

# **Master's Thesis in Behavioral Economics**

## **A Literature review of sources of bias in Self Assessed Health ratings; Proposed experiment to testing framing effects**

### **Abstract:**

This literature review reveals Self Assessed Health rating is all about the context. Context of: the individual; the question; the survey, leading to gaps between observed versus perceived health information, health experience versus health expectations, presentation versus answering of questions. These gaps creates 7 sources of biases in Self Assessed Health ratings: availability of objective health information, subjective health information, use of self-concept in health, over-confidence and optimism, framing effects, social desirability, and satisficing. Potential for such sources of biases should be noted and corrected for when necessary, to facilitate more accurate health policy assessments and cross countries equality comparisons.

**Name: Juliana Koh Pei Hoon**

**Student number: 411342**

**Supervisor: Teresa Bago d'Uva**

**Place: Rotterdam**

**Date: 20<sup>th</sup> April 2016**

## **TABLE OF CONTENTS**

<b>Section 1: Introduction</b>	<b>PG. 3</b>
1.1 Motivation to study sources of biases in SAH ratings	
1.2 Research question	
1.3 Value and contribution	
1.4 Structure of thesis	
<b>Section 2: Approach</b>	<b>PG.6</b>
2.1 Landscape of sources of biases literature in SAH and health self report context	
2.2 The literature review approach	
2.3 The selection criteria for sources of biases in SAH rating	
<b>Section 3: Theories, Proposed Framework, Assumptions</b>	<b>PG.9</b>
3.1 Theories on Cognitive Aspects of Judgment and Decision Making	
3.2 Theories on Non-Cognitive Aspects of Judgment and Decision Making	
3.3 The Cognitive Process of Survey Methodology in SAH ratings	
3.4 Introducing the framework used in this thesis	
3.4.1 The rationale for the proposed framework	
3.4.2 Introducing the framework	
3.5 Definitions and assumptions	
<b>Section 4: Sources of biases in SAH ratings</b>	<b>PG.17</b>
4.1 Pathway 1: The Health information - observed vs. perceived health information	<b>PG.17</b>
4.1.1 The health information environment and conditions:	
4.1.2 Drivers and Source of biases to SAH ratings: (Availability of objective health information)	
4.1.3 Drivers and Source of biases to SAH ratings: (Subjective health information)	
4.2 Pathway 2: The Individual – health experience versus health expectations	<b>PG.22</b>
4.2.1 The individual health environment and conditions:	
4.2.2 Drivers and Source of biases to SAH ratings: (Self-concept)	
4.2.3 Drivers and Source of biases to SAH ratings: (Over confidence and optimism)	
4.3 Pathway 3: The survey – Presentation versus Answering of question	<b>PG. 29</b>
4.3.1 The survey environment and conditions:	
4.3.2 Drivers and Source of biases to SAH ratings: (Framing effects)	
4.3.3 Drivers and Source of biases to SAH ratings: (Social desirability)	
4.3.4 Drivers and Source of biases to SAH ratings: (Satisficing)	
<b>Section 5: Proposed experiment to testing framing effects</b>	<b>PG.41</b>
5.1 Introduction	
5.2 Experiment design	
5.3 Analysis	
5.4 Discussions	
<b>Section 6: Discussions and limitations</b>	<b>PG.53</b>
6.1 Implications of subjectivity as source of bias in SAH	
6.2 Implications of overconfidence, optimism and use self-concept as source of bias in SAH	
6.3 Implications of framing effects as a source of bias in SAH	
6.4 Implications of satisficing as a source of bias in SAH	
6.5 Research limitations and Conclusions for literature review	
<b>References</b>	<b>PG.61</b>

# **A Literature review of sources of bias in Self Assessed Health ratings; Proposed experiment to testing framing effects**

## **Section 1: Introduction**

Self- Assessed health (SAH) is a self-reported overall health categorical variable. It is usually measured via asking a single question to a respondent to rate how their overall health is on a fixed response scale. Typically the response scale range is 1 “very poor”, 2 “poor”, 3 “fair” , 4 “good”, 5 “very good” There are variants to the SAH questions ranging from inverted, to different labelling of scales. Sometimes, it is also asked in comparisons to others or time such as compared to people of your age, or your health 10 years ago, respectively (Lee, 2014; Sturgis et al., 2001).

This basic and universal SAH variable is used to shape health care policies construction and evaluation such as: effectiveness of health intervention and resource allocation, measurement of health equity, monitoring societal health trends, and also as input to constructing other measures like Quality of Life, and substitute for lack of objective health measures (Sturgis et al., 2001; Lee, 2014). SAH is frequently used as it is simple to answer, cheap, easy to administer and in principle, its widespread usage allows it to be compared across studies and countries (Lee, 2014).

### **1.1 Motivation to study sources of biases in SAH ratings**

The widespread use of this SAH variable warrants a deeper understanding of what the SAH variable actually represents, especially if it contains potential sources of biases which should be accounted for. Maddox (1973) states that “SAH clearly measures something more and something less than objective medical ratings”. Being a general and self-reported variable that goes through the individual’s perception, SAH variable is subjective in nature. This subjectivity is both a bane and boon.

The subjectivity in SAH is advantageous in containing much information which cannot be observed by the physician and has been postulated to contribute to SAH’s ability to predict mortality. Idler (1992) called for more research to explore what drives perception and subjectivity of the SAH variable which gives it this ability to predict mortality and morbidity. In response, studies on what constitutes SAH ratings have found SAH to contain both objective health information as well as being rich in subjective information like demographics, social economic advantage (Damain et al., 2008; Lee, 2014), sociological influence of social peer comparison (Fienberg et al., 1985; Festinger, 1954; Van Doorn, 1999; Idler et al., 2004), adaptation and normalization of illness (Groot, 2000; Van Doorn, 1999; Idler and Kasl, 1991; Wilcox et al., 1996; Mossey and Shapiro, 1992), family health history (Van Doorn, 1999), affective aspect of health such as general feeling of well-being, mood and level of energy (Kaplan and Baron –Epel, 2003; Worsley, 1990; Segovia et al., 1989), lack of illness and healthy lifestyle (Segovia et al., 1989; Worsley, 1990) and self-concept of health (Bailis et al., 2003; Smith et al., 1994; Crossley and Kennedy, 2002).

However, just as the subjective SAH variable maybe a carrier of many other useful information, it may also be a carrier of many sources of biases. Sen (2002) commented on the subjective perception of self-reported morbidity where one’s perception of own health may not be in concordance with doctor’s objective evaluation, potentially misleading health care policies. In parallel, SAH variable shares similar characteristics with self-reported morbidity measures in being self-reported and subjective, thus similar challenges.

Subjectivity requires perspective and it is both interpretative and interactive. Research cannot find the truth to a SAH rating but only provide a better educated informed nuance when it understands the perspective taken by the individual to rate SAH. Given the complex resolution of various subjective dynamics in assessing SAH, the final SAH rating reflects the significance of one's perceived health.

To date, research still has not been able to establish a comprehensive view of what constitutes SAH ratings. Currently, even after controlling for the variety of variables and objective health measures, only 40% of SAH variation are explained by the determinants of SAH found (Baron-Epel, 2004; Lee, 2014). This limits the ability to accurately decode the actual perspective the individual has taken to rate one's health. Studying the potential sources of bias in SAH and the cognitive aspects of SAH may aid in revealing the perspective taken by respondent in their SAH rating. This may help address reasons for discordance in objective versus subjective health measures. Moreover, better grasp of the cognitive aspects of SAH will lead to stronger discernment of sources of biases in SAH. Contrary to belief that all biases are bad, one can postulate the potential advantage for sources of bias to be part of the reason of SAH's ability to predict mortality, morbidity and potential healthcare needs – as what is seen could be made by what is unseen. As such, understanding this sources of bias may allow for researchers to control for or model this factors to derive even stronger predictor of mortality, morbidity, longevity (Idler and Benyamani, 1997; Mossey and Shapiro, 1982; Lee, 2014) and health care utilization (Lee, 2014).

Furthermore, understanding sources of biases to SAH, also has potential to unearth new areas of exploration on determinants of SAH cleaned of bias that can be used to improve: health delivery (Hanita, 2000), more accurate health policy evaluations for better health care resource allocation to drive improvement in health for different population (Krause and Jay, 1994; Bago d'Uva et al., 2007), comparative studies of health status equality and equity across populations and countries (Bago d'Uva et al., 2007), studies on relationship between socioeconomic statuses with health for policy making (Lee, 2014; Mosca et al., 2013; Johnston et al., 2009).

## 1.2 Research question

This leads to 4 key research questions for this thesis:

- **RQ1:** What are the contextual influence that drives sources of biases in an individual's SAH rating?
- **RQ2:** What are the sources of biases in SAH rating and what are they all about?
- **RQ3:** What are the implications of these sources of biases in SAH rating?
- **RQ4:** How to create a thought experiment to examine if the higher educated are more susceptible to wording and labelling framing effects of SAH, leading to under rating of SAH compared to lower educated?

## 1.3 Value and contribution

This thesis wish to serve as a catalyst for thinking more seriously about sources of biases in SAH ratings via inciting a deep dive appreciation into the cognitive and non-cognitive aspects of SAH ratings. Current findings on sources of biases in SAH are often multi varied, disperse and the biases may not be directly explored. The aim of this study is to consolidate these different sources of biases in SAH, discovered through different means and research motivations into a singular perspective of the individual having to

craft a SAH rating in survey. The thesis then proceeds to consolidate and frame the sources of biases in SAH ratings induced by this singular perspective. This cumulates in a systemization of the findings through construct of a 3 pathways framework to explore the sources of biases in SAH ratings.

#### **1.4 Structure of thesis**

The rest of the thesis is organized as follows: The context, theories, frameworks used for examining SAH are first discussed in sections 2 and 3. This is followed by an in depth exposition of each source of biases in SAH in section 4. Section 5 discusses implications and limitations of the study. The thesis closes with a thought experiment in Section 6 on framing effects, as a stimulus for thinking about new starting points for future research on testing for sources of biases in SAH. More specifically:

- Section 2: Describe the approach, source of information, derivation and selection of biases;
- Section 3: Introduce the theories and the proposed framework used to explore sources of biases in SAH in this thesis and assumptions. Theories in areas of: cognitive aspects of judgment and decision making, motivations and emotions that influences decision making, cognitive process of surveys will be explored. Jungermann (1983) classifications of three biases typologies (judgmental bias, representative faults, coping defects) will be used to frame the theoretical discussion. This is followed by an introduction and rationale of the proposed framework comprising three pathways of: 1. The health information (observed vs. perceived health information), 2. The individual (health experience vs. health expectation), 3. The survey (what is asked vs. what is answered), for exploring sources of biases in SAH ratings and assumptions used in this thesis
- Section 4: Explore each sources of biases in SAH ratings identified in this study through the three pathways framework. This thesis identifies seven sources of biases to be examined: Availability of objective health information, subjective health information, over confidence and optimism, self-concept, framing effects, social desirability, and satisficing. To facilitate a more structured discussion around the biases, this thesis will adopt the first stage of Keren (1990) debiasing framework which provides a guideline for identification of a bias through an exposition on the; environment; conditions, cognitive/ non cognitive process the induce bias; actual bias description; validity of the bias. These aspects will be considered and organically integrated in the depth discussion on each source of bias.
- Section 5: Propose a hypothetical thought experiment on framing effects
- Section 6: Discuss the implications of sources of biases on SAH ratings and limitations of the research

## Section 2: Approach

This thesis conducts a literature review from research in SAH, self-report health measures, judgment and decision making, survey methodology, social psychology, cognitive psychology and method bias.

### 2.1 Landscape of sources of biases literature in SAH and health self report context

Currently, potential sources of biases of the SAH variable have been directly and indirectly explored in studies in 5 key manners, presented in table 1:

- a. Through contextualizing the cognitive aspects of answering SAH using cognitive process of survey methodology (CPSM)
  - o Potential sources of biases emerging from understanding the cognitive aspects of CPSM
- b. Through cognitive process of judgment and decision making in health self-reports
  - o Potential sources of biases emerging from reviewing the respondent's judgment and decision making process in health self reports
- c. Specific biases being directly explored with respect to SAH rating or health reports
  - o Potential sources of biases were set as hypothesis or derived as outcome of research on SAH ratings and health reports
- d. Studies on SAH subjectivity and what constitutes SAH ratings
  - o Potential sources of biases were offered as explanations for the complexity of SAH measures
- e. Indirectly offered as explanation in observations of mortality studies and SAH reporting heterogeneity studies
  - o Potential sources of biases were offered as explanations for SAH reporting heterogeneity

Each of these sources of biases will be individually explored in section 4.

**Table 1: Potential Sources of biases in SAH and Health self-reports literature**

Study type:	Potential sources of biases covered	Authors	Journal/book titles
a) Contextualize the cognitive aspects of answering SAH using cognitive process of survey methodology	Self-concept, framing effects, subjectivity of health information, social peer comparisons, social desirability, affect and mood, self vs. proxy reporting, question order effect, optimism from normalization of ill health	(Lee, 2014; Scheter, 1993; Fienberg et al., 1985)	Health Survey Methods (first edition, 2015) John Wiley & Sons, Inc., Hoboken, NJ, USA  US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics  The Milbank Memorial Fund Quarterly. Health and Society
b) Cognitive process of judgment and decision making in health self-reports	Satisficing, judgment heuristics like availability/representative/confirmation/affect heuristics, mood and emotions, priming effect, self-image	(Hanita, 2000)	Journal of Clinical Epidemiology
	Interviewer effects, framing effects	(Crossley and Kennedy, 2002; Lumsdaine and Exterkate, 2013; Bowling and Windsor,	Journal of Health Economics, Journal of epidemiology and community health,

c) Specific biases being directly explored with respect to SAH rating or health reports		2008; Brewer et al., 2004; Greene et al, 2014; Baron-Epel and Kaplan, 2001; Sturgis et al., 2001)	European Economic Review, Medical Care, Social Science and Medicine, UK Government
	Illusion of control, optimism, over confidence	(Bowling and Windsor, 2008; Taylor and Brown, 2008, Dunning et al., 2004; Weinstein, 1983; Helweg-Larsen and Shepperd, 2001)	Psychological bulletin, Psychological science in the public interest, Health psychology, Personality and Social Psychology Review, Journal of epidemiology community health
	Satisficing	(Brewer et al., 2004)	Medical Care
	Self-concept	(Bailis et al., 2003; Smith et al., 1994; Scheter, 1993 +.0; Lee, 2014)	Social science & medicine  Health Survey Methods (first edition, 2015) John Wiley & Sons, Inc., Hoboken, NJ, USA  US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics
d) Studies on SAH subjectivity and what constitutes SAH ratings	Subjectivity or health information, social determinants of health, social peer comparisons, optimism via adaptation/ normalization of illness/used of family health history ,affective aspect of health such as general feeling of well-being, mood and level of energy, lack of illness and healthy lifestyle, self-concept of health	(Damain et al., 2008; Lee, 2014; Fienberg et al., 1985; Festinger, 1954; Van Doorn, 1999; Idler et al., 2004; Idler and Kasl, 1991; Groot, 2000;; Wilcox et al., 1996; Mossey and Shapiro, 1992; Kaplan and Baron –Epel, 2003; Worsley, 1990; Segovia et al., 1989; Krause and Jay, 1994; Dunning et al., 2004; Carr et al., 2001)	Journal of health economics, American journal of public health, Journal of Epidemiology and Community Health, Journal of Health and Social Behavior, Social Science & Medicine, Medical care, Journal of Gerontology, The Journals of Gerontology Series B: Psychological Sciences and Social Sciences , Human relations, Research on Aging , Psychological science in the public interest, BMJ Health Survey Methods (first edition, 2015) John Wiley & Sons, Inc., Hoboken, NJ, USA The Milbank Memorial Fund Quarterly. Health and Society
e) Indirectly offered as explanation in observations of mortality studies, SAH reporting heterogeneity studies SAH methodology studies	Subjectivity or health information, social determinants of health leading to different health expectations, social peer comparisons, optimism, overconfidence via adaptation/ normalization of illness	(Bago d’Uva et al., 2008; Bago d’Uva et al., 2011; Greene et al., 2014; Pfarr et al., 2012; Idler, 1993; Idler and Kasl, 1995, 1991; Idler et al., 2004)	International Journal of Epidemiology, Health Economics, Journal of Human Resources, Health economics review, Journal of Gerontology, Journal of Health and Social Behavior  New York University, Leonard N. Stern School of Business, Department of Economic

Whilst many potential sources of biases in SAH are provided as explanation in studies on in observations of mortality studies and SAH reporting heterogeneity studies, as well as inferred from SAH subjectivity and what constitutes SAH ratings; studies on specific biases in SAH ratings and health self-reports are limited currently to include: interviewer effects, framing effects, illusion of control, optimism, over confidence, and self-concept. Beyond this, the most established point of view on potential biases in SAH ratings have been explored through the lens of cognitive process of survey - a field that emerged from the collaboration and synergy of cognitive psychology, survey design and statistics to develop the 4 stages of cognitive process of survey methodology. It is interesting to note the limited number studies

on this to date and most of which were internal initiatives not part of external journal publications. From the literature review table above, there appears to be a wide range of journals interested in issues of subjectivity of SAH ratings. Perhaps more research in the clear need gap for specific biases in SAH ratings to be explored directly, may be of appeal and relevance in these realms too. One other potential source to tap into for further understanding of source of biases in SAH is the area of measurement error and common method biases in behavioral research (Podsakoff et al., 2003; Viswanathan, 2005). Whilst these tend to be generalized for surveys, further research can be commissioned to examine its relevance and impact in sources of biases for SAH rating. Some of these have been explored and adopted in this thesis.

## **2.2 The literature review approach**

This literature review builds on and synthesizes what is currently available. Additionally, complementing it with literature from cognitive aspects of judgment and decision making, cognitive aspect of survey methodology, common method bias and measurement errors. The consolidation, enhancement and alignment of the different perspectives hope to stimulate an overall better understanding of the cognitive processes of SAH ratings.

Furthermore, beyond being descriptive, an exploratory design approach where knowledge on biases and theories are used and integrated so as to gain insights of their relationship with SAH, is adopted. Findings are expressed in a 3 pathways framework for sources of biases in SAH; 1. The source (Health information) - observed vs. perceived health information; 2. The process (individual) - health experience vs. health expectation; 3. The outcome (Survey) - what is asked vs. what is answered. This distinct framework besides being more holistic, also has the advantage of an easier disentanglement of the various sources of biases for further research, prevention, debiasing and self-correction.

## **2.3 The selection criteria for sources of biases in SAH rating**

The selection process for choice of sources of biases in SAH to be studied in the thesis are in accordance to its relevance. Fundamentally, all the sources of biases in SAH have to emerge from one of the 3 macro context of health information, the individual or the survey. Consequently, emerging terms, weak signals, explanations or potential biases were identified from literature in: SAH and its characteristics, specific bias exploration in SAH, cognitive aspects of SAH survey method, and self-reports of health (as reflected in table 1). Finally, there has to be sufficient research in these biases to allow for exposition of the drivers and conditions which will fuel them.

Based on this selection, 7 broad classification of sources of biases have been identified to be studied in this thesis: 1. Availability of objective health information, 2. subjective health information, 3. over confidence and optimism, 4. self-concept, 5. framing effects, 6. social desirability, and 7. satisficing. Drivers and conditions to sources of biases were derived from: readings on SAH literature and biases, inference through cognitive theories in areas of judgment and decision making and survey methodology.

The next section will introduce the assumptions, theories and framework used to investigate these sources of biases in SAH ratings and section 4 will deep dive to explore each of these sources of biases before discussing the implications and limitations in section 5.



## Section 3: Theories, Proposed Framework, Definitions, Assumptions

### 3.1 Theories on Cognitive Aspects of Judgment and Decision Making

As answering of SAH question, is about the individual having to make a decision and judgment of their own health using information they have and providing a self-report, it is important to understand the cognitive process of responding to SAH surveys. To this end, a brief discussion on the cognitive theories in decision making ensues so that we can explore the potential cognitive challenges which an individual may face in answering SAH questions, its drivers and the coping strategies used to establish SAH, which may eventually influence judgment leading to source of biases in SAH reporting.

Herbert Simon's (1955, 1957) concept of bounded rationality purports the observed reality of an individual's decision making being constrained by limited information and cognitive power due to lack of time, money, energy and memory, resulting in non-optimal response in decision making which may violate invariant preferences assumptions (Simon, 1955, 1957; Mullainathan and Thaler, 2000; Tversky and Kahneman, 1974). This triggered the field of inquiry on the cognitive processes in judgment and decision making (Jungermann, 1983; Kahneman, 2003).

Kahneman and Tversky (1974) builds on this bounded rationality theory via introducing two key concepts of intuition and accessibility of information in cognitive process. Judgment and decisions can be made through System 1 thinking which is intuitive, automatic, gut driven, fast versus System 2 thinking which is slower, more analytical, deliberate, effort driven (Kahneman, 2003). The use of intuitive, faster system 1 thinking drives the need for easier, faster accessibility of information, and certain thoughts are deemed to be more accessible than others due to habits and frequency of usage. This need for efficiency from using System 1 thinking, together with the reality of the individual's judgment and decision making being sometimes constrained by limited information and cognitive power, lead to coping strategies being engaged to access information more easily (Kahneman, 2003).

To aid in exploring these coping strategies, this thesis will use Jungermann (1983) three classifications of biases dependent on the source of coping strategies (Jungermann, 1983; Keren, 1990; Caputo, 2012). The first type are *judgmental bias* which involves the use of *judgement heuristics* in decision making that distorts judgement of probability, resulting in systematic errors (Jungermann, 1983; Tversky and Kahneman, 1974). When faced with complex task or task that requires effort, individuals may apply judgment heuristics defined as mental shortcuts, to substitute the target attribute in judgment decision, in order to ease the cognitive burden by promoting accessibility of information (Tversky and Kahneman, 1974; Kahneman, 2003; Keren, 2000; Caputo, 2012). These judgement heuristics primarily include

- **Availability heuristic:** Here is where individuals assign higher probabilities to events which are recalled more easily in memory due to recency or frequency of occurrence and use them for decision making. This lead to bias of overestimation of frequency of events and systematic error (Tversky and Kahneman, 1974; Kahneman, 2003; Caputo, 2012)
- **Representative heuristic:** This is about the neglect of base rate bias and people look the evidence that correspond with desired stereotype and may miss out contextualizing actual sample size or reliability of information (Tversky and Kahneman, 1974; Kahneman, 2003; Caputo, 2012). Example of these include people putting a higher probability to a quiet person being a librarian than a sales person, ignoring the base rate that there are more sales people out

there than librarians, or the thinking there is a higher probability of births if it is boy-girl-boy-girl-boy-girl, than boy-boy-boy-girl-girl-girl, when in reality the probability of next occurrence is independent and equal (Tversky and Kahneman, 1974; Tourangeau, 1984)

- **Confirmation heuristic:** Individuals seek and use selective information that confirms and supports their beliefs, expectations and hypothesis. Overall, being less critical to things consistent to one's beliefs, but more critical and less open if it is inconsistent with one's expectations. (Bazerman and Moore, 2009, pg. 28; Caputo, 2012; Dunning et al., 2004)
- **Anchoring heuristic:** Using what is previously set to influence subsequent decision and may lead to selection bias of information chosen to be used in decision making (Tversky and Kahneman, 1974; Kahneman, 2003)
- **Affect heuristic:** The affect heuristic arises as most judgments are perceived to first stimulate an emotional response before more thorough analytical thinking takes place (Kahneman, 2003). The affect heuristic proposes that the feelings and emotions evoked serve as a mental short cut used for judgment and decision making, instead of use of analytical thinking (Slovic et al., 2002; Kahneman, 2003; Bazerman and Moore, 2009, p.g.9; Kahneman and Ritov, 1994), overall leading to bias due to emotional evaluation and judgement. The affect heuristic is classified as a judgment heuristic when feelings are seen to be a surrogate form of thinking (Kahneman, 2003; Bazerman and Moore, 2009)

In Jungermann (1983) bias categorization, the second type of bias arises from *representational faults* (Jungermann, 1983; Keren, 2000) where different articulations of a problem, leads to a different decision frame and may lead to a different end judgment. This is related to the same root cause of accessibility of information and how frames can serve to enable accessibility to specific information. The core idea is manifested in what Tversky and Kahneman (1981) termed as "*framing effects*" and in "Prospect theory" by Kahneman and Tversky (1979)

- **Framing effect:** Framing serves to highlight and promote accessibility of particular set of information, thereby influencing decisions. "A frame contains the decision maker conception of acts, outcome and contingencies" and can lead to inconsistent outcomes if the frame is changed (Kahneman, 2003). 3 kinds of framing were introduced by Tversky and Kahneman (1979, 1981). These theories on framing shows how presentation of choices in combination, in loss versus gain perspective, or from its ability to reassure a sense of certainty can eventually influence the final choice made.
- **Prospect theory:** The idea of Prospect theory was inspired by the workings of perceptual systems and its characteristics of being reference dependent, with utility being a carrier of loss and gain (Kahneman, 2003; Caputo, 2012; Jungermann, 1983). Prospect Theory is about the use of reference point to enable a perspective based on "changes". It is based on the principle that a perspective based on "changes and differences" makes information more accessible than if the perspective was based on recalling the "absolute" (Kahneman, 2003). The theory propounds risk aversion in gains and risk seeking attitudes in loss leading to inconsistent outcomes (Tversky and Kahneman, 1979)

### 3.2 Theories on Non-Cognitive Aspects of Judgment and Decision Making

The third class of bias classification are called *coping defects* (Jungermann, 1983; Keren, 2000). Non-cognitive aspects like *motivations and moods* may also influence final decision making. These aspects

are complex and beyond the scope of this thesis to present in details. Below covers the more relevant aspects of it to this study. More details can be read up on (Janis and Mann, 1977; Janis and Mann, 1982; Bazerman and Moore, 2009, Isen et al., 1978)

- **Motivation:** Decision making from motivational perspective takes on many aspects. It was initially discussed by Janis and Mann (1977, 1982) theory of motivation as a strategy in decision making. Jungermann (1983) quotes Janis and Mann (1977) in defining this as the use of emotions, motivations and potential *wishful thinking* in decision making under stress but generalized to decision making in general. An example is “defensive avoidance strategy where a person may procrastinate, make wishful rationalization for decision, or selectively inattentive to corrective information” (Jungermann, 1983; Janis Mann, 1982). Wishful thinking subsumed as a potential driver to motivation may also lead to making of decision based on *self-serving* information, leading to failure to assimilate and use relevant information pertaining to assessment of decision at hand, biasing results. The role of motivation in decision making is further explored in literature in Bazerman and Moore (2009, p. g. 85) where conflict between *what we want versus what we think we should do* may lead to “*multiple selves*” and preference reversals (Bazerman and Moore, 2009). Use of motivation in decision making can also create *positive illusions* where people view themselves and the world in a more positive light than is accurate (Taylor and Brown, 1988; Greenwald, 1980; Bazerman and Moore, 2009) and one may even re-write one’s memory to fit desired views (Greenwald, 1980; Bazerman and Moore 2009). This may lead to a sense of *perceived self-control* and *over estimates* one’s ability for self-discipline and control over a situation.
- **Moods and emotions:** Moods are erratic, not necessarily controllable, and influences decision making, resulting in a different outcome from a decision making through pure cognitive process. Self-evaluation can be influenced by transient mood states where positive mood also lead to more positive response and optimism about everything (Hanita, 2000; Isen et al., 1978). Good mood leads to increase reliance on heuristics, use of stereotypes and may bias decision making, in contrast, being in a bad mood makes a person more critical and less use of heuristics (Park and Banaji, 2000; Bazerman and Moore, 2009, p.g. 96; Caputo, 2012) Additionally, emotions can also affect risk attitude and anticipation of regret (Bazerman and Moore, 2009)

### 3.3 The Cognitive Process of Survey Methodology aspects in SAH ratings

Besides looking at the cognitive aspects of judgment and decision making, it is important to discuss the cognitive process of survey methodology. The study of cognitive process of survey methodologies is a field that have emerged based on collaboration and synthesizing learning from cognitive and social psychology with survey methodology (Fienberg et al., 1985; Tourangeau, 1984). There are two key reasons why it is essential to discuss the cognitive process of survey methodology. First, as SAH is executed through a survey, it will also be important to understand the cognitive elements with regards to survey methodology and how these may potentially impact sources of biases in SAH ratings. Second, in the limited but emerging studies to examine cognitive aspects of SAH ratings and health self-reports, the cognitive process of survey methodology has become a cornerstone for such exploration to date (Lee, 2014; Fienberg et al., 1985; Scheter, 1993; Hanita, 2000).

