between pairs of means, while an asterisk (\*) = significance using Tukey HSD with alpha of .05.

Table 9

*Means and descriptives for Consequences of Social Comparison scale and subscales,*

*PSS, Perceived control, and income*

Scale	n	Mean	Std. Dev
Consequences of Social Comparison	266	31.12	23.00
RD	266	-4.70	8.11
RG	266	35.84	17.80
PSS	266	19.84	7.44
Perceived control	266	39.86	6.60
Income	234	34267.77	32949.97

Table 10

*Correlations of Perceived Inequity and Validity Indices*

	PSS	Perceived control	Income	Class
Consequence of Comparison	-.12	-.03	.24**	.30**
RD	-.18**	-.06	.17*	.31**
RG	-.08	-.005	.24**	.25**

\*  $P < .05$ , \*\*  $P < .01$

Table 11

*Correlations of Perceived Inequity and Validity Indices Split by Sex*

	BMI	General Health	General Health of Same Age
Income	.08	.03	.05
<i>Females</i>	.03	.09	.17
<i>Males</i>	.07	-.01	-.01
PSS	-.04	-.10	-.12
<i>Females</i>	-.14	-.11	-.09
<i>Males</i>	.09	-.07	-.09
Perceived control	-.04	-.02	-.08
<i>Females</i>	-.15	-.06	-.16
<i>Males</i>	-.009	-.008	-.06
Class	-.06	.14*	.07
<i>Females</i>	-.26**	.22*	.23*
<i>Males</i>	.06	.09	-.03
RD	-.08	.36**	.30**
<i>Females</i>	-.13	.24*	.22*
<i>Males</i>	-.05	.46**	.37**
RG	-.01	.40**	.36**
<i>Females</i>	-.03	.38**	.41**
<i>Males</i>	.03	.43**	.35**

\*  $P < .05$ , \*\*  $P < .01$

Table 12

*Summary of Multiple Regression Analysis for General Health (n = 224)*

Variable	B	SE(B)	$\beta$	<i>t</i>	Sig. (p)
Income	.00	.00	-.11	-1.76	.08
Class	.02	.03	.03	.50	.62
PSS	-.02	.02	-.08	-1.31	.10
PC	-.01	.02	-.02	-.31	.75
RD	.05	.01	.23	3.26	< .01
RG	.03	.01	.32	4.55	< .01

$R^2 = .24.$

Table 13

*Summary of Correlations of RD and RG to Risk Behaviors (n = 266)*

Variable	RD	RG
Drinking alcohol	-.05	-.08
Rate of exercise	.20**	.16**
Cigarette smoking	-.07	-.08
Not wearing seatbelt	-.06	-.12

\*  $P < .05$ , \*\*  $P < .01$

## REFERENCES

- Adams, J. (1965). Inequity in social exchange. In Berkowitz (Ed.), *Advances in experimental social psychology*, 2. New York: Academic Press.
- Ahrens, A. (1991). Dysphoria and social comparison: Combining information regarding others' performances. *Journal of Social and Clinical Psychology*, 10, 190-205.
- Alicke, M. (2000). Evaluating social comparison targets. In Suls & Wheeler (Eds.) *handbook of Social Comparison: Theory and Research* (pp. 271-294). New York: Kluwer.
- Alicke, M., Klotz, M., Breitenbecher, D., Yrak, T., & Vredenberg, D. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, 68, 804-825.
- Arrowood, A. (1986). Comments on "social comparison theory: Psychology from the lost and found." *Personality and Social Psychology Bulletin*, 12, 279-281.
- Austin, W. (1977). Equity theory and social comparison processes. In Suls & Miller (Eds) *Social Comparison Processes, Contemporary Theory and Research* (pp. 279-305). Halsted Press: New York.
- Beach, S., & Tesser, A. (2000). Self-evaluation maintenance and evolution: Some speculative notes. In Suls & Wheeler (Eds.) *Handbook of Social Comparison: Theory and Research* (pp. 123-140). New York: Kluwer.
- Beck, A. (1976). *Cognitive therapy and the emotional disorders*. New York: International Universities Press.
- Beeghley, L. (2004). *The Structure of Social Stratification in the United States*. Boston, MA: Pearson, Allyn & Bacon