The emergence of potential bias in surveys due to its cognitive aspects was first studied through non-sampling errors in surveys termed as response effects by Bradburn (1983). The various potential cognitive effects were studied through the micro social system of an interview classified into “task”, “interviewer”, “respondents”. Cognitive challenges and influences included framing effects in the “task” aspect, interviewer effects in the “interviewer” aspect and social desirability, mood and emotion influence in the “respondents” aspect (Bradburn, 1983, p.g. 293).

Subsequently, Tourangeau (1984) more formally explored these aspects by creating a framework of the four stages of cognitive aspect in surveys, which is now the commonly used benchmark in this area of studies. The four steps for respondents to be considered as optimizing their response starts with respondents having to: 1. Comprehend and interpret the questions; 2. Access their memories for relevant information; 3. Integrate and making a judgment of the information; 4. Recording and reporting answer. The cognitive aspects that has been explored in the four stages of the survey process are primarily about how accessibility of information affects the types of information being used for comprehension, retrieval , judgement and reporting (Schwarz, 1999; Bradburn, 1983; Tourangeau, 1984; Krosnick, 1991; Podsakoff et al., 2003 ). Below gives a synopsis of the various cognitive aspects to SAH ratings within this process. A more in depth exposition will be carried out for each source of biases in section 4.

- *Comprehension stage:* How the question is asked sets the context. The role of framing and anchoring effect of question order sets the context for evaluation (Schwarz, 1999; Bradburn, 1983; Tourangeau, 1984). Wording and labelling may create subjectivity in understanding of question such as being general question where one is unclear to rate based on: comparing health today with own health in other times (Fienberg et al., 1985; Scheter, 1993); comparing own health versus with others (Fienberg et al., 1985; Scheter, 1993); physical versus emotional versus mental health (Lee, 2014). Beyond this, the individual has to define their own health concept, which involves the trigger of motivation in judgment and decision making to comprehend the question (Lee, 2014).
- *Retrieval stage:* There exist cognitive challenges of recalling all relevant memories or information required for the decision making. Availability heuristic is used to aid difficulty in memory recall. Memories of high frequency or impact tend to be recalled more easily (Schwarz, 1999; Bradburn, 1983; Tourangeau, 1984). Extending to SAH, more extreme and less day to day health issues are more likely to be recalled (Lee, 2014; Fienberg et al., 1985; Scheter, 1993; Hanita, 2000). Framing and priming effect of question order or wording and labeling, drives anchoring and may influence eventual information being accessed (Lee, 2014; Fienberg et al., 1985).
- *Judgment and integration stage:* Limited cognitive power or lack of motivation due to time, energy, lack of incentives constraints one to put in best effort, are often evoked in the stage. Individuals face the cognitive challenge of many schemas reflecting beliefs on own health and having to integrate them or select own resulting schema for final rating (Fienberg et al., 1985) Hence, the individual may truncate information search through use of judgment heuristics like representative heuristics by activating motivations of self –belief/ self-concept of health associated with stereotypes to short cut judgement process (Tourangeau, 1984; Scheter, 1993; Lee, 2014; Hanita, 2000). Again, framing effects can influence judgement by limiting or promoting use of specific information to be used for evaluation of SAH (Hanita, 2000; Fienberg

et al., 1985) Additionally, priming of previous questions or external environment are postulated to also affect mood of individuals and thus their SAH assessment (Fienberg et al., 1985; Scheter, 1993; Hanita, 2000)

- *Reporting stage:* The individual is subjected to external environment and personal state of physical and emotional well-being in this final hurdle. Moods and external context like social desirability and interviewer effect can also influence outcome evaluation (Schwarz, 1999; Bradburn, 1983; Tourangeau, 1984; Lee, 2014; Fienberg et al., 1985; Scheter, 1993; Hanita, 2000). Finally, if one is faces fatigue and does not have the cognitive ability to access and process all relevant information or lack motivation to answer honestly, one will satisfice with just an acceptable response (Hanita, 2000; Krosnick, 1991). This is especially plausible in exposure to a long questionnaire (Krosnick, 1991). Needless to say, the ghost of framing effects continue to exist to the fourth stage of reporting to influence outcome of SAH rating (Lee, 2014; Fienberg et al., 1985; Scheter, 1993).

The above discussion highlights the conditions where SAH ratings may be subjected to cognitive processes that induces sources of biases. Whilst answering a SAH may indeed seem like a simple task, individuals may need to engage in more complex cognitive processes (Lee, 2014). Notwithstanding, answering SAH will at least require cognitive effort and the motivation to engage in the cognitive effort. Without which, judgment heuristics may get activated, biasing response. SAH ratings are also prone to be subjected to external moods and use of motivations in ratings. In summary, current literature has revealed emerging sources of biases in SAH ratings as: the prevalence of *framing effects* through all 4 stages of comprehension, retrieval, judgment and integration, reporting; the influence of use of motivations like *self-concept* of health in comprehension and judgement stage; the external environment and moods driving *social desirability* in reporting stage; the lack of will power and cognitive challenge resulting in the act of *satisficing* in reporting.

### **3.4 Introducing the framework used in this thesis**

#### **3.4.1 The rationale for the proposed framework**

This thesis will build upon the current foundations of cognitive aspects of SAH explored via the cognitive process of survey methodology. Whilst the current framework has strength in highlighting sources of biases in SAH in a logical and experiential survey process manner, it is constrained and confined within the four stages survey methodology process. This results in focus on the survey aspects and the individual's role and the challenges they face within the survey process, as evident from the predominant but limited sources of biases identified through this process as: framing effects, social desirability, to a minor extent satisficing and self-concept. There exist opportunities for a more in depth and holistic understanding of sources of biases in SAH, if a more specific SAH context framework is being used, such as consideration of health information and health experience.

The objective of the proposed framework is to allow for a more flexible, yet targeted exploration of sources of biases in SAH. It brings advantages of: facilitating consolidation of sources of biases through a wider range of literature which directly explored the biases or indirectly used them as explanation for observations in related health self report studies as in table 1; more actionable findings on sources of biases where it is easier to isolate, identify, explore and correct each source of bias, compared to if it

was organically presented within the four stages of comprehension, retrieval, judgment and integration, and reporting; more in depth exploration of each source of bias and wider source of biases.

### 3.4.2 Introducing the framework

This framework is based on identification of 3 macro context in SAH ratings as separate pathways for sources of biases to emerge: the health information (the source); the individual (the transaction); the survey (the outcome). More effective exploration of the sources of biases in SAH ratings will require casting the individual in different frame beyond the survey process, to understanding their health experience, health information they are exposed to.

The complexity of this exploration stems from the fact that SAH is a “general” self-reported question. Being a “general” self-report question, it can be postulated the SAH take on the characteristics of four types of questions mentioned in Tourangeau (1984); factual and behavioral questions; questions that assess of awareness and knowledge of subject; questions on attitude and opinions towards subject; questions that call for reasons or explanations. It is interesting to note when an individual is asked “In general how would you rate you own health?” it is in fact requiring the individual to: have some facts on their health and behavior, needing to have objective knowledge about their health, express their opinion or attitude towards their perceived health, rationalize it. It is in this diverse nature of the concept of a “general” self-report SAH question that potential contradictions may emerge in the 3 macro context of a SAH rating.

- When the “general” question taps into aspects involving actual factual behavior or knowledge, it is analogous to use of *observed objective health information* and *actual health experience*
- In contrast, when the “general” questions takes on the aspects of one’s attitude, opinion or rationalization of own health, this will lead concepts of *perceived subjective health information* and *health expectations*

As such, the multi varied dimensions in a “general” self-report question may breed internal conflicts and tensions, giving rise to contradictions in the observed versus perceived health information, the individual health experience versus their health expectations. On the survey front, there may also be differences between what is asked versus what is answered.

This thesis addresses the complex mechanism of this general self-report SAH question and propose 3 pathways to sources of biases in SAH ratings.

- A. The health information (the source): Observed versus perceived health information
- B. The individual (the process): Health experience versus health expectations
- C. The survey (the outcome): What is asked versus what is answered

First, the health information is the source of information that gets used for rating own health. Ideally, an optimal unbiased health rating can be achieved if objective health measures observed are equivalent to the individual’s perceived health. The sources of biases explored in this first health information pathway are: *availability of objective health information*, *subjectivity of health information*. Second, the individual is a carrier of these information which manifest as health experience and expectations. If there is no contradiction between how the individual experiences their health and their health expectations, then there may be lesser room for sources for bias in SAH rating to emerge from this path. The sources of

biases covered in this second individual pathway are: *over confidence, optimism and self-concept*. Third, the individual eventually expresses their perceived SAH in the survey and may face challenges making it difficult to accurately answer what is asked. The sources of biases associated with this third survey pathway are: *framing effects, social desirability, and satisficing*. Table 2 presents the framework used to explore the sources of biases in SAH in this thesis.

**Table 2: Framework for exploration of sources of biases in SAH ratings**

	Pathway A: Health information (The source)	Pathway B: The individual (The process)	Pathway C: The survey (The outcome)
The principle	Observed vs. Perceived health information	Health Experience vs. Health Expectations	Presentation vs. Answering of question
The sources of bias	<p><b>1. Availability of objective health information</b></p> <p>&lt;Heuristics: availability heuristics&gt;</p>	<p><b>3. Over confidence and optimism</b></p> <p>&lt;Heuristics: affect heuristics, anchoring heuristics, representative heuristic, symmetry rule heuristic, prevalence rule heuristic&gt;</p> <p>&lt;Biases: survivorship bias, conjunction fallacy, self-serving bias, self enhancement, over confidence bias, optimism bias, illusion of control, anchoring and insufficient adjustment, stereo typing, loss aversion, endowment effect, social comparison, false consensus effect, above average effect, omission error ,optimism, over confidence, self-concept &gt;</p>	<p><b>5. Framing effects</b></p> <p>&lt;Heuristics: framing effect, availability heuristic, confirmation heuristic, anchoring heuristics&gt;</p> <p>&lt;Biases: framing effect, response order - recency and primacy effect, question order effect, positivity bias, ambiguity effect, confirmation bias, feedback and learning effect, certainty effect &gt;</p>
	<p><b>2. Subjective health information</b></p> <p>&lt;Heuristics: affect heuristics, anchoring heuristics, representative heuristic&gt;</p> <p>&lt;Biases: anchoring effect and insufficient adjustment, social peer comparisons, subjectivity, bandwagon effect&gt;</p>	<p><b>4. Self concept</b></p> <p>&lt;Heuristics: availability heuristic, confirmation heuristic &gt;</p> <p>&lt;Biases: self-serving bias, self enhancement, social comparison, false consensus effect, above average effect, confirmation bias, omission error &gt;</p>	<p><b>6. Social desirability</b></p> <p>&lt;Biases: Social desirability bias, interviewer effect, courtesy bias, acquiescence, evasiveness, mood &gt;</p>
			<p><b>7. Satisficing</b></p> <p>&lt;Heuristics: availability heuristic, representative heuristic, anchoring heuristic, confirmation heuristic, automacity&gt;</p> <p>&lt;Biases: satisficing, status quo bias, anchoring and insufficient adjustment, emotional bias, consistency effect. mood, social desirability &gt;</p>

### 3.5 Definitions and Assumptions

The following definitions are used in this thesis:

#### 1. Definition of a bias in SAH ratings:

- What is a bias is tricky to define and there are many discussions around it. In this paper, the below perspective is adopted after consolidation of various definitions and contextualized in SAH ratings. A bias is about the use of irrelevant information and/or emotions and motivations and/or influence by moods driving wrong judgment/outcome (Kahneman, 2003) and contributing to different norms and expectations of health between individuals (Carr et al., 2001; Sen, 2002, Salomon et al., 2004; Dunning et al., 2004). Overall, it leads to a deviation from a normatively correct answer (Keren, 1990). Here the normatively correct answer will be one that fits rational assumptions. Biases in SAH may lead to invariant preferences and resulting in SAH ratings which cannot be compared across individuals, populations and studies (Kahneman, 2003; Salomon et al., 2004; Dunning et al., 2004).
- A clear distinction of heuristic versus bias is made in this thesis. As in Kahneman and Tversky Theory in judgement and decision making, heuristics and bias: Heuristics are defined as principles, mental shortcuts used to substitute the actual judgement attribute to reduce complexity of task and assessing probability (Kahneman, 2003). To qualify as a labelled bias, the attribute will have to had engaged use of judgment heuristics, or non-cognitive aspects like motivations and moods, and through this process, results in the use of irrelevant information, emotions and wrong judgment

#### 2. Definition of sources of bias in SAH ratings:

- Sources of biases in SAH ratings in this paper go beyond identifying true biases to also include effects/ heuristics/motivations/moods and conditions, violations which may / may not lead to actual biases.
- Sources of biases has been selected to be the focus in this thesis and not the actual bias. This reflects the complexity of SAH ratings and its interconnectedness between various elements. Working on the level of sources of biases facilitates a richer tapestry to enrich and provide a context/ecosystem for future research in biases in SAH ratings.

#### 3. Assumptions:

- Lack of ill health is generalized and assumed to mean good health, although it may or may not necessarily be the case. Violation and relaxation of this assumption can lead to bias in SAH ratings. Pathway 2 on individual health experience versus expectations examines potential source of biases, if this assumption is relaxed.
- SAH is often communicated as a general health rating at a point in time (not about future health). The researcher is unable to preclude and disentangle use of any future health fears or expectations in health ratings.



## Section 4: Sources of biases in SAH ratings

The sources of bias in each of these pathways will now be explored, by considering the environment, conditions, cognitive/ non cognitive process that induces biases and the actual bias description as proposed by Keren (1990) in identification of a bias.

### 4.1 Pathway 1: The Health information - observed vs. perceived health information

#### 4.1.1 The health information environment and conditions:

##### Observed versus perceived health

The SAH variable is a perception driven self-report variable. Different types of health information are used by different people to construct own SAH rating and an individual can tap into both objective and subjective health information in constructing this response. Observed health can be defined by its ability to be captured by objective health information. This means its measurement can be observable, even if it is not actually observed (Scheter, 1993). Self reports based on personal observation of own health conditions that can subsequently verified by physician is considered as an observed health from objective source of information (Scheter, 1993). Perceived health on the other hand includes various elements which creates a subjective health information dimension to SAH including aspect like: social determinants of health example education which may affect the type of information used and how people report health, social peer comparison, varied determinants of SAH, using an affective heuristic of a general sense of well-being about one's health and influence of moods.

Ideally, one's observed health corresponds perfectly with one's perceived health to avoid potential sources of bias in SAH ratings from emerging. However, in reality, the observed health typically does not equal to the perceived health.

##### Evidence of gap between observed versus perceived health

Although research has found SAH to be strongly correlated with objective health measure (Lee, 2014; Idler and Kasl, 1995, Scheter, 1993), studies have also shown SAH rating to comprise many other subjective dimensions beyond the objective health measure (Damain et al., 2008; Lee, 2014; Fienberg et al., 1985; Festinger, 1954; Van Doorn, 1999; Idler et al., 2004; Groot, 2000; Idler and Kasl, 1991; Wilcox et al., 1996; Mossey and Shapiro, 1992; Kaplan and Baron-Epel, 2003; Worsley, 1990; Segovia et al., 1989; Bailis et al., 2003; Smith et al., 1994; Crossley and Kennedy, 2002).

Additionally, whilst research has shown that socio determinants of health like education and income are used in assessing SAH, a different nature of studies about impact of socio economic factors on health gradients reveal evidence of the potential difference between observed versus perceived health information. Social economic health gradient of income and education were found to be insignificant in its association with perceived subjective self report of health, but significantly correlated with observed objective health measures (Lee, 2014; Mosca et al., 2013; Johnston et al., 2009).

The prevalence of subjectivity in perceived health is also evident in reporting heterogeneity studies. The subjectivity in SAH measures, due to the use of perceived subjective health information, gets translated into reporting heterogeneity of SAH. Reporting heterogeneity in SAH ratings can be described as the systematic difference (under or over estimating) in the mapping of true latent health into SAH reporting of SAH rating for a particular sub group. Different sub groups can may interpret the SAH rating scales 1 to 5 and labels like “very bad”, “fair”, and “very good” differently. This results in differences in stated SAH across different sub groups even if their true latent health were similar.

Studies conducted revealed reality of heterogeneity on self-assessed health rating. Methods like anchoring vignettes (Bago d’Uva et al., 2008), identification of cut point shifts (Lindeboom and Doorslaer, 2004) and construction of health indexes for improved objective health measures (Pfarr et al., 2012) have been used to show evidence of reporting heterogeneity across various socio-economic, demographic and cultural factors (Pfarr et al., 2012; d’Uva et al., 2007; King, 2004; Zimmer et al, 2000; Meijer et al, 2011). Bago d’Uva et al., (2008) found higher educated tend to under report health in developed Europe. Heterogeneity also exist in income level, age, rural urban divide in developing countries like India, China, Indonesia (Bago d’Uva et al., 2007) and country cultural characteristics (Pfarr et al., 2012). Social peer comparisons, justification of economic circumstance, access to health information or even inherent optimism (Bago d’Uva et al., 2007; Bago d’Uva, 2008; Pfarr et al., 2012) have been offered as potential explanations for the observed reporting heterogeneity.

The gap between one’s observed versus perceived health, due to the use of objective versus subjective health information, leads to SAH rating representing a significance of one’s own health. The alchemy of the various subjective elements are so intertwined that perhaps, one can only accept the “subjective truth” as termed by Baillon (2015), of a SAH rating. This “subjective truth” of the SAH rating needs to be taken account of and one must be careful in interpreting results. The below section addresses the drivers and sources of biases in SAH that creates this gap in observed versus perceived health for the individual.

#### **4.1.2 Drivers and Source of biases to SAH ratings: (Availability of objective health information)**

Preferably, the use of health information that is objective, relevant and up to date may enable a more unbiased SAH rating. Nevertheless, objective health information may not always be available to individuals due to lack of access, or ignorance of the need for objective health information due to limited resources used to deal with other priorities in life or the lack of trigger to think about one’s health.

Availability of objective health information may be constrained and influenced by various socioeconomic factors such as social class, income, occupation, education, community , gender, age, access to technology, geographical access to medical care, which all influences the condition people are born grow, live, work and age (World Health Organization). Need for privacy, social stigma, economic ability

to afford health care, geographic inaccessibility to medical care/ health check, lack of health care infrastructure and lack of knowledge are examples of things which may limit access to obtaining objective health information. With the lack of objective health information, information used for SAH ratings will be subjected to **availability heuristic** and may be subjective in nature.

#### **4.1.3 Drivers and Source of biases to SAH ratings: (Subjective health information)**

In reality, much of health information used for SAH assessment goes beyond objective health information like personal medical reports, to use of subjective health information. Subjectivity in health information comes from three key aspects: 1.the different definition and evaluation of good health status; 2. what constitutes a SAH rating to individuals; 3. Affective heuristic and influence of moods driving subjective perception

##### 1. Different definition and evaluation of good health status

###### **Social determinants of health:**

One way to understand how health information and its subjective assessment may be a source of bias in SAH rating, involves studying the determinants of SAH ratings. The definition and evaluation of good health may be influenced by the social determinants of health (Lee, 2014; Kawachi et al., 1999; Damain et al., 1999; Baron- Epel et al., 2008). Whilst SAH rating association with chronic illness and functional status like eating, dressing/undressing, going too toilets, walking, going outdoors are often found to be homogenous across studies, there are more heterogeneous findings with impact of variables like age, sex, marital status, education, social class, lifestyle habits on SAH rating (Damian et al., 1999). This heterogeneity of SAH in social determinants of health shows subjective health information being used due to differing expectations, understanding and evaluation of own health.

Social determinants of health like exposure to stress, social connectivity, unemployment, social support, access to transport, sanitation and technology shape subjective understanding and perception of own health via being the context for an individual to assess and form expectations of one's health. Individuals may peg their own health and expectations with social peer group and community health (Fienberg et al., 1985; Van Doorn, 1999; Idler et al., 2004).

This can be illustrated by the SAH rating of a poorer person versus a wealthier, more educated person who has higher expectations of own health. The lower income individual may have constrained food habits, lack of understanding of what may constitute healthy eating, living in a community with poor sanitation, transportation, and exposure to violence, higher unemployment and lack of health care access, thus giving a lower SAH rating. Their richer counterpart may be more affluent, educated, have access to infrastructure, social capital, trust, a better quality of life, better access to health care, information through internet forums, better understanding of the health knowledge and more empowered for self-diagnosis through information, overall leading to higher expectations and ratings of own health (Sen, 2002; Damain et al., 1999; World Health Organization and Calouste Gulbenkian Foundation,2014; Lee, 2014; Kawachi et al., 1999; Damain et al., 1999; Baron- Epel et al., 2008 ).

To some extent, the influence of social determinants of health as the above, and **social peer comparisons** create health expectations and norms in the community which can be associated with the **bandwagon effect** where watching others leads to conformity in expectations and norm setting. Subsequent, **anchoring** of their health state to this social context driven health norm may lead to **insufficient adjustment** of personal reflection of actual own health state, as well as **representative heuristic** where the individual may ignore base rates of there were more health/unhealthy people in the population that was not associated with their social conditions. Overall, leading to **social comparison bias**.

### **Personal threshold of health and health reference:**

What constitutes good health can mean different things to different people, depending on personal threshold and expectations. This leads to actual “good health” encompassing a wide range of definition and degree of acceptability: being without illness, no restrictions in activities, physical and mental health, “normal health” of having a slight cold and flu (Sturgis et al., 2001). With such a wide corridor of interpretation, personal definition of good health is paralleled to a having make a choice between this and that, having both, or neither. People may also not be aware of health issues which they may have (Sturgis et al., 2001). Moreover, SAH is too “general” a health status question and individuals may not know what frame of reference to take in rating SAH: physical, mental, emotional health (Lee, 2014; Fienberg et al., 1985).

### 2. Constitutions of SAH ratings

Beyond social determinants of health creating subjectivity in health assessment, respondent’s rating of “good health” SAH ratings also encompass other subjectivity drivers of health information: medical, sociological and psychological aspects. Studies on decoding what is used by the individual for their SAH ratings. found SAH variable to contain subjective information like: demographics, social economic advantage (Damain et al., 2008; Lee, 2014), sociological influence of social peer comparison (Fienberg et al., 1985; Festinger, 1954; Van Doorn, 1999; Idler et al., 2004), adaptation and normalization of illness (Groot, 2000; Van Doorn, 1999; Idler and Kasl, 1991; Wilcox et al., 1996; Mossey and Shapiro, 1992), used of family health history (Van Doorn, 1999), affective aspect of health such as general feeling of well-being, mood and level of energy (Kaplan and Baron –Epel, 2003; Worsley, 1990; Segovia et al., 1989), lack of illness and healthy lifestyle like eating health and regularly and avoidance of illness (Segovia et al., 1989; Worsley, 1990) and self-concept of health (Bailis et al., 2003; Smith et al., 1994).

This subjectivity is further complicated by the emphasis different populations take to assess own health. “Non- Whites” tend to use health problems, versus “Whites” who use physical functioning to determine SAH (Krause and Jay, 1994; Lee, 2014). Different aspects are also used to assess “bad health” versus “good health”. Smith et al., (1994) examined constructs of self-rated health amongst 2000 women in Australia aged 45 – 55 years old via its correlation with: beliefs and attitudes in health - such as attitude towards aging, menopause; current health status - through “symptomatology”; health related behaviour - like smoking, exercising, alcohol consumption and preventive health measures. Within this examination, bad health was found to be associated with illness in current health status. Still, self-perceived good health was less about the absence of ill health but correlates with socio demographic

factors like being older, divorced, widowed, working full time and preventive health measures like exercising, reflecting socio economic advantage and self-image.

### 3. Affective heuristic and mood driving subjective perception

The **affect heuristic** is activated sometimes as a mental short cut in the individual self assessed health process. Schechter (1993) used method of cognitive lab work and coding of open ended responses to develop a cognitive model of SAH. Schechter (1993) proposed model shows that health information used to evaluate health comprise two elements: cognitive aspect to SAH - actual knowledge of health conditions, daily health behaviour; affective aspects to SAH -sense of well-being and mood. Feelings and knowledge are ultimately reconciled and integrated to a final SAH response.

Furthermore, Kaplan et al., (2003) supported this by exploring three models of judging health: “biomedical or disease oriented”, “emotions or feeling oriented” and functional related. Their study showed respondents’ use of spontaneous emotional response on general feelings of health. “General feeling of health” – an affective variable was found to be very influential in subjective SAH rating with 76.2%, 67%, 62.8% of those who rated health excellent, good and sub optimal respectively stating its importance. This is in contrast to 55%, 29.3% and 79.4% of the same order for importance of objective health information in SAH rating derived from presence of disease diagnosed by doctors, treatments and medications on is one currently. Worsley (1990) and Segovia et al. (1989) also found the relevance of using general sense of well-being or general feelings of subjective level of energy to answer the question of SAH.

These affective responses are subjective in nature, and do not necessarily correspond to objective health measurements and may lead to **emotional bias**. An affective approach may also be less constant and presents challenge to accuracy of SAH rating as feelings change with time and if one compares health with sometime in past, own SAH rating may differ dependent on context (Andrew, McKennell, 1980; Smith, 1979; Fienberg et al, 1985)

**Moods** are may also be another potential source of influence on subjectivity of health information used in SAH ratings. In good mood, one may reduce the subjective weighting put on recent pain experience (Scheter, 1993). Good mood also leads to positivity bias in self reports of health (Hanita, 2000). The individual may also be influenced the effects on mood of previous questions, resulting in a positive or negative mood context for evaluation of SAH rating (Fienberg et al., 1985).

## 4.2 Pathway 2: The Individual – health experience versus health expectations

### 4.2.1 The individual health environment and conditions:

#### Health experience versus health expectations

Ideally, SAH ratings reflect the individual's actual health experience and can be verified objectively. However, in reality, one's health expectation rather than health experience may be used to rate SAH. This may arise due to actual lack of ill health experience, or if a person is unable to accept own health state and chooses to use personal belief of own health based on health expectation and may also potentially capture desired or aspired for health state. If an individual's health experience equals to their health expectations, then there will not be sources of biases to SAH. Should there be a gap between one's actual health experiences versus health expectations, and respondent reports health expectations instead of actual health experience in SAH, there may then be a bias to SAH ratings.

Drawing on health related quality of life studies which defines the gap between health experience and health expectations as quality of life (Carr et al., 2011), the gap may also be a source of bias for SAH ratings (e.g. a person with chronic illness may not see him/herself as having poor quality of life due to adaptation or own personal expectations of poor health). The line of thought serves well to be adapted to explore sources of biases in SAH rating driven by the gap between individual health experiences and health expectations.

In understanding what causes gap in experience and expectations, health experience is defined as an actual objective or observed health outcome that affects health status. Health experience, being objective in nature, can be used to support SAH being a continuous biological continuum concept of ranging from good to bad (Kaplan and Baron-Epel, 2003; Smith et al., 1994).

In contrast, using health expectations for SAH ratings may render SAH to be not continuous on a scale but each scale to involve different concepts of health (Kaplan and Baron-Epel, 2003; Smith et al., 1994). Health expectations may depend on the model where different aspects such as "medical, psychological, social, economic, spiritual" conditions are used by the individual to conceptualised own health (Larson, 1999; Lee, 2014). Additionally, it is also influenced by what the individual chooses to activate in rating own SAH. If the individual uses: 1. An enduring and static self-concept of health: rated using health expectations and self-concept of health. Actual changing health experience does not influence one's ratings (Lee, 2014; Bailis et al., 2003), or 2. A spontaneous view of health: rated using actual objective health experience and own health expectations does not influence ratings (Lee, 2014; Bailis et al., 2003)

Two sets of evidence exist to show the disparity and gap between health experience and health expectations. Both sets stem from understanding what constitutes a SAH rating. The first set of study is about understanding if there are elements in SAH that are time invariant. Bailis et al., (2003) examined the use of an enduring self-concept or spontaneous health assessment to rate health. Their study found evidence of use of both enduring self-concept of health and spontaneous health assessment to rate SAH. This shows the existence of a gap between actual health experience and health expectation. The second set of studies examines if the elements that constitutes good (optimal) and bad (suboptimal) health in SAH ratings are similar. Kaplan and Baron-Epel (2003) found actual ill health experience and

symptoms to be factors to influence sub-optimal SAH ratings, versus “general feeling” and ability to function daily to influence optimal good health SAH ratings. The authors suggest this to be indicate “personal evaluation of good and bad health to not be a continuity but two different concepts”. Beyond this, SAH ratings were also found to encompass factors of comparing oneself to social groups, which shows the use of health expectations and not just health experience in rating one’s health. Smith et al., (1994) conducted a targeted exploration to assess if SAH is biological continuum. The study looked at the differences in drivers between people who rate SAH better than peers versus people who rated SAH worse than peers. The intention was to use this construct to examine difference in factors used in rating good versus poor health. The findings indicate SAH ratings to not be a biological continuum but that good health and ill health are two separate concepts i.e. good health is not a mirror reflection of bad health. Whilst ill health were associated with ill health experience, good health go beyond not having ill health to include social demographic advantage and self-image. Reversing aspects of bad health is not in itself sufficient to guarantee good health though one may argue it is a necessary condition, vice versa. Overall, supporting ill health SAH ratings being associated with use of health experience but good health ratings being associated with one’s health expectations.

Triangulating the findings of these three studies: Bailis et al., (2003), Kaplan and Baron-Epel (2003) and Smith et al., (1994) provides support for respondent use of both enduring self-concept and spontaneous objective health measure to rate SAH, to reflect evidence that individuals as are not confined to only using actual objective health experience (reflecting use of only spontaneous objective health measures) to rate SAH, but also use of self - concept of health and health expectations. Overall, this use of health expectations and self-concept of health, beyond actual objective health experience explains the non-linear latent underlying utility of a SAH scale. Use of only objective health experience for rating SAH is more likely to support hypothesis of a linear biological continuum of good to bad health SAH rating, where there is an equal increment of the underlying utility as one moves up each rank of the rating scale i.e. the incremental health value the respondent put to rating of 1 to 2 versus 2 to 3 is equivalent.

Based on these studies what could potentially constitute health expectations include: social economic advantage, social peer comparisons, self-image, general feeling, ability to function daily, health improvement goals and health behavioral intentions. The below section will address the sources of biases that potentially lead to the gap between health experience and health expectation. In this particular instance, this can to some extent, be thought of as the drivers to health expectations.

#### **4.2.2 Drivers and Source of biases to SAH ratings: (Self-concept)**

##### What is self-concept of health and its role in SAH ratings?

Lee (2014) introduces the idea of viewing SAH as a form of self-concept of health. Self-concept of health is a more personal and internalized source of information used for SAH rating. It reflects a prior belief a person have about one’s health and allowing one to have a sense of continuity about one’s health (Borawski et al., 1996; Van Doorn, 1999; Schechter, 1993, Bailis et al., 2003). Using self-concept of health to rate SAH involves using one’s health expectation, instead of just using actual health experience. This implies use of self-concept of health in SAH as a potential source of bias in ratings as it may not reflect true health state. Self-concept of health has relevance and particular importance at the

stage of comprehension of SAH question where the respondent decides on which model of own self-concept of health to use to understand the question (Lee, 2014; Scheter 1993).

The psychological aspects and field of self-concept is complex and further understanding of this requires specialized knowledge and learnings. This thesis will draw only on some relevant top line definitions of self-concept for inspiration on what can be learnt, in defining a health self-concept in SAH rating. Most generally, a person's self-concept is defined as "how an individual sees him/herself, how others actually see the individual, and how the individual perceives others to see him or her" (Rosenberg, 1981; Lee, 2014). Many other definitions of the self-concept construct exist and these can be broadly summed up as involving the below key characteristics for self-generalization: self-image, assessment of own areas of strengths and competencies, self-worth, personal value driven, comparisons of self versus others (Markus and Wurf, 1987; Eccles et al., 1989; Marsh, 2004).

To understand how self-concept of health can be used for SAH rating, these general self-concept characteristics mentioned are applied and captured in various models of health conceptualizations. Lee (2014) and Larson (1999) introduces four models of health conceptualization used for defining self-concept of health. First: the medical view of health where one uses actual health experience of ill health to form own concept of health assessment. Second: the WHO wholistic approach of health conceptualization which includes physical, social, mental well-being elements. Third and fourth: wellness and environmental model of health. Overall, these models allowed for broad based conceptualization of health beyond actual ill health experience and are differentiated by the degree of external context such as social, economic spiritual condition in influencing perceived SAH.

This literature review found only 2 studies with evidence of use of self-concept of health in SAH, of which one is more indicative than descriptive. A cognitive interview study carried out by the Questionnaire Design Research Lab at the US National Centre for Health Statistics to understand cognitive aspects of SAH ratings reported that the fast speed at which subjects were able to justify their "good" and "excellent" health ratings indicates the prospect for subjects to already have a preconceived notion of own health status - also known as self-concept (Schechter, 1993). The second more concrete study was carried out by Bailis et al., (2003) to explore what constitutes SAH status with the hypothesis that it mirrors the information a person have about their health and wellness, as well as reflecting the prior beliefs a person have of their own health identity termed as "enduring self-concept". Their findings showed whilst spontaneous assessment of SAH driven by changes in health practice and objective health measures change do drive SAH ratings, "enduring self-concept" is also a significant part of SAH rating. This was evident from respondent's SAH rating in 1994 accounting for 12% of total variance in their 1996 SAH rating.

Beyond these, study on use of self-concept of health for SAH are still nascent and primarily indirectly explored through discerning what constitutes SAH ratings and its subjectivity. In this instance, findings were in line with and reflect Larson's (1999) health conceptualization models of use of medical, social, psychological, economic, spiritual conditions for health ratings. Studies have found SAH ratings to be influenced by aspects like social economic advantage, social peer comparisons, self-image, general feeling of well-being, ability to function daily, health improvement goals/ health behavioral intentions/



perceived control over one's health from healthy living (Van Doorn, 1999; Idler, 1993; Kaplan and Baron-Epel, 2003; Smith et al., 1994; Segovia et al., 1989; Worsley, 1990; Bailis et al., 2003)

#### When does self-concept of health in SAH ratings get triggered?

To understand what may triggers activation of own health self-concept in SAH rating, the field of psychology offers some insight. Taylor and Brown (1988) notes that a person processes information like a "naïve scientist" who observes the world with lots of missing data, errors and bias. The cognitive drain or lack of information leads to difficulties in decision making which requires deliberation (DeWall et al., 2008). In this weakened state, Bargh (1994) introduces an automatic mode being triggered when one is unaware, seeking efficiency, or unable to control intent or outcome. Cialdini and Goldstein (2004) and Janssen et al., (2008) also advocate the role of automaticity as a heuristic for efficient decision making. Automatic process include use of "familiar attitude, automatic evaluation and emotions, automatic trait, non-conscious behavioral mimicry" (Chartrand, 2005). In this instance, the automatic response triggers activation of one's self-concept made up of prior expectations, beliefs or self-serving information as a coping mechanism and heuristic in judgement and decision making.

Relating psychology field back to SAH ratings, long surveys, difficult questions, lack of information on own health may deplete cognitive power, possibly trigger of an automatic response of using one's self-concept of health as an **affect heuristic** and becomes default option used in SAH rating. Hanita (2000) refers to the use of personal values: medical factors (morbidity, life expectancy), self-image value (physical outlook, desired portrayal), social value (acceptance by society, role playing), effort value (lifestyle habits and initiatives) to evaluate SAH and eventually the most accessible set of personal values are used to assess SAH. Furthermore, it is found that information and memory that are ego centric (Greenwald, 1980), positive or assessed in good moods tend to be recalled and retrieved faster by the memory (Hanita, 2000; Greenwald, 1980; Taylor and Brown, 1998). One's self concept of health can be postulated to be ego driven and self-enhancing, thus positive and more easily triggered as a heuristic for rating SAH.

#### Why is the use of self-concept of health a potential source of bias in SAH ratings?

The use of self-concept in SAH rating can become a tool which individuals use to maintain, create or defend one's health status. Activating self-concept of health in SAH potentially leads to **self-serving bias** and sense of **illusionary superiority**. Literature in the psychology arena reveals this need to different and better than others to be emergent in five areas: recall of positive information about oneself more easily than negative ones (Taylor and Brown, 1988), marginalization of own weakness is as a common ailment to all people, but accentuate one's strength as heroic and distinctive (Campbell, 1986), attribute success internally and failure externally (Marsh, 2004) prioritize things which they are more competent at, as being more important (Campbell, 1986), and see self in more positive terms versus how others see them or how they see others (Levinsohn et al., 1980).

Hence, adopting self-concept of health confines one's perspective to own strength, positive belief about self and may lead to information neglect for a more balanced evaluation (Dunning et al., 2004; Fienberg et al., 1985; Bazerman and Moore, 2009). Self-concept of health to facilitate maintenance of positive

belief on one's health via selection of context for social comparison was found by Kaplan and Baron-Epel (2003). Individuals may also elevate one's status by creating a perceived advantage of own health versus others through one's socio-demographic advantage and self-image (Smith et al., 1994). This is classified by Dunning et al., (2004) as the above average effect due to information deficit where the individual identifies self to be better than average and lack the competence to know when they are making a mistake.

Self-concept of health also has a **defensive nature** which permits a person to "transcend" their health problems by qualifying it with "but I am doing fine" (Borawski et al., 1996; Van Doorn, 1999) also cues the existence of use of health self-concept in SAH. Moreover, once triggered, **endowment effect** (Bazerman and Moore, 2009; Tversky and Kahneman, 1978) through the preservation of one's self concept kicks in, especially if one has fought hard to build a healthy self-image. Studies also found people's needs to conform to one's personal identity (Bénabou and Tirole, 2005) and being consistent in one's portrayal of self through invariance of initial response given in an interview is an important part of this confirmation of personal identity (Burger, 1999). Interestingly, the afflicted, depressed and ill were found to be more balanced and realistic in their construct of self-concept compared to the healthy (Taylor and Brown, 1988).

The use of self-concept of health in SAH rating is a softer and more humanistic, albeit potentially bias approach (versus objective health measures) to intercede goal attainment of desired health outcome or own self health image. This raises questions on how reliable and accurate is SAH in reflecting one's true health status. Research on evidence and impact of the use of self-concept of health for SAH is currently limited, and much can still be explored to understand its relevance as a potential source of bias for SAH ratings.

#### **4.2.3 Drivers and Source of biases to SAH ratings: (Over confidence and optimism)**

Optimism and over confidence in SAH ratings can be defined as the difference in over rating of one's perceived SAH versus an objective health measure of it. Both of these are potential sources of biases in SAH ratings that leads to a gap between health experience and health expectations. There are key two scenarios which may drive over confidence and optimism over one's health.

- Scenario 1: One has ill health experience but has adapted to the ill health experience through normalization of illness and sense of survivorship and health transcendence, especially so in older, and people with ill health (Idler, 1993; 2004; Van Doorn, 1999; Wilcox et al., 1996; Carr et al., 2001; Leventhal, 1984; Borawski et al., 1996).
- Scenario 2: Lack of ill health experience leads to high expectations about one's good health (Dunning et al., 2004; Kaplan and Baron-Epel, 2003; Smith et al., 1994)

In the uncertainty of own health state without information on ill health, one seeks to construct own perceived health experience via the use of surrogate information to replicate and make real what one imagined one's health could be. Respondents may use information **anchors** like social peer comparisons, family longevity and normalization of ill health to assess own health (Van Doorn, 1999) leading to

**representative bias** (Kahneman and Tversky, 1972; Tversky and Kahneman, 1974; Grether, 1980). The use of above stated health information forms the basis of perception of own probability of ill health, ignoring any potential base rate or priori of ill health that may exist, resulting in inaccurate judgement about probability of ill health in uncertainty. This, potentially fostering a false sense of security driving over confidence and optimism in one's own health rating with the sense of "this ill health will not happen to me".

Additionally, in the absence of ill health experience, the difficulty of proving one's health is good is substituted by evidence to show one's health is not bad. This sense of **illusion of control** was introduced by Langer (1975) where "factors of skill situations (competition, choice, familiarity, involvement) introduced into chance situations cause the individual to feel inappropriately confident". People fail to judge the lack of uncertainty and this gets expressed as what Langer (1975) termed as a "just world" belief, where good deeds begets goodness, vice versa, and everyone gets what they deserve.

**Optimism** and **over confidence** may occur when one have a healthy lifestyle and health behaviour and SAH ratings are seen to comprise health lifestyle behaviour. (Smith et al., 1994; Segovia et al., 1989; Worsley, 1990). One may also see own good health behaviour and healthy lifestyle as being in control of one's health (Smith et al., 1994; Segovia et al., 1989; Krause and Jay, 1994; Kaplan and Baron-Epel, 2003). The greater the perceived control, the stronger the optimism and over confidence bias (Dunning et al., 2004) . This sense of **perceived self-control** through healthy lifestyle behaviour may be more prevalent amongst younger than older respondents as older people tend to use health problems as reference for health ratings, but younger respondents are inclined to rely on health behaviours for SAH (Krause and Jay, 1994; Kaplan and Baron-Epel, 2003).

In the event, where no obvious ill health symptoms have emerge (irrespective of actual health state), individuals may be incline to be unrealistically optimistic about health risk and under rate own risk to physical disease compared to other people. It may also reduce the individual's motivation to engage in further illness prevention like exercising and health checks (Dunning et al., 2004; Weinstein, 1983; Helweg-Larsen & Shepperd, 2001). Dunning et al., (2004) explains this possibility of unrealistic optimism being a source of bias in SAH through the use of four heuristics; 1. **Symmetry rule heuristic**: no illness and no symptoms of ill health is indeed about good health. However, over use of it increases probability of diagnostic error of serious illness with no symptoms, e.g. hypertension (Dunning et al., 2004); 2. **Prevalence rule heuristic**: some individuals see only rare unusual health symptoms as serious and common place symptoms like cold may be seen as mild and does to connote ill health (Dunning et al., 2004); 3. **Stress-illness rule**: Individuals may subjected to **conjunction fallacy** where one thinks certain conditions are probable than specific ones and misunderstand symptoms to correspond and attribute them to external events like stress or inconvenience instead of identifying with ill health, thereby missing actual cues of potential ill health (Dunning et al., 2004) E.g. having a headache, tiredness during exams, may be actual health symptom of chronic fatigue; 4. **Stereotyping**: One may think I that specific health conditions are more associated with certain stereotypes e.g. cardiac diseases are more associated with men, thus when women have these symptoms, they tend to brush it aside (Dunning et al., 2004). Overall, lack of feedback in SAH ratings may also lead to **confirmation bias** (Dunning et al., 2004).

Another set of drivers to **optimism** and **over confidence** is **adaptation** and **survivorship bias** of having adapted and conquered illness (Idler, 1993; 2004; Van Doorn, 1999; Wilcox et al., 1996; Carr et al., 2001; Leventhal, 1984; Borawski et al., 1996). The sense of having conquered the illness leads to information deficit where the individual does not account for what they do not know about the unknown future or fail to see the power of fear and anxiety in shaping the future, as well as neglecting worse case scenarios (Dunning et al., 2004). Moreover, Individuals may also lower expectations of own health, as one comes into reality with ill health experience (Carr et al., 2001). Thus, under the influence of optimism, SAH rating to be better than what one's health really maybe.

#### Evidence of over confidence and optimism bias in SAH

Optimism and over confidence in SAH ratings can be defined as the difference in over rating of one's perceived SAH versus an objective health measure of it. This can be detected in. Maddox and Douglass (1973) examined this issue of optimism through studying the incongruence of physician versus patient's rating, or two different sets of attitude emerging from same group of ill health respondents. They found discrepancies in physicians' versus respondents' ratings with optimism outnumbering pessimism. 24% of people were found to over-estimate own health compared to physician's rating. This was further supported by cases where the respondent's health declined over initial period and yet 71% kept that positive SAH rating, compared to a lower physician diagnosis of actual health state. In contrast, only 23% showed under rating of own health even if physician indicated improved health.

In a Qualitative study via in depth interviews of 48 respondents aged over than 65 years old, Van Doorn (1999) found existence of two different attitudes of optimism versus realism amongst people who have the same objective health state of serious health problems. 57.7% of respondents were classified to be health optimist defined as on having one or more health problems yet classifying own health as excellent, very good and good. This is relative to 42.3% of respondents defined as poor health realist who rated own health as poor or fair, in line with actual poor health state.

Social comparisons is another area which is used to examine the presence of optimism in SAH ratings. Comparisons are often used for self-enhancement to make one feel "advantaged" (Van Doorn, 1999; Markus and Cross, 1990; Idler et al., 2004; Fienberg et al., 1985). Benchmark selected for comparisons differ across old versus young people, healthy versus sick people, and range from comparing with: other similar sick people, social peers, similar aged peers, and friends. Respondents identified as health optimist by Van Doorn (1999) were seen to make more positive social comparison of self versus others, compared to health realist who made more neutral or negative social comparisons. Breast cancer and arthritis also show those with these illness comparing self with others worse off to enhance health status (Idler et al., 2004).

In exploring how social comparison works across age demographics, reference group theory was applied to ask people aged 65 years old and above to rate their own SAH compared to others of their age (Idler, 1993). "Old-old" was found to be more optimistic than "young-old", seeing their own health to be better than their peers. Kaplan and Baron-Epel (2003) studied subjectivity of SAH measures amongst 383 Israelis, aged 18 to 60+ years old and found healthy young people comparing themselves with their

friends and use general feelings of health to evaluate health status, versus older people who are unhealthy, tend to compare themselves with peers of their age to justify SAH ratings.

The difference between optimism and younger versus older people can be attributed to their points of comparisons and life experience. In contrast to younger health people, where optimism has to be proven by benchmarking with friends, and perceived self-control of health lifestyle for good health, optimism in older people reflects the individual's journey in life, has a sense of rationality and being of a well-deserved reward to the older people. Optimism in older people are driven by sense of having conquered death if they live beyond typical family longevity age, normalization of illness and comparing own success with health versus other similar healthy/unhealthy peers (Van Doorn, 1999; Idler, 1993). Other reasons supporting old-old people to be more optimistic in their SAH rating is the adoption of a more accepting attitude towards life and living with less anger and fear of illness compared to the young-old (Leventhal, 1984)

### **4.3 Pathway 3: The survey – Presentation versus Answering of question**

#### **4.3.1 The survey environment and conditions:**

##### Gap between questions asked versus questions answered

Question context may lead to a gap between what is being asked versus what is being answered. A respondent may be influenced by multi-faceted question context effects like: framing of the questions, external survey environment such presence of interviewers, as well as personal interaction and interest in answering of the survey (Lee, 2014; Crossley and Kennedy, 2002; Schwarz, 1999).