- Bell, D. (1987). Gender differences in the social moderators of stress. In R.C. Barnett, L. Biener, & G.K. Baruch (Eds.), *Gender and stress* (pp. 257-277). New York: The Free Press.
- Berkowitz, L. (1972). Frustration, comparisons, and other sources of emotional arousal as contributors to social unrest. *Journal of Social Issues*, 28, 77-91.
- Bobak, M., Pikhart, H., Rose, R., Hertzman, & Marmot, M. (2000). Socioeconomic factors, material inequalities, and perceived control in self-rated health: cross-sectional data from seven post-communist countries. *Social Science & Medicine*, 51, 1343-1350.
- Bosch, A., Ring, C., de Geus, E., Veerman, E., Amerongen, A. (2003). Stress and secretory immunity. *International Review of Neurobiology*, 52, 213-252.
- Braconier, J., Nilsson, B., Oxelius, V., & Karup-Pedersen, F. (1984). Recurrent pneumococcal infections in a patient with lack of specific IgG and IgM pneumococcal antibodies and deficiency of serum IgA, IgG2 and IgG4. *Scandinavian Journal of Infection Diseases*, 16, 407-410.
- Brandtzaey, P., Baekkevold, E., Farstad, I., Jahnsen, F., Johansen, F., Nilson, E., & Yamanaka, T. (1999). Regional specialization in the mucosal immune system: what happens in the microcompartments? *Immunology Today*, 20, 141-151.
- Breznitz, S. (1990). Theory-based stress measurement? Not yet. *Psychological Inquiry*, 1, 17-19.
- Brickman, P., & Bulman, R. (1977). Pleasure and pain in social comparison. In J. Suls & R. Miller (Eds.), *Social comparison processes: theoretical and empirical perspectives* (pp. 149-186). Washington, DC: Hemisphere.

- Brickman, P. & Kessler, D. (1977). Preference for nonreciprocal social comparison. In Suls & Miller (Eds.), *Social comparison processes: theoretical and empirical perspectives* (pp. 154-160). Washington, DC: Hemisphere.
- Brown, K., & Grunberg, N. (1995). Effects of housing on male and female rats: Crowding stresses males but calms females. *Physiology and Behavior*, 58, 1085-1089.
- Bosch, J., de Geus, E., Ring, C., & Amerongen, A. (2004). Academic examinations and immunity: academic stress or examination stress. *Psychosomatic medicine*, 66, 625-327.
- Brugha, R., & Cragg, D. (1990). The list of threatening experiences: The reliability and validity of a brief life events questionnaire. *Acta Psychiatrica Scandinavica*, 82, 77-81.
- Buunk, B., & Gibbons, F. (2000). Toward an enlightenment in social comparison theory: Moving beyond classic and renaissance approaches. In Suls & Wheeler (Eds.) *handbook of Social Comparison: Theory and Research*. (pp. 487-499). New York: Kluwer.
- Buunk, A., Peiró, J., & Griffioen, C. (2007). A positive role model may stimulate career-oriented behavior. *Journal of Applied Social Psychology*, 37, 1489-1500.
- Cannon, W. (1932). *The wisdom of the body*. New York: Norton.
- Cantril, H. (1965). *The pattern of human concerns*. New Brunswick, N.J.: Rutgers University Press.
- Cassell, J. (1976). The contribution of the social environment to host resistance. *American Journal of Epidemiology*, 104, 107-123.