The interplay of these aspects may create a gap and lead to a difference between what is being asked, versus what is eventually answered. The section explores the various drivers to sources of bias in SAH ratings due to the different causes and outcomes of question context. Three key drivers to influence of question context will be examined: framing effects, social desirability and satisficing.

#### **4.3.2 Drivers and Source of biases to SAH ratings: (Framing effects)**

The way a question is asked sets what Tversky and Kahneman (1981) termed as “decision frame”. Survey and scale is a form of choice problem where different framings of the choice problem can lead to different outcomes. As in life, surveys and answering of scales often requires one to apply rules like - is/ is not, and logic like - up or down, more or less. However, in reality, survey responses may not be as straightforward. They are typically subjected to framing effects, which increases sensitivity to the context formed by the way the question is asked. Typical survey response frames which affects the question context include: wording and labeling, scale frequencies, question order, response order and range of scales (Krosnick, 1991; Krosnick and Alwin, 1987; Toungangeau, 1999; Greene et al., 2014).

Kahneman (2003) describes framing effect as something that influences the accessibility and salience of a particular feature affecting choice decision making, leading to varying preferences dependent on the frame of reference. Framing effects influence the respondent's survey response experience. Framing cast the perspective of what is meant to be seen versus unseen and set about what Schwarz (1999)

termed as the “conversational process involved in question answering process” in helping respondents to interpret the questions and infer what is expected of them. Framing effects can also affect the respondent’s final response to the survey as they cope with limited cognitive power, imperfect information and the cost of thinking (Schechter, 1993; Schechter and Herrmann, 1997; Schaeffer and Presser, 2003; Toungangeau, 1999; Toungangeau and Rasinski, 1988; Sturgis et al., 2001; Schwarz, 1999; Krosnick, 1999; Krosnick and Alwin, 1987; Krosnick and Presser, 2010; Krosnick, 1991; Fienberg et al., 1985; Bowling and Windsor; 2008). Under the influence of framing effects, one may **anchor** on the frame and **availability heuristic** may get activated as a mental short cut, where certain information become more easily accessible and available to the respondent. This enables specific areas to be highlighted as a focus for the respondent in four stages of comprehension, retrieval of memory, judgment and reporting (as described in section 3.3.), thereby enabling efficient information processing and comprehension, and gets used as an anchor in answering SAH question.

The existence and role of framing in survey can best be convinced by adopting a deprivation technique of comparing two outcomes of the existence of framing vs its non-existence. This can be examined by using a potentially two extremes of framed versus unframed questions. An open ended question is good representation of a “question without a frame”. Putting the answer on a list to be chosen as response is a nice “framed question” counterpart to the fluid nature of an open ended question. Schwarz (1999) quotes Schuman and Presser (1981) study on questions and answers in attitude surveys, which shows evidence of spontaneous open ended questions yielding smaller total responses than if the same answers were given in a list for selection. The finding was attributed to the cost of thinking and “autobiographical memory recall”.

This section primarily seeks to explore the how framing may be a source of bias in surveys and self-assessed health ratings. As much of framing effects has been explored in the previous section 3.3, through the lens of cognitive aspects in the four stage survey methodology of comprehension, retrieval, judgement and reporting. As such, this section serves to highlight the various specific drivers to framing effects in SAH and to discuss which of these may be more relevant with respect to SAH ratings.

### **Framing effects in SAH**

Framing effects in SAH ratings exist as the SAH question typically takes different variations. Sturgis et al., (2001) pointed out the commonly found variations in SAH questions; *no natural point of reference* - “How would you rate your overall health / How would you rate your general health on a scale of 1 to 5” (Health Survey for England); *time reference* - “Over the last 12 months would you say your health has been” (General Household Survey); *people reference* - “For someone of your age, how good would you rate your health?” (Health and Lifestyle Survey); *“frequency” scale, rating scale reference* - “Would you say your health is 1. Very good, 2. Good, 3. Fair, 4. Bad, 5. Very Bad” (World Health Organization) and “Would you say your health is 1. Excellent, 2. Very Good, 3. Good, 4. Fair, 5. Poor” (US Health and Retirement Study). Here, five potential different frames are explored for SAH ratings.

### **Type 1: Wording and labelling**

Framing effects from wording and labelling in surveys is something that is widely discussed (Baumeister et al., 1998; Bowling and Windsor; 2008; Conti and Pudney, 2011; Crossley and Kennedy, 2002; Cronbach, 1950; Fienberg et al., 1985; Garland, 1991; Greene et al., 2014; Hofstadter 1985, p. g. 565; Janssen et al., 2008; Krosnick, 1991; Krosnick and Alwin, 1987; Lumsdaine and Exterkate ,2013; Matell and Jacoby, 1972; O’Muircheartaigh et al., 2000; Schechter, 1993; Schwarz, 1999; Sturgis et al., 2001;

Symonds, 1924; Schechter and Herrmann, 1997; Schaeffer and Presser, 2003; Tourangeau, 1984; Toungageau et al., 1991; Toungageau, 1999; Tversky and Kahneman, 1981, 1979; Weng, 2004; Worchester and Burns, 1975). Wording and labelling framing effects can go beyond language ambiguity and a random error, to creating frames that can cast the problem in a different light, skewing more positive or negative responses. The words and labels attached to a numeric value on a scale makes a difference and respondents do not select based on just a numerical choice (Lee, 2014; Bradburn, 1983; Tourangeau, 1984; Schechter, 1993; Schwarz, 1999). Schwarz et al., (1991) found using the same label of “not at all successful” to represent “-5” versus using it to represent “1” on scale of “-5 to 5” and “1 to 10” respectively, bring different meanings to the phrase “not at all successful”. Labelling “-5” as “not at all successful” connotes having actual failures in life versus labelling “1” as “not at all successful” has softer impact of not having any achievements.

There is strong evidence of wording and labelling framing effects in SAH ratings. An unbalanced positively skewed US worded SAH scale “excellent, very good, good, fair, bad” may potentially lead to additive systematic error (Viswanathan, 2005), in this case a **positivity bias**. Bowling and Winsor (2008) found people tend to be more positive on the US SAH scale in a study amongst 11,000 respondents, aged 50 and above in the English Longitudinal Study of Aging. Labelling of first response on scale as “excellent” versus “very good” on SAH rating scale tend to elicit more optimal (“excellent and very good” combined) response (41%) amongst respondents who were shown scale of “excellent, very good, good, fair, bad” versus those who were shown scale of “very good, good, fair, bad, very bad” (26%). Lumsdaine and Exterkate (2013) also examined word and numerical concordance in using two scale “very good”, “good”, “fair”, “bad” (WHO scale), “very bad” versus “excellent”, “very good”, “good”, “fair”, “poor” (US scale) by showing them twice. The impact of different wording labels was reflect in weighted average of 2.34 for the WHO scale and 3.03 for the US scale. 55.7% of the sample are word concordant and 30.6% of the sample are numeric concordant between the two scales.

Evidence of wording and labelling effect was supported by Greene et al. (2014), whose study showed the mean response score of differently labelled scales skewing towards a middle inflated response. However, the mean response score of both the scales were different; the mean response score of “very good, good, fair, bad, very bad” scale is 2.34 - between good and fair; the mean response score of “excellent, very good, good, fair, bad” is 3 – good. Baron-Epel and Kaplan (2001) also found wording and labelling effects in older people who rate SAH better when asked “compared with others of your age, would you say your health is excellent, very good, good, fair or poor” versus if they were exposed to the standard version of SAH question of “would you say your health is excellent, very good, good, fair or poor”.

Overall, the labelling of words can affect the set of sample considerations respondents use to edit their responses to connect their response to the correct answer category (Lee, 2014; Jylha, 1994; Tourangeau, 1984). However, research have also found two dynamics that limit the impact of wording and labelling framing effects. There exists a cushioning effect of gravitation towards “middle or average” responses: “good” rather than “very good”, and “very good” rather than “excellent” tend to be chosen (Greene et al., 2014; Bowling and Windsor, 2008). Second, wording and labelling effect is found to be less prominent in “sub-optimal” health range (poor/bad/very bad) and the increase of response category to “bad and very bad” instead of just “poor” on a scale does not encourage more negative SAH response or correct for the optimism bias (Bowling and Windsor, 2008).

Another aspect of wording and labelling framing effects revolves around how fully labelled and clear the wordings are. Scales that are fully labelled for all response categories were also found to be more reliable than scales that labelled only at the end points (Weng, 2004; Krosnick, 1999). Conti and Pudney (2011) studied the impact of partially (“1 - not at all satisfied, “4 - neither satisfied nor dissatisfied, “7 - completely satisfied”) versus fully labelled (“1 - not at all satisfied, “2 - mostly satisfied” , “3 - somewhat dissatisfied”, “4 - neither satisfied nor dissatisfied, “5 - somewhat satisfied” , “6 - mostly satisfied” , “7 - completely satisfied”) scales on job satisfaction and found large changes in response distribution between the two types of scales, with the biggest distortion recorded at points where the two scales were similarly fully labelled i.e. 1, 4, and 7.

Additionally, clearly labelled questions help reduce doubts and close the gap in response differences due to use of different frequency scales (Schwarz, 1999). Fluidly worded responses like “responsiveness to change of weather” leads to 75% of patients stating symptoms frequency of more than twice a month on a high frequency scale of “twice a month or less to several times day”; in contrast, only 21% had the same report when exposed to a low frequency scale of “never to more than twice a month” (Schwarz, 1999). However, more narrowly and clearly defined/worded labels like “excessive perspiration” leads to a smaller range of difference of symptom reporting of 8% between high and low frequency scales. (Schwarz, 1999).

### **Type 2: Question order**

Question order effects are about the potential influence of prior questions in setting the context, influencing and ensuing judgment on later question response (Schwarz, 1999). People have multiple beliefs encapsulated in schemas formed about certain issues about their state of health (i.e. schemas are mental constructs or conceptual framework which people use to represent and understand various experiences. Different sets of schemas co-exist and the choice of schemas selected to be used in understanding experience or solving problems is dependent on what the issue is about). Prior questions can influence the choice of the winning mental schema being used via evoking an affective evaluation that lingers to secure the heuristic attribute for the subsequent SAH question (Kahneman, 2004; Schwarz, 1999). Preceding questions get assimilated/included or contrasted/excluded in judgement of subsequent questions (Krosnick and Alwin, 1987; Tourangeau, 1991; Bowling and Windsor, 2008; Schwarz, 1999; Lee, 2014).

Questions presented first have a first mover advantage and allow them to set what Schwarz (1999) later termed in his ‘inclusion-exclusion’ model as mental representations. In this model, question order framing effects provide context to aid relevant memory call through creation of “mental representations” by serving as information which is “temporarily accessible” but also subsequently becomes a “representative” benchmark which can be “assimilated or contrasted” for evaluation of one’s final response to the survey question (Schwarz, 1999). “Subtraction hypothesis” was also used by Tourangeau et al., (1991) to test question order effects of asking marriage happiness or general happiness first. Asking general happiness question after marriage happiness, results in a low correlation of 0.2777 with marriage. In contrast, asking general happiness first before marriage questions results in a higher correlation between the two of 0.545.

**Question order effect** has important implications on where SAH question should be placed: upfront for spontaneous response, or latter for learning effect and more cognitive thoughts. Bowling and Windsor (2008) studied the impact of question order effect on SAH rating by randomising two different sets of SAH scales (starting with “excellent” versus “very good”), shown twice to each respondent. Asking SAH after other health related questions was found to lead to more optimal (coded “excellent” or “very good”) SAH reporting. 26% of respondents coded “very good” when SAH question was asked at the



beginning versus 31% of respondents coded “very good” when the question was asked at the end. People were more optimistic with 38% rating optimal health (“excellent and very good”) when shown the scale after completing other health questions versus 33% when SAH rating was obtained upfront. The caveat was a lack of control group in their experiment, making it difficult to identify any potential bias of including two health status measures in a questionnaire (Bowling and Windsor, 2008).

In a comparable way, Lumsdaine and Exterkate (2012) using data from eleven European countries SHARE survey, discovered 41.8 % of respondents changing their SAH response when first shown the WHO SAH scale before the US SAH scale, versus only 28% changed their SAH rating when shown the US SAH version first. The study was however not able to disentangle wording and labelling effect versus question order effect of this change. In another study amongst Australians, Crossley and Kennedy (2002) reported 28% of respondents changing their SAH rating when they were presented with the same SAH scale the second time after a set of health related questions. Learning and feedback effects due to question response order were also offered as explanations in both these studies. Thus far, current studies have found relatively strong evidence of question order framing effects in SAH ratings. Recommendations have been made on presenting the general SAH question upfront to get an untainted spontaneous view of own self-reported health status.

### **Type 3: Response order**

Response order effect is about respondents not giving all response order equal consideration (Holbrook, 2008). Distribution of response is influenced by the order in which it is presented at the beginning or end as options to the respondent and items at both end of the scale tend to more likely to be chosen (Holbrook, 2008). Research has shown items read first to be recorded in long term memory versus those read last are in short term memory, resulting in activation of response order effect with both items at the beginning or at the end to be more likely to be chosen (Krosnick and Alwin, 1987).

More concretely, learning from studies on impressions, information flow and persuasiveness, Krosnick and Alwin (1987) theorized response order effect through the formulation of two outcomes of primacy and recency effects. **Primacy effect** gets activated when items that are listed first have higher probability of being chosen, whereas, items listed later have higher chance of being selected in **recency effect**.

**Primacy effect** is identified as being most prominent when: question is exposed in visual show cards, there is a long questionnaire, long list of responses to choose from (Krosnick and Alwin, 1987). One is perceived to have more cognitive resource of a fresh mind when interacting with information first presented, thus exposed to “deeper cognitive processing (Krosnick and Alwin, 1987). Krosnick and Alwin (1987) found clear evidence of primacy effect in an experiment with a list of attributes with regards to child qualities where attributes like “manners, honest, success, obey, considerate, interested, studious” that were upfront on the list more likely to be picked as the three most important qualities in a child, compared to when the list was presented in a reverse order and response margins between the two lists are as big as 17%. Items in the middle like “clean, amicable, control” did not change much between the two orders of list. The primacy effect was shown only to be significant mainly amongst those with lower cognitive power (measured by education level and word test) and amongst people who have not formed a clear point of view on the issue (Krosnick and Alwin, 1987).

**Primacy and recency effect** has been attributed to the act of satisficing where respondents seek the first acceptable answer rather than processing all available for an optimal response (Simon, 1957; Krosnick and Alwin, 1987). Length of questionnaire or response list options may create fatigue and make it

difficult for memory recall, triggering an easy solution through primacy and recency effect (Krosnick and Alwin, 1987; Toungangeau, 1984; Hermann, 1995; Schechter and Hermann, 1997; Schwarz, 1999; Janssen et al., 2008; Baumeister et al., 1998).

The format in which questions are asked may also encourage response order effects. When questions are presented orally, respondents have less time to process the first question and more like to fall into recency effect. In contrast, primacy effect was found to be more prominent when questions are presented visually (Krosnick and Alwin; 1987; Janssen et al., 2008; Baumeister et al., 1998; Schwarz, 1999).

This literature review has not found studies examining response order effect in SAH ratings. However, one can postulate the potential for SAH ratings being subjected to primacy and recency effect. Perhaps, this could be related to respondents' lack of clear definitive opinion on own health state, depleting cognitive power and subjected to primacy or recency effect. Respondents may also be subjected to primacy and recency effect in SAH ratings from different modes of interview of visual self completion versus aural telephony interviews. Holbrook (2008) suggest questions where responses options fall along a scale such as , "extremely likely, very likely, somewhat likely, slightly likely, or not at all likely", the response order will fall into a logical order and can be examined by looking at distribution changes if scales was reversed from left to right, vice versa. Accordingly, further research can be conducted to examine the impact of rotation of response order and mode of interview on distribution of SAH ratings, signaling primacy and recency effects in SAH ratings.

#### **Type 4: Mid points and number of response categories**

The number of response categories offered on a scale provides a frame which may influence respondent's ratings. Having a wide range of response options allow respondents to more accurately formulate and associated themselves to the relevant category, restricting potential loss of information (Cronbach, 1950). Scales with more response categories are also found to be more reliable than scales with fewer response categories (Weng, 2004; Cronbach, 1950).

Mid-points plays a more important role when there is a small range of response choice on the scale and is often a fail-safe option (Garland, 1991). Matell and Jacoby (1972) portrayed the framing effect of a mid-point through a substantial 20% of respondents picking the mid-point if presented a scale with small range from three to five points scale, but decreases to only 7% picking the mid-point when the number of response option increases to seven to nineteen in a scale increases. In so far as using higher order response scale as a tool to facilitate more accurate judgement, having more options does not guarantee a person's ability to better discriminate as one may still be subjected to choice paralysis and one's psychological ability to differentiate (Symonds, 1924).

Currently, the number of response categories framing effect is potentially less applicable to SAH ratings as it is typically on a 5 point scale with the exception of General Household Survey (UK, 1971 to current) which has a 3 point scale of how one sees own health in last 12 months ranging from " good, fairly good to not good". However, findings by Green et al., (2014) on inflated "middle" response in SAH ratings, as well as Bowling and Windsor (2008) findings on smaller "ceiling effects" of scales starting with "excellent" versus "very good" can be starting points to think about the role of number response categories on SAH ratings.

Offering a bigger scale range to record self-reported health goes beyond facilitating respondent's differentiation of health, to potentially achieve more standard sensitivity over the whole range of underlying health variable and decrease "ceiling and floor" effect where sensitivity gets lost at both

upper and lower end of the SAH scale between the choice ratings (Sturgis et al., 2001). Perhaps, SAH ratings can also prevent loss of information and convenient choice of mid-point by using the item root method (Oldendick, 2008), where questions are first streamed based on how one views own health, leaning towards good or bad side – and if good then it roots to “very good, somewhat good, or more towards good”, vice versa for choice of bad.

#### **Type 5: Mid points and bi-polar scales**

Rating scales are used as a prop to allow respondents to express their inclination and valence towards an appointed issue. Mid-points play a role not just dependent on the number of categories in the scale but also in bi-polar scales. Bi-polar scales can be described as scales with a mid-point that creates two symmetrical but opposing concepts on both sides. Numerically, an example of it can be a scale of “-2, -1, 0, 1, 2”. The role of mid points in a linear versus a bi-polar scale is distinctively different. Mid-points in bi-polar scales typically represents conceptual turning points such as from boring to interesting and may represent “indifference or ambivalence” (Schaeffer and Presser, 2003). This can serve to reduce the cost of thinking and help set a natural point of reference in a rating scale, otherwise non-existent in a linear scale (Groot,2000; Schaeffer and Presser, 2003). In contrast, mid-points on a linear scale is subsumed within its range. For example, mid-point on a bi-polar scale creates a point of reflection from” boring to interesting”, whereas, on a linear scale, the choice changes to playing a more undefined role within a range “not interesting at all, to very interesting” (Schaeffer and Presser, 2003). Thus, in linear scales, one may postulate mid-points taking on a more neither/nor role, versus in bi polar scales where its serves to entail an either/or representation.

This characteristic of bi –polar scale potentially imparts a sense of “pseudo certainty” to the decision making via rule of opposites: either/or, is/is not, and simple logic of degree of each side of opposites. A linear scale of 1 to 10, on the other hand, has no natural point of reference and is more ambiguous or uncertain in nature. Certainty effect may manifest from this framing. More research should be conducted in this area to understand the underlying motivation of this difference.

Linear scales without mid points have also been shown to overcome social desirability bias of trying to please the interviewer in product labelling research on packaged food (Garland, 1991). However, there have been mixed findings on whether the lack of mid-points instigate respondents to be more positive or negative in own rating across both bi-polar and linear scales (Garland, 1991; Worchester and Burns, 1975; O’ Muirheartaigh et al., 1995).

The potential role of such mid-points on a bipolar scale may be interesting to consider for SAH scales. Current SAH scale does not have a natural point of reference to aid response. Respondents express their health perception in a fluid manner based on own sense of health and understanding of the question. Inflated middle responses (Greene et al., 2014) may be one such outcome of self-created point of reference. Implementing a mid –point in a bipolar scale creates a different SAH rating experience, whereby the respondent will be influenced to think in duality and deviations from the mid-point. Although the actual SAH question may remain as “on this scale of -2 to 2 where -2 is very bad and 2 is very good, how would you rate your health in general?” – The existence of a mid-point 0 on this bi-polar scale indirectly reconstructs the SAH question as “how good or bad do you think your SAH is” in respondents mind, and that is used as the frame for evaluating own SAH. Little research has currently been conducted in this area to explore the opportunity for use of bi-polar scales in SAH ratings, its impact and advantages, and can be considered for future work.

### 4.3.3 Drivers and Source of biases to SAH ratings: (Social desirability)

Podsakoff et al., 2003 quotes Crowne and Marlowe (1964) on social desirability as “the need for social approval and acceptance and the belief that it can be attained by means of culturally acceptable and appropriate behaviours”. Social desirability effects in surveys may encourage subjects to rephrase their answers to meet perceived interviewers’ needs. Such inclination to present self in a positive light may lead to additive and systematic error (Tourangeau and Smith, 1996; Schwarz, 1999; Viswanathan, 2005).

Work on social desirability bias stemmed from Cronbach’s (1950) exposition of response set errors in surveys. The tendency to agree with a statement where a person is more likely to answer “true, yes, and agree” regardless of the content was coined as “acquiescence” by Cronbach (1950) and “evasiveness” where they are more like to identify with “uncertain, indifferent”. Social desirability bias were postulated to manifest as “acquiescence” and “evasiveness” in surveys (Bradburn, 1983, pg. 292; Green et al., 2014).

Social desirability effects can be triggered by various scenarios like: perceived lack of confidentiality and anonymity (Tourangeau, 1991; Smith et al., 1994; Aquilino, 1994; Paulhus, 1984), need for positive self portrayal due to social interaction between respondent and interviewers (Taylor and Brown, 1988), perceived socially undesirable or embarrassing behaviours in sensitive questions (Tourangeau and Smith, 1996), mode of interview (Locander et al., 1976; Richardman et al., 1999; Bradburn, 1983; O’Muircheartaigh et al., 2000; Aquilino, 1994; Cronbach, 1950; Brewer et al., 2004; Conti and Pudney, 2011; Crossley and Kennedy, 2002; Conti and Pudney, 2011), interviewer effect (Viswanathan, 2005; Taylor and Brown, 1988; Bradburn, 1983), affiliation and relationship building goals (Cialdini and Goldstein, 2004; Fehr and Schmidt, 2006; Engelmann and Strobel, 2004), and Pygmalion effect (Rosenthal, 1987). The below exposes each of these dimensions.

#### Confidentiality and anonymity

Social desirability has also been ascribed by Paulhus (1984) as one of “unintentional self-deception or intentional impression management”. When subjugated to disclosure of illegal, sensitive or socially undesirable behaviour, the subjects’ willingness to be open depends on reassurance of self-protection afforded by the degree of confidentiality of information and anonymity of responses (Aquilino, 1994). In such instance, impression management was found to be a more pertinent focus than self-deception in addressing social desirability bias in self-reports (Paulhus, 1984).

#### Positive self portrayal

As part of impression management, respondents are likely to engage in face saving, courtesy bias and power dynamics in enacting social desirability. The social interaction between the respondent and interviewer triggers adoption of norms and strategies to enhance positive self-concept for positive self-enhancement, making one look and feel better about oneself (Taylor and Brown, 1988; Swann, 1984; Bradburn, 1983; Richardman et al., 1999).

#### Affiliation goals

On another note, social desirability bias can also be motivated by an overly helpful behaviour of the need to please. Affiliation and reciprocity goals are used to gather approval from others to build intimacy in relationships (Cialdini and Goldstein, 2004). The interviewer effect and face to face mode of interview are also subjected to behavioral bias driven by the “Pygmalion effect” (Rosenthal, 1987) where people perform better when they are expected to do so.

### Sensitive questions

Having to please and cover is more prominent in sensitive questions context. Sensitive questions are susceptible to interview mode and social desirability bias. Tourangeau and Smith (1996) describes sensitive questions as “being of a private nature, or one that raises concern if reported truthfully”. However, this aspect may be less relevant to SAH questions, as it can be generalized to be a sensitive question as such.

### Interviewer effect

The presence of an interviewer in the survey can lead to interviewer effect. As the personal is political, the interviewer carries a personal identity and respondents may not perceived the interviewer just as in instrument who delivers questions. Instead, the interviewer may be viewed as being representative of particular socio-economic, demographic groups and encompassing particular attitudes (Bradburn, 1983, pg. 293; Viswanathan, 2005; Taylor and Brown, 1988; Bradburn, 1983). This sets expectations and respondents tap into what is expected of them rather than what is in them; health may become a cultural construct and response instead of an objective personal health review. Overall, driving systematic error in amongst some sub-groups (Bradburn, 1983, pg. 293; Viswanathan, 2005; Taylor and Brown, 1988; Bradburn, 1983).

### Mode of interviews

Different modes of interview like telephone, face to face with interviewer, self- completion via mail vs, interviewer, and online computer interviews also influence social desirability bias (Locander et al., 1976; Richardman et al., 1999; Bradburn, 1983; O’Muirheartaigh et al., 2000; Aquilino, 1994; Cronbach, 1950; Brewer et al., 2004; Conti and Pudney, 2011; Crossley and Kennedy, 2002; Conti and Pudney, 2011). Answering sensitive questions through self-completion questionnaire reduces social desirability bias compared to interviewer (telephone or face to face) administered questions and closes the gender gap difference in reporting of number of sexual partners (Tourangeau and Smith, 1996). Aquilino (1994) showed “greater social distance” made available through use of self- completion questionnaire to provide confidentiality and elicit highest response of cocaine and marijuana drug use, followed by face to face interview and worse being telephone interviews. Richardman et al., (1999) conducted an extensive literature review, meta-analysis of these studies and an experiment on computer aided versus pen and paper self-completion. Their study concluded social desirability can be attributed to the physical or verbal presence of an interviewer. Presence of interviewer in survey also raises respondent’s consciousness of their external survey environment. In a job satisfaction study, Conti and Pudney (2011) found changes in distribution of job satisfaction when the question was asked by face to face interviewer versus via a self-completion questionnaire. The social desirability bias driven by response mode interviewer effect was especially pronounced in the presence of kids and partners. “Not in front of the kids” and “need to put on a good show” for the interviewer led to upward bias of job satisfaction and “need to maintain economic bargaining power, thus not showing too much satisfaction in front of partner” leads to downward bias in response of job satisfaction.

### Evidence of social desirability bias in health reporting

Social desirability has been found to affect self-reported health measures such as over reporting of self physical activity versus objective health measures. Adams et al., (2005) regressed social desirability (personality traits) on difference of objective versus self-reported physical activity measurements and

found significant relationship between over reporting of own physical activity with social desirability (Adams et al., 2005; Greene et al., 2014).

Mode of interviews is another common method used to test for social desirability bias. However, there may be other confounding factors which cannot be disentangled, and at best there are associations between mode of interviews and social desirability bias. Crossley and Kennedy (2002) puts forth the possibility for social desirability bias due to interviewer mode -when the treatment group was asked their SAH a second time via face to face interview, versus first time SAH response via written questionnaire, the distribution of middle category (food) was reported less, signaling potential social desirability bias due to interviewer mode. However, the authors acknowledge potential of measurement error and learning effects

Beyond assessing effects on an aggregate level, Brewer et al., (2004) examines the presence of social desirability bias in different interview modes by capturing person-level variable. Social desirability scores were elicited from a set of questions that links social desirability to a personality trait of low self-reports of socially undesirable behaviour. Although the authors found evidence of people with high social desirability scores reporting less health syndromes, there was no significant effect of mode of interview through examination of interaction effects (mail survey versus telephone interview) on social desirability.

The question to ponder is the relevance of social desirability bias for SAH rating. Beyond the potential of respondents being helpful and building affinity with interviewer leading to social desirability bias, what other characteristics of SAH rating would warrant respondents to be subjected to social desirability bias? SAH rating, unlike other health measures on dieting, physical exercise, body perception, drug usage, is of a more general, less sensitive and private nature. Perhaps, the social desirability bias may be of lesser source of behavioral bias in SAH rating, compared to other sources of behavioral bias like over confidence, optimism, self-concept and framing effects.

#### **4.3.4 Drivers and Source of biases to SAH ratings: (Satisficing)**

Satisficing in survey can be described as an act where the individual is unable (due to lack of cognitive power) and/or unwilling (due to lack of motivation) to put in conscious effort to go through the four stages of cognitive aspects of survey (Krosnick, 1991; Anand, 2008) proposed by Tourangeau (1984) comprising: comprehension, memory recall, integration and judgment, and recording of response.

This literature review has found only one study on satisficing in health symptom reporting. Brewer et al., (2004) studied the presence of satisficing across two different modes of interview of phone surveys versus self completed mail surveys. Their study found evidence of satisficing as there were less symptom reports in phone versus mail surveys. Moreover, the nature of symptoms not reported on the phone were milder symptoms, presumed to need more effort to recall.

Beyond this, satisficing in SAH has not been formally explored but offered as explanations to the potential cognitive difficulties respondents may face in going through the four stage cognitive process of survey (Sturgis et al., 2001). Greene et al., (2014) presents a time-cost argument for decreasing respondents' motivations to pay attention to the question and satisfice by quickly answering and picking a middle response option.

As there have not been many formal studies to date on satisficing in SAH ratings, this section will explore the possibilities for the relevance and act of satisficing in SAH ratings. A key question to ask is - is

satisficing is a relevant source of bias for SAH ratings? SAH questions are perceived to be easy to answer, and indeed correctly so, as the individual live with their health daily and are aware of their health behaviour and experiences. Moreover, the SAH question seems sufficiently simple and manageable of just putting a rating on a scale of 1 to 5.

However, whilst it is not difficult to answer a SAH rating, it may still demand some effort from the respondent. The previous sections have shown some of the efforts which a respondent has to make as part of the SAH rating experience. In what is observed versus perceived health information, individuals may not have all the relevant health information to make an accurate judgement about own health and they may have to choose what type of objective versus subjective health information to use for rating SAH. This is followed by a need to integrate actual health experience with health expectations where the respondent has to engage in self-concept of health, perceived control over health and reconcile it with actual health experience. Finally, the respondent may have to deal with the framing effects of SAH question to decode how best to represent one's resolved health state. They may also be further influenced by external survey environment like interviewer effects and moods.

Accurate rating of SAH is mainly dependent on the respondent's ability and willingness; Able: Have the information and cognitive ability to work through one's health rating; Willing: Even if they had the cognitive powers, are they willing to invest the time and effort to decode the answer for themselves and then to report it honestly. Ideally, all the four stages of cognitive aspects of survey are properly thought through by the respondent to give an optimized answer (Anand, 2008; Krosnick, 1991). In reality, one may satisfice in answering SAH if no full conscious effort is made in engaging each stage, or if certain stages of recall of memory, integration and judgment of information are skipped for a quick answer.

Krosnick (1991) highlights the three conditions which nurtures satisficing. The potential inclination to satisfice in SAH is dependent on; 1. *The respondent's ability to complete the task* (the cognitive sophistication of the individual, their experience with the health topic and ability to consolidate and reconcile any conflicting attitude about one's health); 2. *The respondent's motivation to complete the task* (importance of topic to the respondent, sense of accountability for wrong or no response, actual enjoyment in engagement of thinking); 3. *The level of task difficulty* (impact of length of questionnaire and framing effects) (Krosnick, 1991; Anand, 2008).

Here, a tentative synthesis of challenges in answering SAH are sketched out using the four cognitive stages of survey. It provides rationale for further research in verifying role of satisficing in SAH ratings.

- **Comprehension of questions:** Respondents are required to have an understanding of health concept, decide which definition of health concept they are using and nature of health status (Lee, 2014). There may also be difficulty in answering a "general" SAH question (Sturgis et al. 2001). Framing effects may also affect understanding of response scale (Lee, 2014). Furthermore, mode of interview like telephone versus self-administered versus face to face can influence the amount of time pressure given to respondent for cognitive thinking (Schechter and Hermann, 1997; Tourangeau, 1999)
- **Recall of memory:** Respondents face various challenges to recall health events needed to rate SAH. In health symptoms recall: One tend to recall more prominent and not day to day health events (Schwarz, 1999; Schechter, 1997; Fienberg et al., 1985; Krosnick, 1991). Schechter (1993) suggest that retrieval of information on things like pains, aches are based on subjective weights respondents put on the factor. In healthcare consumption recall: People may be forgetful and

recall wrong time frame through telescopic memory (Bradburn, 1983). Inability to recall may lead to satisficing by using belief of maintenance of status quo in health from the past (Krosnick, 1991) and use of desired memory like family longevity and own healthy behavior for educated guess (Van Doorn, 1999; Schechter and Hermann, 1997).

- **Integration and judgment:** Fienberg et al., (1985) notes the difficulty of this stage for the respondent to make an overall judgment on SAH. The respondents have to integrate the different information of actual ill health experience, health lifestyle, own self concept of health, comparison of own health at other points in times versus comparing with others. All these factors may pressure the respondent to tap into **availability and representative heuristic** to satisfice and give the first available satisfactory answer. Additionally, Podsakoff et al., (2003) suggest respondents need to maintain consistency within self and the urge for portrayal of consistency may not truly reflect how they feel about an issue. **External moods and affective emotions** may also affect judgment and integration of information (Hanita, 2000; Podsakoff et al., 2003; Schechter, 1993). Positive moods are found to lead to positive evaluation of almost all aspects of life versus negative mood drives more critical assessment but its effect is unpredictable (Hanita, 2000; Podsakoff et al., 2003). Being in a good mood also lead to heavier reliance on heuristics and being less thoughtful or critical about one's response (Park and Banaji, 2000; Caputo, 2012). It is important to control where possible for the mood of a respondent at time of response and how it may affect their health assessment and SAH ratings. Finally, If the cognitive requirements to integrate and judge health information is too much and depletes one's self control and will power, respondents may lapsed into **automacity** and chooses to process information in a "mindless manner without careful consideration" (Hanita, 2000; Bargh; 1994; Baumeister et al., 1998; Cialdini and Goldstein, 2004; Chartand, 2005; DeWall et al., 2008; Evans et al., 2011; Janssen et al., 2008; Mullainathan and Thaler, 2000; Robertson and Barling, 2004)

The above highlights the cognitive challenges a respondent may face in self-reported health. Should they succumb to these challenges, compromise will be made fostering the below potential outcomes of satisficed SAH ratings. These satisficing outcomes have been defined by Krosnick (1991) in his seminal piece on satisficing.

- **Outcome 1: Weak satisficing** (Krosnick, 1991): The first reasonable alternative is picked by the respondent as an acceptable quick response. Note that weak satisficing of acquiescence is less relevant here due to the nature of SAH scale not being about agreeing/disagreeing but requesting a point of view on one's health.
- **Outcome 2: Strong satisficing:** (Krosnick, 1991): There are four possibilities of strong satisficing in SAH ratings. 1. Respondents gravitate towards "do not know" options, perhaps the parallel in SAH would be an inflated middle responses (Greene et al., 2014). 2. Respondents may also choose to maintain **status quo** of health. 3. Respondents randomly chose options if there are no feedback or repercussions on them. 4. Lack of differentiation in outcome between questions with the same rating scales. (E.g. If SAH was asked at the end of the questionnaire after a series of health related question with similar scale of 1 to 5). Use of similar rating scales across questions although efficient for the researcher, may encourage satisficing where respondents give the same response across all questions with same rating scale (Krosnick, 1999).



## Section 5: Proposed experiment to testing framing effects

### 5.1. Introduction:

#### Motivation:

Bago d’Uva et al. (2008) established differential health reporting by education level with higher educated under rating health compared to the lower educated, especially in the Belgium, Germany and the Netherlands in period 2004 – 2005.

The authors analysed data from WHO-MCS of respondents’ classification of their own health in 6 domains of mobility, cognition, pain, sleep, breathing and emotional health –an example of this is response to mobility question: “Overall, in the last 30 days, how much of a problem did you have moving around” on a scale of “none, mild, moderate, severe, extreme”. Additionally, they also had access to vignette ratings in a particular health domain representing different levels of difficulty. This approach was used to establish heterogeneity in reporting via identification of shifting cut points from mapping vignette description rating with self-reported health.

#### The research question and hypotheses:

With this context, this thought experiment aims to explore if the higher educated are influenced by the concept of wording and labelling framing effects in the self-reported health scale leading to under rating of health compared to the lower educated. This thought experiment test the below two hypothesis:

- H1: There is wording framing effects in SAH amongst the high educated compared to the lower educated.
- H2: Wording framing effect may influence the high educated to under-rate health compared to low educated , reflecting pessimism of the higher educated.

The words and labels of a scale promotes anchoring heuristics, leading to ease of retrieval of particular information set by the anchor of words and labels (Kahneman, 2003). Such effects may lead to non-random systematic error (Viswanathan, 2005). Studying this potential wording and labelling framing effects will enable better understanding if it contributes as one the drivers to the heterogeneity in health reporting found in Bago d’Uva et al. (2008). Previous research by Bowling and Windsor (2008) have shown evidence that US SAH scale starting with “excellent” promotes optimism bias. In this context, we would like to explore if such potential wording framing effects may have affected the low and high educated differently, capturing attitude of pessimism or realism in the higher educated. This experiment exposes the higher educated to a positive US SAH scale as well to see if they still under rate health on such a scale. If the higher educated also succumb to under rating of the more positive US SAH scale, we may be able to explain Bago d’Uva et al. (2008) finding of reporting heterogeneity between high and low educated to be in part due to pessimism or realism of the higher educated. Overall, this also raises the need to take into consideration the choice of scales used for SAH assessment in interpretation and comparisons of findings.

### **The approach to answer the research question:**

This thought experiment will highlight a possible way to investigate if one of the reasons for the heterogeneity found in Bago d’Uva et al. (2008) study is in part driven by the higher educated under rating health due to wording framing effects of the scale used to elicit six health domains SRH. We will illustrate how this issue may be examined via comparing responses to vignettes using a WHO labelled SAH scale of “very good, good, fair, bad, very bad” versus US labelled SAH scale of “excellent, very good, good, fair, bad”.

Whilst Bago d’Uva et al. (2008) study adopted the World Health Survey (WHS) scale of “none, mild, moderate, severe, extreme” for self-rated health in a particular health domain, which is different from the scale of a non - domain driven SAH, the effects of its word and labelling range may be seen to activate a similar response as a WHO SAH scale of “very good, good, fair, bad, very bad”. Similarly, if we were to construct a more positively skewed scale of SRH of a health domain e.g. “none, very occasional limitations, mild, moderate, severe”, this scale will likewise better correspond to a US SAH scale of “excellent, very good, good, fair, bad”.

As such, we postulate the higher educated may be exposed to wording and labelling framing effects of the WHS SRH health domain scale, i.e. if the response scale was labelled as “none, very occasional limitations, mild, moderate, severe”, there may have been lesser heterogeneity between higher and lower educated respondents. Using the parallels of WHS health domain SRH scale with WHO SAH scale, and a hypothetical more positively skewed health domain SRH scale with US WHO scale, we will seek to capture this by examining high versus low educated individuals responses to vignette using WHO labelled SAH (“very bad, bad, fair, good, very good”) vs US labelled SAH (“poor, fair, good, very good, excellent”) scale. We choose to explore this via using SAH scales instead of the WHS health domain SRH scale labelling, to support and further expound on this thesis role to flash out of sources of biases in SAH.

### **Contribution of this experiment:**

Previous attempts to study this wording and labelling framing effects were often conducted with dual objective of examination of question order effects as well as framing effects. Both Lumsdaine and Exterkate (2013) and Bowling and Windsor (2008) exposed respondents to SAH ratings twice and randomly assigned either WHO SAH or US SAH scale as the first question. These studies provided empirical evidence of presence of wording and labelling framing effects, and these two scales are not interchangeable (Lumsdaine and Exterkate, 2013), with the US SAH scale showing more optimism bias but a lower ceiling effect (Bowling and Windsor, 2008). Validity and effectiveness of cross country comparisons of studies as well as detection of improvement of health status in follow studies will be affected by choice of scale (Bowling and Windsor, 2008; Lumsdaine and Exterkate, 2013).

This thought experiment contributes by using a different measurement tool of vignettes to assess wording and labelling framing effects. Vignettes have been used for identifying reporting heterogeneity and correction of reporting bias. It is interesting to assess if this vignette method can also be an alternative to current methods of using own SAH to identify wording and labelling framing effects. Furthermore, even if, previous studies showed the SAH twice by randomly assigning respondents to the US versus WHO version of SAH scale as the first, it may still be difficult to fully disentangle the question

order effect from wording labelling framing effects in those approaches as the respondents would still have been exposed and influenced by the initial anchor from the first SAH question shown to them. Here, we clearly isolate this issue to explore only the wording and labelling effects of a scale. This is possible as unlike previous studies, we will not show the respondents the SAH scale twice, but by using only one measure of rating of vignette. Respondents will then be randomly assigned to rate this vignette using with the WHO version or US version of SAH scale. In this exercise, we assume both assumptions of vignette consistency: “where the individual uses the same response scales to rate the vignettes and their own situations” and vignette equivalence: “where the description is perceived by all to correspond to the same state” (Bago d’Uva et al., 2011) hold.

## 5.2. The experiment design

### Describing the sample

This survey will be using the LISS Panel. The LISS Panel consists of 4500 households, comprising 7000 individuals. The LISS panel uses a true probability sample of households from the population register by Statistics Netherlands. The experiment will be fielded as part of the LISS Panel Core online survey in August 2016 – September 2016. Respondents are incentivized for completion of the overall questionnaire as reward for work effort.

For purpose of analysis, we will only be using the responses of healthy individuals: who have coded self assessed health as 4 or 5 in the LISS Panel Core Survey 2014; non-smokers; exercises; aged 25 – 40 years old. The sample pool will ensure mix of low and high educated with low educated set at below 7 years of education. The restrictions of healthy respondents and use of a panel has two implications: 1. There may not be an equal number of low versus high educated respondents used for the final analysis; 2. Whilst power test will be conducted to check if we manage to generate the minimum sample, this is not assured as we are relying on a fix set of panel with restrictions. We will also provide a summary statistics comprising maximum, minimum, mean and standard deviation of the sample for key variables like age, gender, health status, household income, and ethnicity.

### Describing the task

Demographics of respondents will be available from the panel, and not recorded in this experiment. This experiment involves only two questions: first involving an own self assessed health question, and second a rating of one vignette using either the WHO SAH scale, or the US SAH scale

**Question 1:** Both high and low educated respondents will first be asked to rate their own SAH on a scale of 1 to 5 using the WHO SAH scale (assuming LISS panel use WHO version of SAH). “Thinking about your own health, how would you rate your health on a scale of “5 -very good”, “4- good”, “3- fair”, “2- bad”, “1-very bad”

**Question 2:** A vignette representing a healthy individual adapted from WHO MCS survey, for example “Mary has no problems in moving around using her hands, arms and legs. She jogs 4 kilometres twice a week without any problems.” will be presented to the respondents. The same vignette will be used for all high and low educated respondents to reduce any possible confounders due to vignette descriptions.

This vignette will be explained to all individuals - that they are to look at this description and to rate what they think the health of this person is. Respondents will then be randomly assigned into 2 groups of those exposed to WHO SAH scale for rating this vignette, versus those using the US SAH scale to rate this vignette. I.e. Within high educated respondents, which scale the individuals will use for rating the vignettes will be chosen by random assignment. This is repeated for the low educated respondents. Respondents will be asked: "Overall, looking at this person in the description, what you think is the health of this person?" The response categories will be: for US SAH scale - "5-excellent", "4-very good", "3-good", "2-fair", "1-poor"; for WHO SAH scale - "5-very good", "4-good", "3-fair", "2-bad", "1-very bad".

### **The experimental design rationale:**

The key challenge in this study is to create a context for exploring if framing effects exist for higher educated compared to the lower educated, and if this framing effects result in under rating of health amongst the high educated reflecting pessimism. For this, we first need to establish both the premise for exploration and a control benchmark.

- This thought experiment adopt the following premise: High and low educated are both not affected by framing effects in WHO SAH scale, and they both rate the WHO SAH scale similarly.
  - This premise is derived from deliberate selection of low and high educated respondents who both rate their own self assessed health on a WHO SAH scale as 4 or 5 in Q1 of this experiment, at the analysis stage.
    - It is important to note that this premise is contrary to what has been found empirically in studies by Bago d'Uva et al. (2008) where high and low educated are found to use different "scales" with shifting cut points when rating WHO-SAH.
  - However, this is suggested purely for purpose of analysis of potential wording framing effects in this thought experiment.
  - Thus, any evidence of framing effects on higher educated will be based on the difference in US SAH scale rating between the low and high educated.
- Additionally, there is a further necessary restriction for this to work. We presuppose that the low educated are not affected by wording and labelling framing effects. The experiment will need to validate if this holds.
  - In accordance to empirical evidence by Bago d'Uva et al. (2008) of higher educated under-rating WHO SAH scale compared to lower educated, this allows us to postulate that the lower educated may not be affected by wording framing effects on different US or WHO SAH scale.
  - As such, if there are differences in US SAH ratings between low and high educated respondents, this difference implies the higher educated being exposed to wording framing effects.
- With the above, the control benchmark is set by defining a scenario and generating a set of respondents where there is no framing effects between low and high educated in one particular frame (taken as WHO SAH scale here).

- This control is achieved by selecting only low and high educated respondents who both rate 4 or 5 on their own self assessed health in Q1 of this experiment using a WHO SAH scale.
- We presuppose here that in Q2 of the experiment, both the low and high educated will rate this vignette similarly on a WHO SAH scale, especially since the vignette has been selected to reflect a healthy individual, and both the high and low educated respondents selected for the analysis are those who rate their own self assessed health in Q1 WHO-SAH a 4 or 5. The experiment will need to validate if this holds. This assumes vignette equivalence and vignette consistency assumptions.
- This control will then allow for any subsequent wording and labelling effects in high versus educated respondents to be picked up by the alternative treatment US SAH scale frame
- Furthermore, if wording framing effects exist amongst the higher educated, we want to then assess if it causes them to under-rate health reflecting a pessimist mindset towards own health.
  - We investigate this by comparing the proportion of low versus high educated people rating health as optimal after rescaling and combining both WHO and US SAH scale response of “excellent/very good”. If the proportion of high educated people rating optimal health level on these two combine scales are significantly lower than low educated people, and the dual of bigger proportion of higher educated who rate health at sub optimal health level of “poor/very bad/bad” of the rescaled and combined WHO and US SAH responses are significantly bigger than the lower educated, we may conclude that high educated under-rate.

To ensure all variations in vignette scale ratings is attributed insofar only to wording and labelling framing effects, the key in enabling this, in addition to randomization of WHO-SAH versus US-SAH scale allocated to respondents in rating of the vignette, lies in three aspects of the experimental design:

1. First: Removal of possible confounders driven by individual heterogeneity (e.g. use of self-concept of health, optimism, subjectivity, even satisficing) which may influence rating on a scale, through the use of vignettes instead of individual SAH health for evaluation of framing effects
  - Use of vignette facilitates a common ground of evaluation amongst all respondents compared to use of an individual’s perspective on SAH. Its indirect projective approach helps to minimize these other possible confounders.
2. Second: Removal of any possible confounders that has to do with actual health state change which may influence rating on a scale, through selection of only healthy high and low educated respondents. As mentioned above, selection of only healthy respondents is paramount to this experiment, even if we were to conduct randomization of WHO SAH vs. US SAH scales for vignette ratings.
  - The importance and role of this healthy stock of respondents lies in the analysis stage. It helps us to achieve the condition that all respondents selected for analysis rate 4 or 5 on own WHO SAH in Q1, and subsequently postulate for both low and high educated to rate vignette WHO-SAH scale similarly. These are two conditions are necessary for the

experiment to establish if wording framing effects exist amongst higher educated in the US-SAH scale.

- To ensure healthy stock of respondents, only individuals who rated own WHO SAH rating as 4 or 5 in LISS Core Survey 2014 will be put in the experiment. We will also only use respondents from year 2014, to minimize potential health shocks or effects of any changes in healthcare system reform in the Netherlands between now and then. Additionally to increase probability of still being healthy, all respondents will be within narrow age band of 25 – 40 years old, engaged in sports (LISS lifestyle survey) and non-smokers (LISS health tobacco, alcohol survey)
  - Respondents own WHO SAH rating in Q1, will be compared to their 2014 ratings, and we will eradicate individuals who have a different WHO SAH rating in 2014 with 2016 ratings. This is to ensure subsequent health ratings of vignettes do not pick up factors due to actual change in health state, but only due to wording framing effects
3. Third: Getting healthy low and high educated individuals to rate on a healthy vignette, to extract variation only from wording framing effects
- To capture wording framing effects of the scale, the experiment tries to minimize any potential subjective bias by using a sample of vignette quite close to the actual respondent's health state. The vignette selected should be reflective of someone quite healthy. Following Bago d'Uva et al. (2008), healthy vignettes from World Health Survey (WHS); WHO-MCS will be used. Vignettes from other surveys such as SHARE or ELSA do not include healthy vignettes with the lowest level often starting with some mild form of health discomfort.
  - Additionally, at the analysis stage, as the vignette selected for this experiment reflects that of a healthy individual, we will remove all outlier responses in vignette ratings, to ensure it does not taint the wording framing effects we are looking for (e.g. rate a healthy vignette as unhealthy).

### 5.3 Analysis carried out on vignette ratings using different US SAH versus WHO SAH scales:

The below comparisons will be made to examine the hypothesis. These are illustrated in table 3a, table 3b and table 4 .