- Chapin, S. (1928). A quantitative scale for rating the home and social environment of middle-class families in an urban community: a first approximation to the measurement of socio-economic status. *Journal of Educational Psychology, 19*, 99-111.
- Clow, A. (2001). The physiology of stress. In Jones and Bright (eds.), *Stress: Myth, Theory and Research* (pp. 47-61). London: Pearson Education Limited.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum
- Cohen, S., Karmarck, T., & Memelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*, 385-396.
- Cohen, S. (1986). Contrasting the Hassles Scale and the Perceived Stress Scale: Who's really measuring appraised stress? *American Psychologist, 41*, 716-718.
- Cohen, S., Frank, E., Doyle, W., Skoner, D., Rabin, B., & Gwaltney, J. (1998). Types of stressors that increase susceptibility to the common cold in healthy adults. *Health Psychology, 17*, 214-223.
- Coon, D. (2005). *Psychology: A Modular Approach to Mind and Behavior*. New York: Wadsworth Publishing.
- Cook, T., Crosby, & Hennigan, K. (1977). The construct validity of relative deprivation. In Suls & Miller (Eds) *Social Comparison Processes, Contemporary Theory and Research* (pp. 307-333). New York: Halsted Press.
- Coste, J., Strauch, G., Letrait, M., & Bertagna, X. (1994). Reliability of hormonal levels for assessing the hypothalamic-pituitary-adrenocortical system in clinical pharmacology. *British Journal of Clinical Pharmacology, 38*, 474-479.

- Cox, T. (1993). *Stress Research and Stress Management: Putting Theory to Work*. (Health and Safety Executive Contract Research Report No 61/1993). Suffolk, UK: Health and Safety Executive.
- Cox, B., Borger, S., & Enns, M. (1999). Anxiety sensitivity and emotional disorders: Psychometric studies and their theoretical implications. In S. Taylor (Ed.), *Anxiety Sensitivity: Theory, research and treatment of the fear of anxiety* (pp. 115-148). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Crosby, F. (1979). Relative deprivation revisited: A response to Miller, Bolce, and Halligan. *The American Political Science Review*, 73, 103-112.
- Crosby, F. (1982). *RD and Working Women*. New York: Oxford University Press.
- Davies, J. (1962). Toward a theory of revolution. *American Sociological Review*, 27, 5, 5-19.
- Davis, J. (1959). A formal interpretation of the theory of relative deprivation. *Sociometry*, 22, 280-296.
- Davis, J. (1966). The campus as a frog pond: An application of theory of relative deprivation to career decisions for college men. *Academic Journal of Sociology*, 72, 17-31.
- Davis, M., Matthews, K., & Twamley, E. (1999). Is life more difficult on Mars or Venus? A meta-analytic review of sex differences in major and minor life events. *Annals of Behavioral Medicine*, 21, 83-97.
- Deinzer, R., & Schuller, N. (1998). Dynamics of stress-related decrease of salivary

- immunoglobulin A (SIgA): Relationship to symptoms of the common cold and studying behavior. *Behavioral Medicine*, 23, 161-170.
- Delongis, A., Folkman, S., & Lazarus, R. (1988). Hassles, health, and mood: Psychological and social resources as mediators. *Journal of Personality and Social Psychology*, 54, 486-495.
- Eibner, C., & Evans, W. (2005). Relative deprivation, poor health habits and mortality. *Journal of Human Resources*, 40, 591-620.
- Eisert, D. & Kahle, L. (1982). Self-evaluation and social comparison of physical and role change during adolescence: A longitudinal analysis. *Child Development*, 53, 98-104.
- Ellaway, A. & Macintyre, S. (2007). Is social participation associated with cardiovascular risk factors? *Social Science & Medicine*, 64, 1384-1391.
- Evans, P., Der, G., Ford, G., Hucklebridge, F., Hunt, K., & Lambert, S. (2000). Social class, sex, and age differences in mucosal immunity in a large community sample. *Brain, Behavior, and Immunity*, 14, 41-48.
- Festinger, L. (1950). Informal social communication. *Psychological Review*, 57, 271-282.
- Festinger, L. (1954a). A theory of social comparison processes. *Human Relations* 7, 117-40.
- Festinger, L. (1954b). Motivation leading to social behavior. In M. Jones (Ed.), *Nebraska symposium on motivation* (Vol. 2, pp. 191-218). Lincoln: University of Nebraska Press.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford

University Press.

Fougeyrollas, P., Noreau, L., Bergeron, H., Cloutier, R., Dion, S.A., St-Michel, G.

(1998). Social consequences of long term impairments and disabilities: conceptual approach and assessment of handicap. *International Journal of Rehabilitation Research, 21*, 127-141.

Gerard, H. (1963). Emotional uncertainty and social comparison. *Journal of Abnormal and Social Psychology, 66*, 568-573.