- H1: There is wording framing effect in SAH amongst the high educated compared to the lower educated
  - The difference in higher versus lower educated rating of vignettes using WHO SAH (table 3a)
  - The difference in higher versus lower educated rating of vignettes using US SAH (table 3a)
  - The difference in WHO SAH versus US SAH scale vignette ratings for high educated (table 3b)
  - The difference in WHO SAH versus US SAH scale vignette ratings for low educated (table 3b)
- H2: Wording framing effect may influence the high educated to under-rate health compared to low educated, reflecting pessimism of the higher educated.
  - The difference in proportion of lower versus higher educated people in optimal health (“excellent/good”) combined WHO SAH and US SAH health rating of vignettes (table 4)
  - The difference in proportion of lower versus higher educated people in sub optimal health (“poor/very/bad”) combined WHO SAH and US SAH health rating of vignettes (table 4)

#### ***The overall process for analysis of Hypothesis 1 will involve the below 4 steps:***

1. Establish if wording framing effects do not affect low educated. Establish if low educated rate the vignette on a US-SAH scale vs. rating vignette on a WHO-SAH scale differently.
  - This is achieved by examining if the proportion of low educated who rate “excellent/very good” in recoded US SAH scale for vignettes is significantly different from the proportion of low educated who rate “excellent/very good” in recoded WHO SAH scale for vignettes for low educated (***Table 3b: comparing 1’ and 3’***).
2. Establish if both the low and high educated rate the vignette on a WHO-SAH scale similarly
  - This is achieved by examining if the percentages of responses on “very good and good” in the WHO SAH scale for vignette rating amongst the low and high educated are significantly different (***Table 3a: comparing 3 and 4***).
  - If both low and high educated rate the vignette on a WHO SAH scale similarly, complemented with our curation of selection of only healthy respondents for this analysis (rate own self assessed health in Q1 on WHO SAH scale as 4 or 5, also rate own self assessed health in LISS Core survey 2014 as 4 or 5, and healthy behaviour or non-smokers and exercising), any potential wording framing effects will have to be captured by the vignette rating using US SAH scale

3. Establish if low and high educated rate the vignette on a US-SAH scale similarly
  - This is achieved by examining if the percentages of responses on “excellent and very good” in the US-SAH scale for vignette rating amongst the low and high educated are significantly different (**Table 3a: comparing 1 and 2**).
  - If both low and high educated rate the vignette on a US-SAH scale differently, then it implies framing effects exist for higher educated compared to the lower educated. This is so as it was shown in step 1 that the lower educated are not affected by wording framing effects and since both low and high educated rate vignette on WHO-SAH similarly as in step 2, then any differences in vignette rating using a US-SAH scale can be attributed to wording framing effects in the high educated.
4. Establish if wording framing effects affect high educated. Establish if high educated rate the vignette on a US-SAH scale vs. rating vignette on a WHO-SAH scale differently
  - This is achieved by examining if the proportion of high educated who rate “excellent/very good” in recoded US SAH scale for vignettes is significantly different from the proportion of high educated who rate “excellent/very good” in recoded WHO SAH scale for vignettes for high educated (**Table 3b: comparing 2’ and 4’**).
  - If high educated is found to significantly rate vignette using US-SAH scale differently from if rating vignette using WHO-SAH scale, together with above steps 1 to 3, this provides evidence in support of the existence of wording framing effects in the higher educated compared to the low educated.

Whilst it may seem only necessary to conduct analysis on evidence of wording framing effects by comparing either low educated rating of vignette on US-SAH versus WHO-SAH scale and similarly for the high educated (as in table 3b), or, by comparing low versus high educated vignette rating on a WHO-SAH scale and similarly on the US-SAH scale (as in table 3a), there is value to administering both sets of analysis. Primarily, these are interconnected as illustrated in above section on overall process of analysis. If priority has to be given, analysis of table 3b will be more pertinent: comparing low educated rating of vignette on US-SAH versus WHO-SAH scale is akin to a necessary condition for this hypothesis analysis to work; comparing high educated rating of vignette on US-SAH versus WHO-SAH scale is comparable to a sufficient condition. However, both sets of analysis in table 3a versus table 3b play different roles. The value of understanding low versus high educated vignette rating on a WHO-SAH scale and similarly on the US-SAH scale (as in table 3a) is its role in illuminating what drives the framing effects in higher educated and further evidence of framing effects in higher educated, overall providing more support for the hypothesis and a stronger story.



**Testing Hypothesis 1: There is wording and labelling framing effect in SAH amongst the high educated compared to the lower educated**

As above, the below analysis in table 3a and 3b will be carried out. As all tests are between independent samples, and the sample may not be normally distributed, we will use non parametric test of Mann Whitney test at 5% level of significance will be carried out to assess if the two randomly assigned SAH scale samples come from the same population. It will test to see if the median is different between the two SAH scales.

*Table 3a: Analysis of H1:*

- *Establish if both the low and high educated rate the vignette on a WHO-SAH scale similarly*
- *Establish if both the low and high educated rate the vignette on a US-SAH scale similarly*

	Low educated	High educated	
US SAH scale	1	2	<p><b>Step 3: Compare 1 and 2</b></p> <p>If the null of low educated and high educated rate US SAH scale similarly <u>is rejected</u> – supports evidence of framing effects in high educated</p>
WHO SAH scale	3	4	<p><b>Step 2: Compare 3 and 4</b></p> <p>If null that low educated and high educated rate WHO SAH scale similarly <u>is not rejected</u> - provide support of potential framing effects if US SAH scale is used.</p> <p>If both low and high educated rate the vignette on a WHO SAH scale similarly, complemented with our curation of selection of only healthy respondents for this analysis (rate own self assessed health in Q1 on WHO SAH scale as 4 or 5, also rate own self assessed health in LISS Core survey 2014 as 4 or 5, and healthy behaviour or non-smokers and exercising), any potential wording framing effects will have to be captured by the vignette rating using US SAH scale</p>
	<p><b>** 1 is percentage of responses on “excellent and very good”</b> in the vignette rating using US SAH scale amongst the low educated</p> <p><b>** 2 is percentage of responses on “excellent and very good”</b> in the vignette rating using US SAH scale amongst the high educated</p> <p><b>** 3 is percentage of responses on “very good and good”</b> in the vignette rating using WHO SAH scale amongst the low educated</p> <p><b>** 4 is percentage of responses on “very good and good”</b> in the vignette rating using WHO SAH scale amongst the high educated</p>		

Table 3b: Analysis of H1:

- Establish if wording framing effects do not affect low educated. Establish if low educated rate the vignette on a US-SAH scale vs. rating vignette on a WHO-SAH scale differently.
- Establish if wording framing effects affect high educated. Establish if high educated rate the vignette on a US-SAH scale vs. rating vignette on a WHO-SAH scale differently

	Low educated	High educated
US SAH scale	1'	2'
WHO SAH scale	3'	4'
	<p><b>Step 1: Compare 1' and 3'</b></p> <p>If null that low educated rate US SAH and WHO SAH similarly is not rejected – supports evidence that wording and labelling framing effects does not affect low educated</p>	<p><b>Step 4: Compare 2' and 4'</b></p> <p>If null that high educated rate US SAH and WHO SAH scale similarly is rejected – supports evidence of high educated being exposed to wording and labelling effect in the US SAH scale</p>
	<p>** Following Bowling and Windsor (2008) we will compare the responses of both US and WHO scale by combining and recoding them. The recoded scales are “excellent/very good”, “good”, “fair”, “poor/bad/very bad”</p> <p>** 1' refers to proportion of low educated who code “excellent/very good” in combined recoded US SAH scale for vignettes for low educated</p> <p>** 3' refers to proportion of low educated who code “excellent/very good” in combined recoded WHO SAH scale for vignettes for low educated</p> <p>** 2' refers to proportion of high educated who code “excellent/very good” in combined recoded WHO SAH scale for vignettes for high educated</p> <p>** 4' refers to proportion of high educated who code “excellent/very good” in combined recoded WHO SAH scale for vignettes for high educated</p>	

Whilst Hypothesis 1 is about finding evidence of framing effects amongst high educated compared to the lower educated, Hypothesis 2 is about examining given the framing effect in Hypothesis 1, how does it affect rating of health. Hypothesis 2 is about wording framing effect influencing the higher educated to under-rate health compared to the lower educated, reflecting pessimism of the higher educated.

**The overall process of analysis of Hypothesis 2 will involve the below steps:**

- This is achieved by comparing the difference in proportion of lower versus higher educated people in optimal health (“excellent/good”) combined WHO SAH and US SAH health rating of vignettes (table 4). The rationale for this combination is as follows:
  - The above examination of framing effects in Hypothesis 1 is based on the premise of controlling and using one of the scales WHO-SAH scale as the benchmark to illustrate potential evidence of framing effects via the US-SAH scale
  - As such, we do not know in reality if the framing effects is in the US-SAH or WHO-SAH scale
  - Therefore, it does not matter which scales are used as long as we have established that such a wording framing effect does exist.
  - Consequently, to examine if there is an under-rating of health due to the framing effects, this analysis propose to combine the responses of vignette

rating of both the US-SAH and WHO-SAH scale for each of the low and high educated respectively.

- If the proportion of low educated who rate vignette on both US and WHO scale as excellent/very good is larger than the proportion of high educated who rate US and WHO scale as excellent/very good, this provides support that higher educated under rate health compared to lower educated, reflecting pessimism
- Similarly, comparisons are also made on difference in proportion of lower versus higher educated people in sub optimal health (“poor/very/bad”) combined WHO SAH and US SAH health rating of vignettes (table 4)
- If the proportion of low educated who rate US and WHO scale as fair/poor/bad/very bad is smaller than proportion of high educated who rate US and WHO scale as fair/poor/bad/very bad, this provides support that higher educated under rate health compared to lower educated, reflecting pessimism

**Testing Hypothesis 2: Wording framing effect may influence the high educated to under-rate health compared to low educated, reflecting pessimism of the higher educated**

The below analysis in table 4 will be carried out. Following Bowling and Windsor (2008) we will compare the responses of both US and WHO scale by recoding them. The recoded scales are “excellent/very good”, “good”, “fair”, “poor/bad/very bad”. As all test are between independent samples, and the sample may not be normally distributed, we will use non parametric test Mann Whitney test at 5% level of significance will be carried out to assess if the two randomly assigned SAH scale samples come from the same population. It will test to see if the median is different between the two SAH scales.

*Table 4: Analysis of H2: The difference in proportion of lower versus higher educated people in optimal health and poor health rating*

	Low educated		High educated		
	US SAH scale	WHO SAH scale	US SAH scale	WHO SAH scale	
Excellent/Very good	1	2	3	4	Compare (1+2) with (3+4) If reject the null that proportion of low educated who rate US and WHO scale as excellent/very good is same or smaller than the proportion of high educated who rate US and WHO scale as excellent/very good – provides support that higher educated under rate health
Good					
Fair	5	6	7	8	Compare (5+6+9+10) with (7+8+11+12) If reject the null that the proportion of low educated who rate US and WHO scale as fair/poor/bad/very bad is same or bigger than proportion of high educated who rate US and WHO scale as fair/poor/bad/very bad – provides support that higher educated under rate health
Poor/bad/very bad	9	10	11	12	
	** 1 refers to proportion of low educated who code “excellent/very good” in combined recoded US SAH scale for vignettes for low educated ** 2 refers to proportion of low educated who code “excellent/very good” in combined recoded WHO SAH scale for vignettes for low educated ** 3 refers to proportion of high educated who code “excellent/very good” in combined recoded US SAH scale for vignettes for high educated ** 4 refers to proportion of high educated who code “excellent/very good” in combined recoded WHO SAH scale for vignettes for high educated				

	<p>** 5 refers to proportion of low educated who code “fair” in combined recoded US SAH scale for vignettes for low educated</p> <p>** 6 refers to proportion of low educated who code “fair” in combined recoded WHO SAH scale for vignettes for low educated</p> <p>** 7 refers to proportion of high educated who code “fair” in combined recoded US SAH scale for vignettes for high educated</p> <p>** 8 refers to proportion of high educated who code “fair” in combined recoded WHO SAH scale for vignettes for high educated</p> <p>** 9 refers to proportion of low educated who code “poor/bad/very bad” in combined recoded US SAH scale for vignettes for low educated</p> <p>** 10 refers to proportion of low educated who code “poor/bad/very bad” in combined recoded WHO SAH scale for vignettes for low educated</p> <p>** 11 refers to proportion of high educated who code “poor/bad/very bad” in combined recoded US SAH scale for vignettes for high educated</p> <p>** 12 refers to proportion of high educated who code “poor/bad/very bad” in combined recoded WHO SAH scale for vignettes for high educated</p>
--	--

#### 5.4: Limitations of the thought experiment

This proposed experiment has two advantages of providing strict control and limitations to minimize confounders and attempting to isolate wording framing effects for examination. Due to the strict restrictions set, this experiment has strength of internal validity of findings.

However, there are many caveats and potential weaknesses to this bricolage approach of achieving randomization with experiments together with restrictions set for sample used for analysis. External validity is a key challenge for the following reasons: This is primarily an experimental set up which may not be representative of real world data; Use of LISS panel data only includes Dutch respondents; Control for healthy stock of respondents together with use of a healthy vignette may be seen to be too restrictive even if it facilitates isolation of wording framing effects. Furthermore, analysis of the hypotheses hinges on first finding no framing effects amongst the low educated, and second for both low and high educated to rate vignette on WHO- SAH similarly. Ensuring sample size through power test may also not be relevant in this instance as the sample from LISS panel is set. Additionally, the sample restrictions we set for analysis of hypotheses further deems to this to be even less achievable. Finally, the test of hypothesis for higher educated under rating is based on combining scales of wording labels. Similar test should be further conducted by combining scales based on numeric concordance.

## Section 6: Discussions

SAH is about the context of the individual, the question framing and survey. Section four on sources of bias to SAH highlights the need to be aware of the potential challenges and caveats of taking this measurement literally for policy making and comparisons of health achievement between countries and across time. Understanding of the context and the unseen influences of SAH facilitates more informed decision making.

Of all the potential sources of biases found in this literature review, social desirability bias probably has the least influence on SAH ratings. Social desirability bias is typically more relevant and associated with sensitive issues. One may be inclined to think there may not be sufficient sensitive issues in a SAH rating to warrant social desirability bias being a big issue in SAH ratings compared to other sources of behavioral bias like over confidence, optimism, self-concept and framing effects. As such, discussions on implications of sources of biases in SAH will be contained within subjectivity of health information, over confidence and optimism, self-concept of health, framing effects and satisficing.

### 6.1 Implications of subjectivity of health information as source of bias in SAH

The subjectivity of SAH measure includes: health expectations, understanding and rationalization of own perceived health states, to personal definition and thresholds of good health. Such subjective inference about one's health may lead to bias in SAH rating.

Systematic bias through reporting heterogeneity in SAH has been found for different SES groups by education and income level (Bago d'Uva et al, 2008; Etile and Milcent; 2006) and by age and gender (Van Doorslaer and Gerdtham, 2003). Bago d'Uva et al. (2008) showed that taking SAH at face value might underestimate SES inequality resulting in some statistically insignificant differences in health inequality. In their study, there appears to be no SES health inequality prior to correction for some health domains and countries. However, subsequent vignettes correction using EQ5D inspired six domains of mobility, cognition, pain, sleep, breathing, and emotional health based on World Health Survey and SHARE health domains reveals significant health inequality where older Europeans tend to under rate health, perhaps due to high health expectations and knowledge (Bago d'Uva et al., 2008). These findings highlight the importance of correcting for subjectivity in SAH measurements to ensure unbiased SES inequality assessment. Beyond individual country assessment of health distribution equities, such heterogeneity also needs to be addressed for better cross country comparisons of SAH by international organizations like WHO (Jylhä et al., 1998; Salomon et al., 2004).

To address the above challenge, below discusses possibilities with regards to the role of vignettes and its possible extensions beyond correcting for reporting heterogeneity via comparing sub group respondents rating of a similar vignette is compared with their own SAH ratings.

**1. Development of additional vignette sets to elicit definitions and thresholds of good health:** One aspect of subjectivity in SAH is due to different definitions and expectations of what constitutes good health. Whilst this is currently addressed via vignette correction of current EQ5D six dimensions of mobility, cognition, pain, self-care, usual activities, and affect, it may not be sufficiently robust to capture the nuances of what constitute good health. Perhaps, there is further opportunity to consider

development of additional vignette description sets specific to correcting for heterogeneity stemming from differential expectations, definitions and threshold of good health which may not be sufficiently represented or captured by the current vignette descriptions. Sturgis et al. (2001) discusses the range of “good health” meaning: being without illness, no restrictions in activities, physical and mental health, “normal health” of having a slight cold and flu. The additional vignettes are not categorized into specific health domains, but encompass an overall health experience to elicit people’s definition of good health, for example: “Paul has no visible signs of illness”; “Ingritta experiences no restrictions in daily activities”; “Lillian has slight cold and flu”, “Jodi has sound and peaceful sleep nightly”. Respondents are then asked to rate on a scale of 1 to 5 for each of these vignettes on the extent they think Paul, Ingritta, Lillian and Jodi have good health. Having specific set of vignette developed with the perspective of eliciting definition or personal threshold of good health serves three key purposes. First, it provides a descriptive context to understand why the respondents may be picking the choices they do in the EQ5D inspired vignettes used by WHO – MCS, World Health Survey and SHARE . Second, this allows for a two stage correction process, to decompose the heterogeneity that arises from subjective definition of good health versus heterogeneity that stems from other subjective aspects of SAH rating. The relevance of this decomposition is the opportunity to isolate differences in rating due to one’s perspective or reference point with regards to definition of health, which may not be related to issues of health equity. After first correcting for vignette domain rating using these additional good health definition vignettes, it is more plausible to attribute remaining heterogeneity as evidence of inequalities in health. Third, instead of the above two stage correction, these additional health definition vignettes may also be used as correction weights for responses to domain driven vignettes. The purpose is to account for these different references points and definitions used for good health by different individuals to weigh the responses given to domain vignettes, potentially resulting in more accurate representation of heterogeneity, untainted by original reference point of good health definition.

**2. Development of a validated universal measurement correction weights (using vignettes):** Just as EQ5D provides a general validated valuation of QOL values set, perhaps, similar initiatives can be undertaken for development of a validated universal measurement correction weights for different SES levels, and cultural nuances using vignette correction method. This correction values can then be conveniently applied to policies for more accurate measurement of SES inequalities, healthcare effectiveness evaluation and comparable cross country comparisons. This universal vignette correction weights will also benefit developing countries which may not have financial resources to collect such data. Further research can also be engaged to explore transferability criteria of such weights across countries and SES segments.

**3. The case for non-correction of heterogeneity:** Whilst correction of heterogeneity in SAH arising from biases are important when serving the purpose of SES inequality assessment and inputs for health policy effectiveness, SAH is also assessed in studies to see if it contains information about health which may subsequently predict mortality risk (Idler and Kasl, 1991; Idler, 1993; Wilcox et al., 1996). The non-objective health aspects that SAH captures may have specific predictive value, and SAH is to reflect a kind of “life spirit” and a person’s sense of existence (Stenback, 1964; Idler, 1993). Thus, depending on the objective of the study, not all SAH reporting heterogeneity may need needs to be eradicated. If one wants to tap into the predictive value of SAH, there may be value in preserving the subjectivity. The power of SAH is its containment of both objective and subjective health information. More qualitative research can be conducted to understand what drives this subjectivity in SAH, and its role in

complementing SAH's key strength of also being an objective carrier of health information, instead of simply correcting for its heterogeneity. Moreover, these qualitative findings can be ploughed back as inputs to enrich the current EQ5D inspired vignette descriptions - to better reflect the mindset of the respondents, when corrections are needed. Furthermore, it may be interesting to explore the tradeoff between a corrected SAH for more accurate portrayal of its true value and its predictive ability of mortality and morbidity risk i.e. establish the difference in predicative power of mortality/morbidity of a corrected versus non corrected SAH measure.

## **6.2 Implications of over confidence, optimism and use of self-concept as sources of bias in SAH**

Over-confidence and optimism creates a potential dual problem for respondents over time. On one hand, being over confident reduces perceived need for prevention of health issues (Weinstein, 1983). On the other hand, one may postulate a similar situation to QALY calculations where the patient over rates quality of life of current health state due to adaptation, leading to under-funding of treatment due to low elicited health benefits (van Gils et al., 2011). Here, policy makers using inflated over-confident SAH values may wrongly conclude no need to fund particular health initiatives to this over confident group. The dynamics of the two elements of respondents not engaging in preventive measures, and policy makers not funding certain health preventive aspects creates myopia leading to short run savings, but potential longer run cost, should actual illness emerge subsequently.

Much of research on evidence of over confidence and optimism on SAH has been studies done with elderly patients, survivors (Idler, 1993; Van Doorn, 1999; Maddox and Douglass, 1973) and studies on optimism in younger healthy respondents is mainly constructed through examining social peer comparisons (Kaplan and Baron-Epel, 2003). To some extent these findings on optimism in SAH ratings from survivorship and adaptation to illness may be subjected to selection bias.

The use of self-concept health in SAH rating is real and currently embedded within the subjectivity of SAH rating. However, research have yet to develop a deep understanding of the impact of use of such self-concept of health on the validity, objectivity and accuracy of SAH rating. The limited inquiry of self-concept of health in SAH lies mainly in the area of what constitutes SAH ratings, the health conceptualization models of use of medical, social, psychological, economic, spiritual conditions for health ratings (Lee, 2014; Larson, 1999; Kaplan and Baron-Epel, 2003; Smith et al., 1994) and testing for existence of use of a self-concept in SAH (Bailis et al., 2003; Kaplan and Baron-Epel, 2003; Smith et al., 1994).

Below highlights some opportunities for further research in this area of over-confidence, optimism and use of self-concept in SAH ratings:

**1. Identification of over confidence in SAH:** More research should be conducted amongst a wider target to understand the impact of over confidence and optimism on SAH, especially amongst those with no current ill health. Dunning et al., (2004) offered four heuristics of 1. Symmetry rule of no ill health is

indeed good health, 2. Prevalence rule where only rare symptoms are seen to be ill health, 3. Stress illness rule where discomfort are attributed to stress instead of actual ill health and 4. Stereotyping where certain illness are assumed to be associated with particular sets of people, as explanations for overconfidence and optimism in healthy people. However, there has been no empirical test of these heuristics. Perhaps, these heuristics could serve as starting points to construct experiments to examine causal effects of over confidence and optimism on SAH ratings. The four heuristics can be used to simulate over –confidence and optimism scenarios and randomly assign similarly healthy respondents into “normal” versus simulated “over-confidence” scenarios to see casual effect on SAH ratings.

**2. Feedback mechanisms to reduce over confidence and optimism:** Furthermore, provision of some form of feedback to encourage self-awareness within the SAH answering process may help introduce a sense of realism to reign in any excessive over-confidence and optimism.

**3. Stability of SAH ratings and removal of sources of biases in SAH rating due to over-confidence, optimism and use of self-concept of health, through use of panel data and fixed effects estimator:** Seemingly unrelated to over-confidence and optimism, there exist an element of affect in SAH ratings. Kaplan and Baron-Epel (2003) found 62.8% - 76.2% of respondents attributing SAH rating to their general feeling of health. This raises potential examination of internal validity of the SAH ratings as a respondents feelings about own health may change over time influenced by moods, external context, survey context etcetera. Thereby, it is plausible that a different SAH rating may be elicited from the same individual over two different days, leading to questioning of validity of SAH rating for policy evaluation, or health assessment. Further studies may be conducted to explore the short term stability of a SAH measurement and if there are any theories which may explain consistency of a SAH rating. Whilst SAH may not be stable, it may be interesting to explore the use of panel data and fixed effects estimators to examine this issue. If a person’s self-concept of health, personality of over-confidence and optimism is inherent and unchanging, these sources of biases may be purged through removal of time invariant individual heterogeneity using fixed effects estimators, leading to more consistent SAH rating across time periods. Any other variations in SAH that causes non consistency of SAH should then be due to actual objective health state changes. Moreover, if there exist further external context like moods and feelings which are truly random, then the SAH rating obtained will still be unbiased.

**4. Further research on determinants of self-concept of health:** Future exploration can include more empirical and/or experimental evidence of the determinants of self-concept of health.

**5. Examination of the relevance and feasibility of vignette correction for use of self-concept of health in SAH:** The relationship between use of self-concept of health in SAH with over-confidence and optimism can also be examined and this will have relevance on assessing the use of vignette correction approach to get rid of use of self-concept of health in SAH ratings. However, there are possibilities that self-concept of health may be subjected to potential violations of both the assumptions of “vignette equivalence: where the description is perceived by all to correspond to the same state” (Bago d’Uva et al., 2011) and “vignette consistency: where the individual uses the same response scales to rate the vignettes and their own situations” (Bago d’Uva et al., 2011) – raising question on the relevance of the vignette approach in correcting the use of self-concept of health. Due to the defensive nature of use of



self-concept of health and endowment effect, an individual may engage self-concept of health in rating own SAH but may not find it necessary to use it in rating the vignettes which does not pertain to self – thereby potential violation vignette consistency assumption. Possibly vignette equivalence assumption may also be violated, depending on the difference in the degree of detachment from self and apathy between individuals in rating these vignettes (i.e. if all individuals are equally detached or attached to self-concept of health then the vignette description may still be perceived by all to correspond to the same state). If the assumptions do not hold for use of self-concept of health in SAH ratings, then other forms of corrections will be needed.