Gilbert, D. (2002) *The American Class Structure: In An Age of Growing Inequality*. Belmont, CA: Wadsworth

Girdler, S., Jammer, L., & Shapiro D. (1997). Hostility, testosterone, and vascular reactivity to stress: Effects of sex. *International Journal of Behavioral Medicine, 4*, 242-263.

Glaser, R. (2005). Stress-associated immune dysregulation and its importance for human health.: A personal history of psychoneuroimmunology. *Brain, Behavior, and Immunity, 19*, 3-11.

Glaser, R., Rabin, B., Chesney, M., Cohen, S., & Natelson, B. (1999). Stress-induced immunosuppression. *Journal of the American Medical Association, 281*, 2268-2270.

Goethals, G. (1986). Social comparison theory: Psychology from the lost and found. *Personality and Social Psychology Bulletin, 12*, 261-278.

Goethals, G., & Darley, J. (1977). Social comparison theory: An attributional approach. In Suls & Miller (Eds.) *Social Comparison Processes: Theoretical and Empirical Perspectives*. (pp. 259-278). Washington, DC: Hemisphere.

- Goethals, G. & Klein, M. (2000). Interpreting and inventing social reality: Attribution and constructive elements in social comparison. In Suls & Wheeler (Eds.) *handbook of Social Comparison: Theory and Research*. (pp. 23-44). New York: Kluwer.
- Goethals, G., Messick, D., & Allison, S. (1991). The uniqueness bias: Studies of constructive social comparison. In Suls & Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 149-173). Hillsdale, NJ: Lawrence Erlbaum.
- Gough, H. (1952). Predicting social participation. *Journal of Social Psychology*, 35, 227-233.
- Guimond, S., & Dambrun, M. (2002). When prosperity breeds intergroup hostility: The effects of relative deprivation and relative gratification on prejudice. *Personality and Social Psychology Bulletin*, 28, 900-912.
- Gurney, J., & Tierney, K. (1982). Relative deprivation and social movements: A critical look at twenty years of theory and research. *The Sociological Quarterly*, 23, 33-47.
- Gurr, T. (1968). A causal model of civil strife: A comparative analysis using new indices. *The American Political Science Review*, 62, 1104-1124.
- Gurr, T. (1970). *Why men rebel*. Princeton, NJ: Princeton University Press.
- Hakmiller, K. (1966). Threat as a determinant of downward comparison. *Journal of Experimental Social Psychology*, Suppl. 1, 32-39.
- Heider, F. (1944). Social perception and phenomenal causality. *Psychological Review*, 51, 358-373.

- Herman, J., Prewitt, C., & Cullinan, W. (1996). Neuronal circuit regulation of the hypothalamo-pituitary-adrenocortical stress axis. *Critical Reviews in Neurobiology, 10*, 371-394.
- Hinkle, L. (1973). The concept of 'stress' in the biological and social sciences. *Science, Medicine and Management, 1*, 31-48.
- Idler, E., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior, 38*, 21-37.
- Jex, S. M., Beehr, T. A. & Roberts, C. K. (1992). The meaning of occupational stress items to survey respondents. *Journal of Applied Psychology, 77*, 623-628.
- Jones, F. & Bright, J. (2001). *Stress, Myth, Theory and Research*. London: Prentice Hall.
- Jones, F., & Kinman, G. (2001). Approaches to studying stress. In Jones and Bright (eds.), *Stress: Myth, Theory and Research* (pp. 17-45). London: Pearson Education Limited.
- Kasl, S. (1978). Epidemiological contributions to the study of work stress. In C. L. Cooper and R. L. Payne (Eds.), *Stress at work*. Chichester: Wiley and Sons.
- Kelly, H. (1973). The process of causal attribution. *American Psychologist, 28*, 107-128.
- Kendall, P. & Lazarsfeld, P. (1950). Problems of survey analysis. In Merton & Lazarsfeld (Eds.) *Continuities in Social Research: Studies in the Scope and Method of "The American Soldier"*. (pp. 133-196). Glencoe, IL: The Free Press.
- Keverne, E., Nevison, C., & Martel, F. (1999). Early learning and the social bond. In C.

- Carter, I Lederhendler, & B. Kirkpatrick (Eds.) *The integrative neurobiology of affiliation* (pp. 263-274). Cambridge, MA: MIT Press.
- Klein, W., & Kunda, Z. (1993). Maintaining self-serving social comparisons: Biased reconstruction of one's past behaviors. *Personality and Social Psychology Bulletin, 19*, 732-739
- Latanté, B. (1966). Studies in social comparison- Introduction and overview. *Journal of Experimental Social Psychology, Suppl. 1*, 1-5.
- Lazarus, R. (1990). Theory-based stress measurement. *Psychological Inquiry, 1*, 3-13.
- Lazarus, R., & Folkman, S. (1984). *Stress, Appraisal and Coping*. New York: Springer.
- Lazarus, R., DeLongis, A., Folkman, S., & Gruen, R. (1985). Stress and adaptional outcomes: The problem of confounded measures. *American Psychologist, 40*, 770-779.
- Lazarus, R. (1990). Theory-based stress measurement. *Psychological Inquiry, 1*, 3-13.
- Lerner, M. (2003). The justice motive: Where social psychologists found it, how they lost it, and why they may not find it again. *Personality and Social Psychology Review, 7*, 388-399.
- Lewin, K., Dembo, T., Festinger, L., & Sears, P. (1944). Level of aspiration. In J. Hunt (Ed.), *Personality and behavior disorders, 1*, 333-378. New York: Ronald Press.
- Maier, S., & Watkins, L. (1998). Cytokines for psychologists: Implications of bidirectional immune-to-brain communication for understanding behavior, mood, and cognition. *Psychological Review, 105*, 83-107.
- Marmot, M., Adelstein, A., Robinson, N., & Rose, G. (1978). The changing social class distribution of heart disease. *British Medical Journal, 2*, 1109-1112.

- Marmot, M. (2005). *The Status Syndrome: How Social Standing Affects Our Health and Longevity*. New York: Holt Paperbacks.
- McGrath, J. (1976). Stress and behavior in organizations. In M. Dunnette (Ed.), *Handbook of Industrial and organizational Psychology*. Chicago: Rand McNally
- Merton, R., & Kitt, A. (1950). Contributions to the theory of reference group behavior. In Merton & Lazarsfeld (Eds.) *Continuities in Social Research: Studies in the Scope and Method of "The American Soldier"*. (pp. 40-105). Glencoe, IL: The Free Press.
- Miller, D. Turnbull, W. & McFarland, C. (1988). Particularistic and universalistic evaluation in the social comparison process. *Journal of Personality and Social Psychology*, 55, 908-917.
- Morrison, D. (1973). Some notes toward theory on RD, social movements, and social change. In R.R. Evans (Ed.), *Social Movements: A Reader and Source Book*. Chicago: Rand McNally.
- Nissen, H. (1954). Comments on Professor Festinger's paper. In Jones (Ed.), *Nebraska symposium on motivation*. (Vol. 2, pp. 219-223). Lincoln: University of Nebraska Press.
- Okami, P. (1992). Intolerable grievances patiently endured: Referent cognitions and group conflict as mediators of anti-Jewish sentiment among African-Americans. *Political Psychology*, 13, 727-753.
- Pawlow, L., & Jones, G. (2005). The impact of abbreviated progressive muscle relaxation on salivary cortisol and salivary immunoglobulin A (sIgA). *Applied Psychophysiology and Biofeedback*, 30, 375-387.