### **6.3 Implications of framing as a source of bias in SAH**

Studies by Lumsdaine and Exterkate (2013), Bowling and Windsor (2008), Baron-Epel and Kaplan (2001), Greene et al., (2014) have shown evidence of wording and labelling framing effects on SAH ratings where response may change their ratings when exposed to a differently or fully labelled scale as opposed to a partially labelled scale. Response order and the resulting primacy and recency effects brought about by different modes of interview may also make SAH ratings incomparable across countries if the survey was executed in different way of phone versus paper face to face interviews. Question order effects exist in SAH ratings where SAH questions placed at the end of a questionnaire (Bowling and Windsor, 2008) or when SAH is asked twice in an interview were found to be more positively skewed, after isolating for mode of administration effects (Crossley and Kennedy, 2002). Overall, framing effects needs to be accounted when using SAH for policy making and assessment. For instance if different countries adopt: different versions of SAH questions such as WHO versus US versions; different modes of interviews to elicit SAH; different placement of the SAH in the questionnaire, respondents will be exposed to different framing effects and SES inequality may be wrongly concluded. Other than these studies of framing, research on framing effects in SAH health is still nascent. Below highlights some opportunity for further research:

**1. Call for more varied research studies on framing effects:** 1. Studies to assess differences in framing effects on different SES groups; 2. Studies on primacy effect where items listed first are more likely to be selected in visual format like face to face paper interview, versus recency effect where last items get better recalled in oral format like telephone interviews (Krosnick and Alwin; 1987; Janssen et al., 2008; Baumeister et al., 1998; Schwarz, 1999) and assessed response order effects on SAH as this literature review has not found studies examining response order effects on SAH. It may be that response order framing effect is less relevant from SAH as it is just a linear scale of 1 to 5, and this can be verified in future work.

**2. Vignette approach for detection and correction of framing effects:** Vignettes approach may also be used to detect and correct for framing effects on SAH ratings between different groups of respondents. The thought experiment in the above section highlights the use of the vignette approach to test for wording and labelling framing effects in reporting heterogeneity of under rating of health amongst high educated.

## 6.4 Implications of satisficing as a source of bias in SAH

Satisficing is probably one of the hardest source of biases to assess, find evidence and to correct. However, it is something that is very real, easily activated and justified by the respondent and relates to the lack of cognitive ability and/or willingness to put in thorough thought in responding to the SAH question. Much of this stems from the fact that there is neither a reward nor penalty for a truthful or thoughtful answer. Various response can be cultivated to discourage satisficing such as: clear wordings and fully labelled scales (Schechter, 1997; Conti and Pudney, 2012; Schwarz, 1999), provide explanations of what constitutes SAH rating (Toungangeau, 1999), open ended questions to record feelings about own health before answering SAH can be used to promote self-awareness and reduce satisficing (Fienberg et al., 1985; Hermann, 1995). Some preliminary thoughts on addressing satisficing are highlighted below.

**1. Self-identification:** Correcting for satisficed response could involve removal of self-reported inaccurate response. Hermann (1997) suggested giving options to respondents to state if their SAH rating was a “wild guess” or “educated guess”, and “wild guess” responses can just be removed from data.

**2. Controlling for related bias instead:** It is also possible to explore if there exist specific associations and relationship between specific bias and a specific outcome for satisficing (e.g. use of optimism being more associated with strong satisficing of status quo versus framing effects more associated with weak satisficing of selecting first acceptable response). Satisficing may then be indirectly corrected for by focusing on correcting for the related bias instead.

**3. Creation of penalty or reward structure/nudges:** Other considerations involves creation of penalty or reward within the survey or experiment structure to reel in satisficing behavior.

## 6.5 Research limitations and Conclusions on Literature Review

The pathways and sources of biases in SAH found in this literature review are constrained by current limitations of specific sources of bias research in the SAH ratings arena. Whilst inferences have also been drawn from other disciplines such as survey methodologies and judgement and decision making studies, these may not been directly explored in SAH rating studies.

For future research in sources of bias for SAH, the definition of bias used will need to be considered. Although this literature review is anchored on sources of biases to SAH to account for the complexity of the experience of SAH ratings and its interconnectedness with other elements to include discussions on effects/ heuristics/motivations/moods and conditions, violations which may / may not lead to actual biases, it is still important to discuss the definition of bias as defined in this paper in section 3.5.

In this thesis, a bias is about the use of irrelevant information and/or emotions and motivations and/or influence by moods driving wrong judgment/outcome (Kahneman, 2003) and contributing to different norms and expectations of health between individuals (Carr et al., 2001; Sen, 2002, Salomon et al., 2004;

Dunning et al., 2004), leading to a deviation from a normatively correct answer (Keren, 1990). A clear distinction has also been made between using a heuristic versus an actual bias.

The discussion on sources of biases presented in this paper are based on the respondent's application of a heuristic due to lack of cognitive power. Some critics argue that the use of judgment heuristics may not lead to bias as it has been explored in a static condition, versus in a dynamic environment where it could be viewed as functional and part of a continuous process (Keren, 1990; Hogarth, 1981). Moreover, not all framing effects are about biases. Some could just be due to different stakeholders of respondent and researcher holding different perspectives to the problem space (Keren, 1990). Whilst there may be different theories on what may lead to a bias, for something to be defined as a bias, it needs to be compared to a "unique correct solution", and currently there lacks a universal criterion for which the biases can be benchmarked against.

Overall, this literature review serves to highlight three key points. Firstly, SAH is all about the context of the individual (who they are inherently, what influences one's health experience and expectations), and the context of the question (framing of the question) and survey (the mode of interview and survey experience). Just as subjectivity of SAH is all about the importance of context, framing effects highlight the same red flag. It is important for policy makers and researchers to keep this in mind when measuring and assessing impact of SAH in SES inequalities and healthcare policy evaluations and country comparisons.

Secondly, more research should be conducted in the areas of assessing presence of these biases (availability of objective health information, subjective health information, use of self-concept in health, over-confidence and optimism, framing effects, social desirability, and satisficing), its impact on SAH evaluation and correction using other econometric methods. More exploration can also be done to further exploit the power of the vignette correction method for other biases, and perhaps an even a broader vision of working towards development of a universal correction weights for different SES and cultural inequalities.

Lastly, it is important to realise that SAH as a measurement tool was not developed from theories, but as a tool that is easy and cheap to gather information on health (Lee, 2014). Although very simple, SAH is a very powerful tool and its theoretical value on its ability to predict mortality and morbidity was discovered through empirical descriptive studies (Lee, 2014). There is however still limited theoretical understanding of SAH and most of this is through application the cognitive aspect of survey methodology on process of answering SAH (Lee, 2014). Having a deeper understanding of the sources of biases in SAH and theories around these biases may serve to expand the scope for theorising SAH. Other areas of theoretical support include collaboration between behavioral theorist, health economist, and survey designers to assess possibility of applying Prospect Theory to the framing effects of a SAH measurement. Possible avenues of exploration include: First establishing how reference points are formed in self - assessed health evaluation in the SAH scale and what causes it to shift. Perhaps, the way the scales are labelled may influence reference point formation of self assessed health evaluation especially in the presence/absence of a mid-point or neutral option (e.g. fair); Second, deeper studies into understanding what drives a health experience to be considered as gains versus loss and how these are influenced by initial health expectations. This will cumulate into derivation of a risk seeking (convex) versus risk averse (concave) utility function for health as well as risk attitudes in the probability

weighting reflecting optimism/pessimism in SAH ratings; Third, this can also be further applied to study rating of SAH in chronic versus varying health states; Fourth, how incorporating Prospect Theory in SAH measurements can account for sources of biases and identification of which ones are addressed.

To this end, this thesis hopes to inspire and nurture a starting ground to further exploration of detection and correction of SAH biases and strengthen theoretical development of this very simple, yet special and important SAH measurement tool.

## References:

- Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2), 179-211
- Alexandrov, A. (2010). Characteristics of single-item measures in Likert scale format. *The Electronic Journal of Business Research Methods*, 8(1), 1-12
- Anand, S. (2008). Satisficing. In Paul J. Lavrakas (Ed.), *Encyclopedia of Survey Research Methods*. (pp. 798-800). Thousand Oaks, CA: Sage Publications, Inc. doi: <http://dx.doi.org/10.4135/9781412963947.n518>
- Andrews, F. M., & McKennell, A. C. (1980). Measures of self-reported well-being: Their affective, cognitive, and other components. *Social Indicators Research*, 8(2), 127-155.
- Apesteguia, J., & Ballester, M. A. (2009). A theory of reference-dependent behavior. *Economic Theory*, 40(3), 427-455
- Aquilino, W. S. (1994). Interview mode effects in surveys of drug and alcohol use: A field experiment. *Public opinion quarterly*, 58(2), 210-240.
- Ariely, D. (2008). *Predictably irrational* (p. 20). New York: HarperCollins
- Bago d'Uva, T.B., O'Donnell, O., and Van Doorslaer, E. (2008a). 'Differential Health Reporting by Education Level and its Impact on the Measurement of Health Inequalities among Older Europeans'. *International Journal of Epidemiology*, 37(6): 1375-83.
- Bago d'Uva, T., Van Doorslaer, E., Lindeboom, M., and O'Donnell, O. (2008b). 'Does Reporting Heterogeneity Bias the Measurement of Health Disparities?' *Health Economics*, 17(3): 351-75.
- Bago d'Uva, T.B., Lindeboom, M., O'Donnell, O. and Van Doorslaer, E. (2011). 'Slipping Anchor? Testing the Vignettes Approach to Identification and Correction of Reporting Heterogeneity'. *Journal of Human Resources*, 46(4): 875-906
- Bailis, D. S., Segall, A., & Chipperfield, J. G. (2003). Two views of self-rated general health status. *Social science & medicine*, 56(2), 203-217
- Baillon, A. (2015). *Subjective Truths* (No. EIA-2015-063-MKT). Erasmus Research Institute of Management (ERIM), ERIM is the joint research institute of the Rotterdam School of Management, Erasmus University and the Erasmus School of Economics (ESE) at Erasmus University Rotterdam.
- Baker, M., Stabile, M., and Deri, C. (2004). 'What do self-reported, objective, measures of health measure?' *Journal of Human Resources*, 39, 1067-1093.
- Bandura, A. (1978). The self system in reciprocal determinism. *American psychologist*, 33(4), 344.
- Baron-Epel, O., Weinstein, R., Haviv-Mesika, A., Garty-Sandalon, N., & Green, M. S. (2008). Individual-level analysis of social capital and health: a comparison of Arab and Jewish Israelis. *Social science & medicine*, 66(4), 900-910.
- Baron-Epel, O., & Kaplan, G. (2001). General subjective health status or age-related subjective health status: does it make a difference?. *Social Science & Medicine*, 53(10), 1373-1381.

- Baron, J., Beattie, J., & Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning: II. Congruence, information, and certainty. *Organizational Behavior and Human Decision Processes*, 42(1), 88-110.
- Bastardi, A., Uhlmann, E. L., & Ross, L. (2011). Wishful Thinking Belief, Desire, and the Motivated Evaluation of Scientific Evidence. *Psychological Science*, 22(6), 731-732
- Bassett, W., R.L., 1999. Outlook F., Lumsdaine, Outcomes, and Optimism. Unpublished manuscript
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American psychologist*, 54(7), 462.
- Bargh, J. A. (1994). The four horsemen of automaticity: Intention, awareness, efficiency, and control as separate issues.
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: is the active self a limited resource?. *Journal of personality and social psychology*, 74(5), 1252
- Bazerman, M., & Moore, D. A. (2009). *Judgment in managerial decision making*. New York, N.Y., Wiley
- Beatty, P., & Schechter, S. (1994). An examination of mode effects in cognitive laboratory research. In *Proceedings of the Section on Survey Research Methods, American Statistical Association* (pp. 1275-1280).
- Beatty Douglas Herrmann Cathy Puskar Jeffrey Kerwin, P. (1998). 'Don't Know' Responses in Surveys: Is What I Know What You Want to Know and Do I Want You to Know It?. *Memory*, 6(4), 407-426.
- Bénabou, R., & Tirole, J. (2005). Incentives and prosocial behavior (No. w11535). National Bureau of Economic Research.
- Benartzi, S., & Thaler, R. H. (1993). Myopic loss aversion and the equity premium puzzle (No. w4369). National Bureau of Economic Research.
- Bickel, W. K., & Vuchinich, R. E. (Eds.). (2000). *Reframing health behavior change with behavioral economics*. Psychology Press
- Blaxter, M. (2003). *Health and lifestyles*. Routledge
- Blishen, B., & Atkinson, T. (1980). Anglophone and Francophone differences in perceptions of the quality of life in Canada. *The Quality of Life: Comparative Studies* (Sage, London), 25-40.
- Bobak, M., Pikhart, H., Rose, R., Hertzman, C., & 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(9), 1343-1350
- Borawski, E. A., Kinney, J. M., & Kahana, E. (1996). The meaning of older adults' health appraisals: Congruence with health status and determinant of mortality. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 51(3), S157-S170.
- Bound J. (1991). 'Self-reported versus objective measures of health in retirement models', *Journal of Human Resources*, 26(1): 106-138.

Bowling, A., & Windsor, J. (2008). The effects of question order and response-choice on self-rated health status in the English Longitudinal Study of Ageing (ELSA). *Journal of epidemiology and community health*, 62(1), 81-85.

Bowling, A. (2014). *Research methods in health: investigating health and health services*. McGraw-Hill Education (UK)

Bradburn, N. M. (1983). Response effects. *Handbook of survey research*, 1, 289-328.

Brewer, N. T., Hallman, W. K., Fiedler, N., & Kipen, H. M. (2004). Why do people report better health by phone than by mail?. *Medical care*, 42(9), 875-883.

Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative?. *Journal of personality and social psychology*, 36(8), 917.

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: mindfulness and its role in psychological well-being. *Journal of personality and social psychology*, 84(4), 822.

Butler, J.S., Burkhauser, R.V., Mitchell, J.M. and Pincus, T.P. (1987). 'Measurement Error in Self-Reported Health Variables', *The Review of Economics and Statistics*, 69(4): 644-

Burger Jerry M. The foot-in-the-door compliance procedure. A multiple-process analysis and review. *Personal Soc Psychol Rev* 1999;3(4):303–25

Cabinet Office and Institute for Government (2010), *MINDSPACE: Influencing Behaviour through Public Policy*

Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). *The quality of American life: Perceptions, evaluations, and satisfactions*. Russell Sage Foundation.

Campbell, J. D., Trapnell, P. D., Heine, S. J., Katz, I. M., Lavalley, L. F., & Lehman, D. R. (1996). Self-concept clarity: Measurement, personality correlates, and cultural boundaries. *Journal of personality and social psychology*, 70(1), 141.

Campbell, J. D. (1986). Similarity and uniqueness: the effects of attribute type, relevance, and individual differences in self-esteem and depression. *Journal of personality and social psychology*, 50(2), 281.

Caputo, A. (2013). A literature review of cognitive biases in negotiation processes. *International Journal of Conflict Management*, 24(4), 374-398.

Carr, A. J., Gibson, B., & Robinson, P. G. (2001). Is quality of life determined by expectations or experience?. *Bmj*, 322(7296), 1240-1243.

Chandran, S., & Menon, G. (2004). When a day means more than a year: Effects of temporal framing on judgments of health risk. *Journal of Consumer Research*, 31(2), 375-389.

Chartrand, T. L. (2005). The role of conscious awareness in consumer behavior. *Journal of Consumer Psychology*, 15(3), 203-210

Chipperfield, J. G. (1993). Incongruence between Health Perceptions and Health Problems Implications for Survival among Seniors. *Journal of Aging and Health*, 5(4), 475-496.

- Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annu. Rev. Psychol.*, 55, 591-621.
- Conti, G., & Pudney, S. (2011). Survey design and the analysis of satisfaction. *Review of Economics and Statistics*, 93(3), 1087-1093.
- Coleman, J. S., & Coleman, J. S. (1994). *Foundations of social theory*. Harvard university press.
- Costa, P. T., McCrae, R. R., & Zonderman, A. B. (1987). Environmental and dispositional influences on well-being: Longitudinal follow-up of an American national sample. *British Journal of Psychology*, 78(3), 299-306
- Cronbach, L.J. (1950). 'Further Evidence on Response Sets and Test Design', *Educational and Psychological Measurement*, 10(1): 3-31.
- Crossley, T. F., & Kennedy, S. (2002). The reliability of self-assessed health status. *Journal of health economics*, 21(4), 643-658
- Cutler, D. M., Richardson, E., Keeler, T. E., & Staiger, D. (1997). Measuring the health of the US population. *Brookings papers on economic activity. Microeconomics*, 217-282.
- Curt, C. L., & Zechmeister, E. B. (1984). Primacy, recency, and the availability heuristic. *Bulletin of the Psychonomic Society*, 22(3), 177-179
- Caroll GD, Choi JJ, Laibson D, et al. Optimal defaults and active decisions. *Q J Econ.*2009; 124(4):1639-74.
- Crosen, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic literature*, 448-474.
- Crowne, D. P., & Marlowe, D. (1964). *The approval motive: Studies in evaluative dependence* (p. 27). New York: Wiley.
- Damian, J., Ruigomez, A., Pastor, V., & Martin-Moreno, J. M. (1999). Determinants of self assessed health among Spanish older people living at home. *Journal of Epidemiology and Community Health*, 53(7), 412-416
- Danielsen, R. Cognitive biases in the five most disputed terms; drivers; consequences; debiasing techniques
- Daykin, A. R., & Moffatt, P. G. (2002). Analyzing ordered responses: A review of the ordered probit model. *Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences*, 1(3), 157-166.
- Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology/Psychologie canadienne*, 49(3), 182.
- Deeg, D. J., & Kriegsman, D. M. (2003). Concepts of self-rated health: specifying the gender difference in mortality risk. *The Gerontologist*, 43(3), 376-386.
- Delquié, P. (1993). Inconsistent trade-offs between attributes: New evidence in preference assessment biases. *Management Science*, 39(11), 1382-1395.



- DeSalvo, K. B., Bloser, N., Reynolds, K., He, J., & Muntner, P. (2006). Mortality Prediction with a Single General Self-Rated Health Question. *Journal of general internal medicine*, 21(3), 267-275.
- DeWall, C. N., Baumeister, R. F., & Masicampo, E. J. (2008). Evidence that logical reasoning depends on conscious processing. *Consciousness and cognition*, 17(3), 628-645.
- Diener, E. (1994). Assessing subjective well-being: Progress and opportunities. *Social indicators research*, 31(2), 103-157.
- Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & Van Baaren, R. B. (2006). On making the right choice: The deliberation-without-attention effect. *Science*, 311(5763), 1005-1007.
- Douglas R. Hofstadter. (1985). *Metamagical themas: Questing for the essence of mind and pattern*. Section VII: "Sanity and Survival"
- Dowd, J. B., & Zajacova, A. (2010). Does self-rated health mean the same thing across socioeconomic groups? Evidence from biomarker data. *Annals of epidemiology*, 20(10), 743-749.
- Duffy, F. D., & Holmboe, E. S. (2006). Self-assessment in lifelong learning and improving performance in practice: physician know thyself. *Jama*, 296(9), 1137-1139.
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment implications for health, education, and the workplace. *Psychological science in the public interest*, 5(3), 69-106
- Eccles, J. S., Wigfield, A., Flanagan, C. A., Miller, C., Reuman, D. A., & Yee, D. (1989). Self-concepts, domain values, and self-esteem: Relations and changes at early adolescence. *Journal of personality*, 57(2), 283-310.
- Emmons, R. A. (1986). Personal strivings: An approach to personality and subjective well-being. *Journal of Personality and Social psychology*, 51(5), 1058
- Engelmann, Dirk, and Martin Strobel (2004), "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments," *American Economic Review*, 94, 857-869.
- Etile, F., and Milcent, C. (2006). 'Income-related Reporting Heterogeneity in Self-Assessed Health: Evidence from France'. *Health Economics*, 15: 965-81
- Ettner, S.L., 1996. New evidence on the relationship between income and health. *Journal of Health Economics* 15, 67–85.
- Evans, A. M., Dillon, K. D., Goldin, G., & Krueger, J. I. (2011). Trust and self-control: The moderating role of the default. *Judgment and Decision Making*, 6(7), 697-705
- Ezzati M., Martin H., Skjold S., Vander Hoorn, S. and Murray C.J. (2006). 'Trends in national and state-level obesity in the USA after correction for self-reported bias: analysis of health surveys', *Journal of the Royal Society of Medicine*, 99(5):250–257
- Fayers, P. M., & Sprangers, M. A. (2002). Understanding self-rated health. *The Lancet*, 359(9302), 187-188.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly journal of Economics*, 817-868.

- Fehr, E., & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism—experimental evidence and new theories. *Handbook of the economics of giving, altruism and reciprocity*, 1, Chapter 8
- Ferraro, K. F. (1980). Self-ratings of health among the old and the old-old. *Journal of Health and Social Behavior*, 377-383.
- Festinger, L. (1954). A theory of social comparison processes. *Human relations*, 7(2), 117-140
- Fienberg, S. E., Loftus, E. F., & Tanur, J. M. (1985). Cognitive aspects of health survey methodology: an overview. *The Milbank Memorial Fund Quarterly. Health and Society*, 547-564.
- Fischer, G. W., & Hawkins, S. A. (1993). Strategy compatibility, scale compatibility, and the prominence effect. *Journal of Experimental Psychology: Human Perception and Performance*, 19(3), 580.
- Fiske, S. T., & Taylor, S. E. (2013). *Social cognition: From brains to culture*. Sage.
- Finucane, M. L., Slovic, P., Mertz, C. K., Flynn, J., & Satterfield, T. A. (2000). Gender, race, and perceived risk: The 'white male' effect. *Health, risk & society*, 2(2), 159-172.
- Frank, R. G. (2004). Behavioral economics and health economics (No. w10881). National Bureau of Economic Research
- Frederick, Shane, George Loewenstein, and Ted O'Donoghue (2002), "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40, 351-401.
- Fry, T.R.L. and Harris, M.N. (2005). 'The DOGIT Ordered Generalised Extreme Value Model', *Australian and New Zealand Journal of Statistics*, 47(4): 531-542
- Gal, D., & Rucker, D. D. (2011). Answering the unasked question: Response substitution in consumer surveys. *Journal of Marketing Research*, 48(1), 185-195
- Gallese, V. 2003. The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical Transactions of the Royal Society*, 358: 517-528.
- Garland, R. (1991). 'The Mid-Point on a Rating Scale: Is it Desirable?', *Marketing Bulletin*, 2, 66-70, Research Note 3..
- Geertz, C. (1973). Thick description: Toward an interpretive theory of culture. In *The interpretation of cultures. Selected essays* (pp. 1-30). New-York: Basic Books.
- Gneezy, Uri, and Jan Potters (1997), "An Experiment on Risk Taking and Evaluation Periods," *Quarterly Journal of Economics*, 112, 631-645
- Gold, M., Franks, P., & Erickson, P. (1996). Assessing the health of the nation: the predictive validity of a preference-based measure and self-rated health. *Medical care*, 34(2), 163-177.
- Greene, W. H., Harris, M. N., & Hollingsworth, B. (2014). Inflated Responses in Measures of Self-Assessed Health
- Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American psychologist*, 35(7), 603.

- Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *The Quarterly Journal of Economics*, 537-557
- Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of health economics*, 19(3), 403-420.
- Groves, R. M., Fultz, N. H., & Martin, E. (1992). Direct questioning about comprehension in a survey setting. *Questions about questions: inquiries into the cognitive bases of surveys*. New York: Russell Sage, 49-61.
- Hanita, M. (2000). Self-report measures of patient utility: should we trust them?. *Journal of clinical epidemiology*, 53(5), 469-476.
- Helliwell, J. F., & Putnam, R. D. (2004). The social context of well-being. *Philosophical transactions-royal society of London series B biological sciences*, 1435-1446.
- Herrmann, D. (1995). Reporting current, past, and changed health status: What we know about distortion. *Medical care*, AS89-AS94
- Helweg-Larsen, M., & Shepperd, J. A. (2001). Do moderators of the optimistic bias affect personal or target risk estimates? A review of the literature. *Personality and Social Psychology Review*, 5(1), 74-95.
- Henslin, J. M. (1967). Craps and magic. *American Journal of Sociology*, 316-330
- Hertzman, C., Power, C., Matthews, S., & Manor, O. (2001). Using an interactive framework of society and lifecourse to explain self-rated health in early adulthood. *Social science & medicine*, 53(12), 1575-1585
- Holbrook, A. (2008). Response Order Effects. In Paul J. Lavrakas (Ed.), *Encyclopedia of Survey Research Methods*. (pp. 755-757). Thousand Oaks, CA: Sage Publications, Inc. doi: <http://dx.doi.org/10.4135/9781412963947.n488>
- Hofstadter, D. H. (1980). Gödel, Escher, Bach: An Eternal Golden Braid;. Section VI "The location of meaning". Section XX: "Strange loops or Tangled Hierarchies"
- Idler, E.L., Kasl, S.V., 1995. Self-ratings of health: do they also predict change in functional ability? *Journal of Gerontology* 50B, S344–S353
- Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of health and social behavior*, 21-37.
- Idler, E. L., Hudson, S. V., & Leventhal, H. (1999). The Meanings of Self-Ratings of Health A Qualitative and Quantitative Approach. *Research on Aging*,21(3), 458-476.
- Idler, E.L. and S. Kasl. 1991. "Health Perceptions and Survival: Do Global Evaluations of Health Status Really Predict Mortality?" *Journal of Gerontology: Social Sciences* 46:S55-S65.
- Idler, E. L. (1993). Age differences in self-assessments of health: Age changes, cohort differences, or survivorship?. *Journal of Gerontology*, 48(6), S289-S300.
- Idler, E., Leventhal, H., McLaughlin, J., & Leventhal, E. (2004). In sickness but not in health: self-ratings, identity, and mortality. *Journal of Health and Social Behavior*, 45(3), 336-356.