- Pettigrew, T., Christ, O., Meertens, R., van Dick, R., & Zick, A. (2008). Relative deprivation and intergroup prejudice. *Journal of Social Issues, 64*, 385-401.
- Pollock, K. (1988). On the nature of social stress: production of a modern mythology. *Social Science and Medicine, 26*, 381-392.
- Poundstone, W. (1992) *Prisoner's Dilemma*. New York: Doubleday.
- Powers, B. A. (2004). Self-perceived health of elderly institutionalized people. *Journal of Cross-Cultural Gerontology, 3*, 299-321.
- Rasmussen, P., Willingham, J., & Glover, T. (1996). Self-esteem stability, cynical hostility, and cardiovascular reactivity to challenge. *Personality and Individual Differences, 21*, 711-718.
- Reckman, R. & Goethals, G. (1973). Deviancy and group-orientation as determinants of group composition preferences. *Sociometry, 36*, 419-423.
- Redelmeier, D., & Singh, S. (2001). Survival in Academy Award-winning Actors and Actresses. *Annals of Internal Medicine, 134*, 955-962.
- Reiss, S., Peterson, R., Gursky, D., & McNally, R. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behavioural Research and Therapy, 24*, 1-8.
- Runciman, W.G. (1966). *RD and Social Justice*. London: Routledge.
- Rutter, M., Kreppner, J., & O'Connor, T. (2001). Specificity and heterogeneity in children's response to profound institutional privation. *British Journal of Psychiatry, 179*, 97-103.
- Schachter, S. (1959). *The psychology of affiliation*. Stanford, CA: Stanford University Press.

- Scwartz, N., & Strack, F. (1991). Evaluating one's life: A judgment model of subjective well-being. In F. Strack, M. Argyle, & Schwartz, N. (Eds.), *Subjective well-being: An interdisciplinary perspective* (pp. 27-47). New York: Pergamon.
- Seyle, H. (1956). *The stress of life*. New York: McGraw-Hill book Co.
- Seyle, H. (1976). *The stress of life*. (revised edition). New York: McGraw-Hill book Co.
- Shirtcliff, E., Granger, D., Schwartz, & Curran, M. (2000). Use of salivary biomarkers in biobehavioral research: cotton-based sample collection methods can interfere with salivary immunoassay results. *Psychoneuroendocrinology*, *26*, 165-173.
- Spada, M., Nikčević, A., Moneta, G., & Wells, A. (2008). Metacognition, perceived stress, and negative emotions. *Personality and Individual Differences*, *44*, 1172-1181.
- Stouffer, S., Suchman, E., DeVinney, L., Star, S., & Williams, R. (1949). *Studies in Social Psychology in World War II: The American Soldier. Vol. 1, Adjustment During Army Life*. Princeton, NJ: Princeton University Press
- Stouffer, S. (1950). Afterthoughts of a contributor. In Merton & Lazarsfeld (Eds.) *Continuities in Social Research: Studies in the Scope and Method of "The American Soldier"* (197-211). Glencoe, IL: The Free Press.
- Suls, J. (1977). Social comparison theory and research: An overview from 1954. In Suls & Miller (Eds) *Social Comparison Processes, Contemporary Theory and Research*. (pp. 1-20). Halsted Press: New York
- Suls, J. (2000). Opinion Comparison: The role of the corroborator, expert, and proxy in social influence. In Suls & Wheeler (Eds.) *handbook of Social Comparison: Theory and Research* (pp. 105-122). New York: Kluwer.

- Suls, J. & Wheeler, L. (2000). A selective history of classic and neo-social comparison theory. In Suls & Wheeler (Eds.) *handbook of Social Comparison: Theory and Research* (pp. 3-19). New York: Kluwer.
- Taylor, S., Klein, L., Lewis, B., Gruenwald, T., Gurung, R., & Updegraff, J. (2000). Biobehavioral responses to stress in females; Tend-and-befriend, not fight-or-flight. *Psychological Review*, *107*, 411-429.
- Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 181-227). New York: Academic Press.
- Tesser, A. (1991). Emotion in social comparison and reflection processes. In Suls & Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 117-145). Hillsdale, NJ: Lawrence Erlbaum.
- Tesser, A., & Collins, J. (1988). Emotion in social reflection and comparison situations: Intuitive, systematic, and exploratory approaches. *Journal of Personality and Social Psychology*, *55*, 695-709.
- Tesser, A., & Smith, J. (1980). Some effects of friendship and task relevance on helping: You don't always help the one you like. *Journal of Experimental Social Psychology*, *16*, 582-590.
- Thompson, W. & Hickey, J. (2005). *Society in Focus*. Boston, MA: Pearson, Allyn & Bacon
- Vanneman, R. D. & Pettigrew, T. (1972). Race and relative deprivation in the urban United States, *Race* *13*, 461-486.
- Webber, C. (2007). Revaluating relative deprivation theory. *Theoretical Criminology*,

11, 95-118.

- Wheeler, L. (1962). *Desire: A determinant of self-evaluation through social comparison*. In prep doctoral dissertation, University of Minnesota, Minneapolis, MN.
- Wheeler, L. (1966). Motivation as a determinant of upward comparison. *Journal of Experimental Psychology*, 2, 27-31.
- Wheeler, L. (1991). A brief history of social comparison theory. In Suls & Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 3-21). Hillsdale, NJ: Lawrence Erlbaum.
- Wheeler, L., & Miyake, K. (1992). Social comparison in everyday life. *Journal of Personality and Social Psychology*, 62, 760-773.
- Wheeler, L. (2000). Individual differences in social comparison. 141-158.
- Wheeler, L. (1966). Motivation as a determinant of upward comparison. *Journal of Experimental Social Psychology*, (Supplemental 1), 27-31.
- Wheeler, L., & Zuckerman, M. (1977). Commentary. In Suls & Miller (Eds.) *Social Comparison Processes* (pp. 335-357). New York: Halsted Press
- Williams, R., Haney, T., Lee, K., Kong, Y., Blumenthal, J., & Whalen, R. (1980). Type A behavior, hostility, and coronary atherosclerosis. *Psychosomatic Medicine*, 42, 539-549.
- Wills, T. (1981). Downward comparison principles in social psychology. *Psychological Bulletin*, 90, 245-271.
- Wills, T. (1991). Similarity and self-esteem in downward comparison. In Suls & Wills

- (Eds.), *Social comparison: Contemporary theory and research* (pp. 51-78). Hillsdale, NJ: Lawrence Erlbaum.
- Wilkinson, R.G., (2005). *The impact of inequality*. Routledge, London: New Press.
- Windsor, T., Anstey, K., Butterworth, P., & Rodgers, B. (2008). Behavioral approach and behavioral inhibition as moderators of the association between negative life events and perceived control in midlife. *Personality and Individual Differences, 44*, 1080-1092.
- Wood, J. (1996). What is social comparison and how should we study it? *Personality and Social Psychology Bulletin, 22*, 520-537.
- Yitzhaki, S. (1979). RD and the Gini coefficient. *Quarterly Journal of Economics, 93*, 321-324.
- Zinbarg, R., Mohlman, J., & Hong, N. (1999). Dimensions of anxiety sensitivity. In S. Taylor (Ed.), *Anxiety Sensitivity: Theory, research, and treatment of the fear of anxiety* (pp. 83-114). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Zinbarg, R., Brown, T., Barlow, D., & Rapee, R. (2001). Anxiety sensitivity, panic, and depressed mood: A reanalysis teasing apart the contributions of the two levels in the hierarchical structure of the Anxiety Sensitivity Index. *Journal of Abnormal Psychology, 110*, 372-377.

## VITA

Graduate College  
University of Nevada, Las Vegas

Samuel A. Holland

### Degrees:

Bachelor of Science, Psychology, 2001  
James Madison University

Master of Science, Psychological Science, 2003  
James Madison University

### Special Honors and Awards:

Graduate Student Association Grant, 2009

### Publications:

Trice, A. D., Holland, S. A., & Kitsantas, A. 2001. The effects of mother's developmental expectations, health knowledge, and locus of control, on children's health knowledge and locus of control. *Journal of Genetic Psychology*.

Trice, A. D., Holland, S. A., & Gagne, P. E. 2000. Voluntary class absences and other behaviors in college students: An exploratory study. *Psychological Reports*. 87 179-182

Trice, A. D., & Holland, S. A. 2001. *Heroes, antiheroes, superheroes, dolts: Changing images of masculinity in popular American films, 1921-1999*. Jefferson, NC: Macfarlane Press

Dissertation Title: Relative Deprivation, Relative Gratification, Status, and Health

### Dissertation Examination Committee:

Chairperson, Jeffrey M. Kern, Ph. D.

Committee Member, David Copeland, Ph. D.

Committee Member, Cortney Warren, Ph. D.

Graduate Faculty Representative, Timothy Bungum, Ph. D