- Idler, E. L. (1999). Guest Editor's Introduction Self-Assessments of Health: The Next Stage of Studies. *Research on Aging*, 21(3), 387-39
- Izen, A. M., Shalke, T. E., Clark, M., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior: A cognitive loop?. *Journal of personality and social psychology*, 36(1), 1.
- Janssen, L., Fennis, B. M., Pruyn, A. T. H., & Vohs, K. D. (2008). The path of least resistance: Regulatory resource depletion and the effectiveness of social influence techniques. *Journal of Business Research*, 61(10), 1041-1045.
- Janis, I. L., & Mann, L. (1982). A theoretical framework for decision counseling. *Counseling on personal decisions: Theory and research on short-term helping relationships*, 47-72.
- Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. Free Press.
- JUNGERMANN, H., SCHOLZ, R.W. (ED.) 1983. *Decision making under uncertainty ---The two camps on rationality*. Elsevier Science Publisher B.V. (North---Holland).
- Jylhä, M. (2009). What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Social science & medicine*, 69(3), 307-316
- Jobe, J. B., & Mingay, D. J. (1991). Cognition and survey measurement: History and overview. *Applied Cognitive Psychology*, 5(3), 175-192.
- Johnson, Eric J., et al. (2012) "Beyond nudges: Tools of a choice architecture." *Marketing Letters* 23.2: 487-504.
- Johnston, D. W., Propper, C., & Shields, M. A. (2009). Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient. *Journal of health economics*, 28(3), 540-552.
- Jungermann, H. (1983). The two camps on rationality. In R.W. Scholz (ed.), *Decision Making under Uncertainty*, (pp. 63-86). Amsterdam: Elsevier
- Jürges, H., Avendano, M., Mackenbach, J.P., 2008. Are different measures of self-rated health comparable? An assessment in five European countries. *European Journal of Epidemiology* 23 (12), 773–781
- Jürges, H. (2007). True health vs response styles: exploring cross-country differences in self-reported health. *Health economics*, 16(2), 163-178.
- Jylhä, M., Guralnik, J. M., Ferrucci, L., Jokela, J., & Heikkinen, E. (1998). Is self-rated health comparable across cultures and genders?. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 53(3), S144-S152.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *The journal of economic perspectives*, 193-206
- Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *American psychologist*, 58(9), 697.

- Kahneman and Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, 47, 263-292.
- Kahneman, D., & Tversky, A. (1982). Variants of uncertainty. *Cognition*, 11(2), 143-157.
- Kahneman, D., & Ritov, I. (1994). Determinants of stated willingness to pay for public goods: A study in the headline method. *Journal of Risk and Uncertainty*, 9(1), 5-37.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Kaplan, G., & Baron-Epel, O. (2003). What lies behind the subjective evaluation of health status?. *Social Science & Medicine*, 56(8), 1669-1676.
- Kawachi, I., Kennedy, B. P., & Glass, R. (1999). Social capital and self-rated health: a contextual analysis. *American journal of public health*, 89(8), 1187-1193
- Kessler, Judd B., and C. Yiwei Zhang, forthcoming. "Behavioral Economics and Health."
- Keller, S. D., & Ware, J. E. (1996). Questions and answers about SF-36 and SF-12. *Medical Outcomes Trust Bulletin*, 4(3), 1-13
- Kerkhofs, M. and Lindeboom, M. (1995). 'Subjective health measures and state dependent reporting errors', *Health Economics*, 4: 221-235
- Kierkegaard, Soren (1843) *Fear and Trembling*
- Kim Ly, Nina Mažar, Min Zhao and Dilip Soman (2013) "A practioner's guide to nudging" Rotman School of Management, University of Toronto
- King, G., Murray, C. J., Salomon, J. A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American political science review*, 98(01), 191-207
- Keren, G. (1990). Cognitive aids and debiasing methods: can cognitive pills cure cognitive ills?. *Advances in Psychology*, 68, 523-552
- Klesges, R.C., Debon, M., Ray, J.W. (1995) 'Are self-reports of smoking rate biased? Evidence from the Second National Health and Nutrition Examination Survey' *Journal of Clinical Epidemiology*, 48(10): 1225-1233.
- Knowles, E. S., & Condon, C. A. (1999). Why people say "yes": A dual-process theory of acquiescence. *Journal of Personality and Social Psychology*, 77(2), 379
- Krause, N. M., & Jay, G. M. (1994). What Do Global Self-Rated Health Items Measure?. *Medical care*, 32(9), 930-942
- Kristensen, N. and N. Westergaard-Nielsen (2007), "Reliability of job satisfaction measures", *Journal of Happiness Studies*, 8, 273-292.
- Krosnick, J.A. and Alwin, D.F. (1987). 'An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement', *Public Opinion Quarterly*, 51(2): 201-219. *Cognition and Survey Research*, 111-131, New York: Wiley

- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied cognitive psychology*, 5(3), 213-236
- Krosnick, J. A. (1999). Survey research. *Annual review of psychology*, 50(1), 537-567.
- Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. *Handbook of survey research*, 2, 263-314
- Krosnick, J. A. (1999). Maximizing questionnaire quality. *Measures of political attitudes*, 2, 37-58
- Krueger, A.B. and D.A. Schkade (2008), "The reliability of subjective well-being measures", *Journal of Public Economics*, 92(8-9), 1833-1845.
- KRISTOFFERSEN, I. (2010). The metrics of subjective wellbeing: Cardinality, neutrality and additivity\*. *Economic Record*, 86(272), 98-123
- Lahelma, E., Martikainen, P., Laaksonen, M., & Aittomäki, A. (2004). Pathways between Socioeconomic Determinants of Health. *Journal of Epidemiology and Community Health* (1979-), 327-332.
- Langer, E. J. (1975). The illusion of control. *Journal of personality and social psychology*, 32(2), 311
- Langendyk, V. (2006). Not knowing that they do not know: self-assessment accuracy of third-year medical students. *Medical education*, 40(2), 173-179.
- Larson, J. S. (1999). The conceptualization of health. *Medical Care Research and Review*, 56(2), 123-136.
- Lee, S. (2014) Self-Rated Health in Health Surveys, in *Health Survey Methods* (ed T. P. Johnson), John Wiley & Sons, Inc., Hoboken, NJ, USA. doi: 10.1002/9781118594629.ch8
- Lessler, J., Tourangeau, R. and Salter, W. (1989). Questionnaire design in the cognitive research laboratory. *Vital and Health Statistics, Series 6, No. 1* (DHHS Publication No. PHS 89- 1076). Washington, DC: US. Government Printing Office.
- Leventhal, E. A. (1984). Aging and the perception of illness. *Research on Aging*, 6(1), 119-135.
- Lewinsohn, P. M., Mischel, W., Chaplin, W., & Barton, R. (1980). Social competence and depression: the role of illusory self-perceptions. *Journal of abnormal psychology*, 89(2), 203.
- Lindeboom M, van Doorslaer E. Cut-point shift and index shift in self-reported health. *J Health Econ* 2004; 23: 1083-1099.
- Little, D. (2012). Explanatory autonomy and Coleman's boat. *THEORIA. Revista de Teoría, Historia y Fundamentos de la Ciencia*, 27(2), 137-151.
- Locander, W., Sudman, S., & Bradburn, N. (1976). An investigation of interview method, threat and response distortion. *Journal of the American Statistical Association*, 71(354), 269-275.
- Lumsdaine, R. L., & Exterkate, A. (2013). How survey design affects self-assessed health responses in the Survey of Health, Ageing, and Retirement in Europe (SHARE). *European Economic Review*, 63, 299-307

- Lunde, L., and Løken, K.V. (2011). "How Are You Feeling"? Assessing Reporting Bias in a Subjective Measure of Health by Quantile Regression'. Working Paper 08/11. Bergen: University of Bergen
- Maddox, G. L., & Douglass, E. B. (1973). Self-assessment of health: A longitudinal study of elderly subjects. *Journal of Health and Social Behavior*, 87-93.
- Miilunpalo, S., Vuori, I., Oja, P., Pasanen, M., & Urponen, H. (1997). Self-rated health status as a health measure: the predictive value of self-reported health status on the use of physician services and on mortality in the working-age population. *Journal of clinical epidemiology*, 50(5), 517-528
- Manderbacka, K. (1998). Examining what self-rated health question is understood to mean by respondents. *Scandinavian Journal of Public Health*, 26(2), 145-153.
- Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual review of psychology*, 38(1), 299-337
- Marsh, H. W., Parada, R. H., & Ayotte, V. (2004). A multidimensional perspective of relations between self-concept (Self Description Questionnaire II) and adolescent mental health (Youth Self-Report). *Psychological Assessment*, 16(1), 27.
- Marsh, H. W. (1990). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review*, 2(2), 77-172.
- Matell, M.S. and Jacoby, J. (1972). 'Is there an Optimal Number of Alternatives for Likert Scale Items? Effects of Testing Time and Scale Properties', *Journal of Applied Psychology*, 56(6), 506-509.
- McClendon, M. J. (1991). Acquiescence and recency response-order effects in interview surveys. *Sociological Methods & Research*, 20(1), 60-103
- Means, B. Autobiographical memory for health-related events. US Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Center for Health Statistics, 1989.
- Meijer, E., Kapteyn, A., Andreyeva, T., 2011. Internationally comparable health indices. *Health Economics* 20, 600–619
- Mellers, B. A. (1998). Decision-Making. *Encyclopedia of Cognitive Science*
- Milkman, K. L., Chugh, D., & Bazerman, M. H. (2009). How can decision making be improved?. *Perspectives on Psychological Science*, 4(4), 379-383.
- Mosca, I., Bhuachalla, B. N., & Kenny, R. A. (2013). Explaining significant differences in subjective and objective measures of cardiovascular health: evidence for the socioeconomic gradient in a population-based study. *BMC cardiovascular disorders*, 13(1), 64.
- Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological review*, 115(2), 502.
- Mossey, J. M., & Shapiro, E. (1982). Self-rated health: a predictor of mortality among the elderly. *American journal of public health*, 72(8), 800-808.
- Mynttinen, S., Sundström, A., Koivukoski, M., Hakuli, K., Keskinen, E., & Henriksson, W. (2009). Are novice drivers overconfident? A comparison of self-assessed and examiner-assessed driver

competences in a Finnish and Swedish sample. *Transportation research part F: traffic psychology and behaviour*, 12(2), 120-130.

Mu, R. (2014). 'Regional Disparities in Self-Reported Health: Evidence from Chinese Older Adults'. *Health Economics*, 23(5): 529-49.

Mullainathan, S., & Thaler, R. H. (2000). Behavioral economics (No. w7948). National Bureau of Economic Research.

Musashi, M., & Munenori, Y. (2005). *The book of five rings: A classic text on the Japanese way of the sword*. Translated by, Thomas Cleary, Shambhala Publications, Inc.

Murray CJL. 1996. Epidemiology and morbidity transitions in India. In *Health, Poverty and Development in India*, Dasgupta M, Chen CL, Krishnan TN (eds). Oxford University Press: Delhi, 122–147.

Murray CJL, Ozaltin E, Tandon A, Salomon J et al. 2003. Empirical evaluation of the anchoring vignettes approach in health surveys. In *Health Systems Performance Assessment: Debates, Methods and Empiricism*, Murray CJL, Evans DB (eds). World Health Organization: Geneva

Neal, D. T., Wood, W., Labrecque, J. S., & Lally, P. (2012). How do habits guide behavior? Perceived and actual triggers of habits in daily life. *Journal of Experimental Social Psychology*, 48(2), 492-498

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2), 175. P 175-177

Nieminen, T., Martelin, T., Koskinen, S., Aro, H., Alanen, E., & Hyyppä, M. T. (2010). Social capital as a determinant of self-rated health and psychological well-being. *International Journal of Public Health*, 55(6), 531-542

Niedenthal, P. M., Cantor, N., & Kihlstrom, J. F. (1985). Prototype matching: A strategy for social decision making. *Journal of Personality and Social Psychology*, 48(3), 575.

Nisbett, R., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.

Nygren, T. E., Isen, A. M., Taylor, P. J., & Dulin, J. (1996). The influence of positive affect on the decision rule in risk situations: Focus on outcome (and especially avoidance of loss) rather than probability. *Organizational behavior and human decision processes*, 66(1), 59-72.

O'Donoghue, Ted, and Matthew Rabin (1999), "Doing it now or later," *American Economic Review*, 89, 103-124

Oldendick, R. (2008). Response Alternatives. In Paul J. Lavrakas (Ed.), *Encyclopedia of Survey Research Methods*. (pp. 750-752). Thousand Oaks, CA: Sage Publications, Inc. doi: <http://dx.doi.org/10.4135/9781412963947.n485>

O'Muircheartaigh, C., Gaskell, G., & Wright, D. B. (1995). Weighing Anchors: Verbal and Numeric Labels for Response Scales. *Journal of Official Statistics*, 11(3), 295-307.

O'Muircheartaigh, C., Krosnick, J. A., & Helic, A. (2000). Middle alternatives, acquiescence, and the quality of questionnaire data. *Annu. Meet. Am. Assoc. Public Opin. Res.*, Fort Lauderdale, FL.



- Okun, M. A., Stock, W. A., Haring, M. J., & Witter, R. A. (1984). Health and subjective well-being: A meta-analysis. *The International journal of aging and human development*, 19(2), 111-132.
- Park, J., & Banaji, M. R. (2000). Mood and heuristics: the influence of happy and sad states on sensitivity and bias in stereotyping. *Journal of personality and social psychology*, 78(6), 1005.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of personality and social psychology*, 46(3), 598
- Pfarr, C., Schmid, A., & Schneider, U. (2012). Reporting heterogeneity in self-assessed health among elderly Europeans. *Health economics review*, 2(1), 1-14.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology*, 88(5), 879.
- Pohl, R. (Ed.). (2004). *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory*. Psychology Press
- Porter Abbott, H. (2008). *The Cambridge Introduction to Narrative* (2nd edition). (pp. 1-27; 67-82). Cambridge: Cambridge University Press.
- Powdthavee, N. (2009). Ill-health as a household norm: evidence from other people's health problems. *Social Science & Medicine*, 68(2), 251-259.
- Powell, M., & Ansic, D. (1997). Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of economic psychology*, 18(6), 605-628.
- Pratkanis, A. R. (1989). The cognitive representation of attitudes. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 71-98). Hillsdale, NJ: Erlbaum.
- Pratt, John W. (1964), "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 122-136
- Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: divergent perceptions of bias in self versus others. *Psychological review*, 111(3), 781
- Quesnel-Vallée, A. (2007). Self-rated health: caught in the crossfire of the quest for 'true' health?. *International Journal of Epidemiology*, 36(6), 1161-1164.
- Reed, M. B., & Aspinwall, L. G. (1998). Self-affirmation reduces biased processing of health-risk information. *Motivation and Emotion*, 22(2), 99-132.
- Ritov, I., & Baron, J. (1992). Status-quo and omission biases. *Journal of risk and uncertainty*, 5(1), 49-61.
- Richman, W. L., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology*, 84(5), 754.
- Robertson, J. L., & Barling, J. (Eds.). (2014). *The Psychology of Green Organizations*. Oxford University Press. ( page 50)

- Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: evidence for an accessibility model of emotional self-report. *Psychological bulletin*, 128(6), 934.
- Rodin, J., & McAvay, G. (1992). Determinants of change in perceived health in a longitudinal study of older adults. *Journal of Gerontology*, 47(6), P373-P384
- Rohwedder, S., Haider, S. J., & Hurd, M. D. (2006). Increases in wealth among the elderly in the early 1990s: How much is due to survey design?. *Review of Income and Wealth*, 52(4), 509-524.
- Rosenberg, M. (1989). Self-concept research: A historical overview. *Social forces*, 68(1), 34-44.
- Rosenberg, M. (1981). The self-concept: Social product and social force. *Social psychology: Sociological perspectives*, 593-624.
- Rosenthal, R. (1987). "Pygmalion" Effects: Existence, Magnitude, and Social Importance. *Educational Researcher*, 37-41.
- Salomon, J. A., Tandon, A., & Murray, C. J. (2004). Comparability of self rated health: cross sectional multi-country survey using anchoring vignettes. *Bmj*, 328(7434), 258.
- Samuelson W, Zeckhauser RJ. Status quo bias in decision making. *J Risk Uncertainty*.1988; 1:7-59.
- Sarver, V. T., Jr. (1983). Ajzen and Fishbein's "theory of reasoned action": A critical assessment. *Journal for the Theory of Social Behavior*, 13, 155-163.
- Savage, L. J. (1961). *The Foundations of Statistics Reconsidered*. University of Calif Press
- Schaeffer, N. C., & Presser, S. (2003). The science of asking questions. *Annual Review of Sociology*, 65-88
- Schechter, S., & Herrmann, D. (1997). The proper use of self-report questions in effective measurement of health outcomes. *Evaluation & the health professions*, 20(1), 28-46.
- Schechter, S., Beatty, P., Willis, G. B., Schwarz, N., Park, D. C., Knauper, B., & Sudman, S. (1998). Asking survey respondents about health status: judgment and response issues. *Cognition, aging, and self-reports*, 265-283.
- Schechter, S. (1993). Investigation into the cognitive processes of answering self-assessed health status questions. US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.
- Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: assessment and implications of generalized outcome expectancies. *Health psychology*, 4(3), 219
- Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological science*, 18(5), 429-434
- Schwarz, N. (1999). Self-reports: how the questions shape the answers. *American psychologist*, 54(2), 93.
- Schwarz, N., Knäuper, B., Hippler, H. J., Noelle-Neumann, E., & Clark, L. (1991). Rating scales numeric values may change the meaning of scale labels. *Public Opinion Quarterly*, 55(4), 570-582.

- Schwarzer, R. (1994). Optimism, vulnerability, and self-beliefs as health-related cognitions: A systematic overview. *Psychology and health*, 9(3), 161-180.
- Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. *Measures in health psychology: A user's portfolio. Causal and control beliefs*, 1, 35-37.
- Schweitzer, M. (1994). Disentangling status quo and omission effects: An experimental analysis. *Organizational Behavior and Human Decision Processes*, 58(3), 457-476.
- Schopenhauer, A. (2012). *The world as will and representation* (Vol. 1). Courier Corporation.
- Segovia, J., Bartlett, R. F., & Edwards, A. C. (1989). An empirical analysis of the dimensions of health status measures. *Social science & medicine*, 29(6), 761-768.
- Sen A. (2002) "Health: Perception versus Observation", *British Medical Journal*, April 2002.
- Sen, A (1977). "Rational Fools: A Critique of the Behavioural Foundations of Economic Theory," *Philosophy and Public Affairs* 6 (Summer 1977);
- Shah, D., Kumar, V., & Kim, K. H. (2014). Managing Customer Profits: The Power of Habits. *Journal of Marketing Research*, 51(6), 726-741
- Sheng, M. L., Hsu, C. L., & Wu, C. C. (2011). The asymmetric effect of online social networking attribute-level performance. *Industrial Management & Data Systems*, 111(7), 1065-1086.
- Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, 99-118.
- Simon, H. A. (1965). *Administrative behavior* (Vol. 4). New York: Free Press.
- Simon, J. G., De Boer, J. B., Joung, I. M. A., Bosma, H., & Mackenbach, J. P. (2005). How is your health in general? A qualitative study on self-assessed health. *The European Journal of Public Health*, 15(2), 200-208.
- Singh-Manoux, A., Dugravot, A., Shipley, M. J., Ferrie, J. E., Martikainen, P., Goldberg, M., & Zins, M. (2007). The association between self-rated health and mortality in different socioeconomic groups in the GAZEL cohort study. *International Journal of Epidemiology*, 36(6), 1222-1228
- Singh-Manoux, A., Martikainen, P., Ferrie, J., Zins, M., Marmot, M., & Goldberg, M. (2006). What does self rated health measure? Results from the British Whitehall II and French Gazel cohort studies. *Journal of epidemiology and community health*, 60(4), 364-372.
- Skinner, E. A. (1996). A guide to constructs of control. *Journal of personality and social psychology*, 71(3), 549.
- Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). Rational actors or rational fools: Implications of the affect heuristic for behavioral economics. *The Journal of Socio-Economics*, 31(4), 329-342.
- Smith, Adam (1790) (Pg 18 – 20, 30) *The Theory of Moral Sentiments*
- Smith, T. W. (1979). Happiness: Time trends, seasonal variations, intersurvey differences, and other mysteries. *Social Psychology Quarterly*, 18-30.

- Smith, J.P., 1999. Health bodies and thick wallets: the dual relation between health and economic status. *Journal of Economic Perspectives* 13, 145–166
- Smith, A. M., Shelley, J. M., & Dennerstein, L. (1994). Self-rated health: biological continuum or social discontinuity?. *Social Science & Medicine*, 39(1), 77-83
- Smith, T. W. (2013). The Impact of Alternative Response Scales on Measuring Self-ratings of Health. *Survey Practice*, 1(1)
- Sonstroem, R. J., & Potts, S. A. (1996). Life adjustment correlates of physical self-concepts. *Medicine & Science in Sports & Exercise*.
- Spranca, M., Minsk, E., & Baron, J. (1991). Omission and commission in judgment and choice. *Journal of experimental social psychology*, 27(1), 76-105
- Stets, J. E., & Burke, P. J. (2000). Identity theory and social identity theory. *Social psychology quarterly*, 224-237.
- STENBACK, A. (1964). Physical health and physical disease as objective fact and subjective experience. *Archives of general psychiatry*, 11(3), 290-301.
- Sterman, J. D. (1989). Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management science*, 35(3), 321-339.
- Stock, S., Miranda, C., Evans, S., Plessis, S., Ridley, J., Yeh, S., & Chanoine, J. P. (2007). Healthy Buddies: a novel, peer-led health promotion program for the prevention of obesity and eating disorders in children in elementary school. *Pediatrics*, 120(4), e1059-e1068
- Sturgis, P., Thomas, R., Purdon, S., Bridgwood, A., & Dodd, T. (2001). Comparative review and assessment of key health state measures of the general population.
- Sutter, M., Kocher, M. G., Glätzle-Rützler, D., & Trautmann, S. T. (2013). Impatience and uncertainty: Experimental decisions predict adolescents' field behavior. *The American Economic Review*, 103(1), 510-531.
- Swann, W. B. (1984). Quest for accuracy in person perception: a matter of pragmatics. *Psychological review*, 91(4), 457.
- Symonds, P. M. (1924). On the Loss of Reliability in Ratings Due to Coarseness of the Scale. *Journal of Experimental Psychology*, 7(6), 456
- Tagore, R. (1958). *Our Universe*. Translated by Indu Dutt.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological bulletin*, 103(2), 193.
- Taylor, S. E., Lichtman, R. R., & Wood, J. V. (1984). Attributions, beliefs about control, and adjustment to breast cancer. *Journal of personality and social psychology*, 46(3), 489.
- Teppa, F., Thorp, S., & Bateman, H. (2015). Family, friends and framing: A cross-country study of subjective survival expectations.

Thaler RH, Sunstein CR. *Nudge: improving decisions about health, wealth, and happiness*. New Haven: Yale University Press; 2008

Thaler RH. (2009). Opting in versus opting out. *New York Times* [Internet]. 2009 Sep 26[cited 2013 Mar 25]. Available from:<http://www.nytimes.com/2009/09/27/business/economy/27view.html>

Thompson, S. C., Armstrong, W., & Thomas, C. (1998). Illusions of control, underestimations, and accuracy: a control heuristic explanation. *Psychological bulletin*, 123(2), 143

Tourangeau, R., Rasinski, K. A., Bradburn, N., & D'Andrade, R. (1989). Belief accessibility and context effects in attitude measurement. *Journal of Experimental Social Psychology*, 25(5), 401-421.

Tourangeau, R., K.A. Rasinski and N. Bradburn (1991), "Measuring happiness in surveys: a test of the subtraction hypothesis", *Public Opinion Quarterly*, 55, 255-266.

Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological bulletin*, 103(3), 299.

Tourangeau, R. (1999), "Context effects on answers to attitude questions", in M.G. Sirken, D.J. Herrmann, S. Schechter, N. Schwarz, J.M. Tanur and R. Tourangeau (eds.),

Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions the impact of data collection mode, question format, and question context. *Public opinion quarterly*, 60(2), 275-304.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological bulletin*, 133(5), 859

Tourangeau, R. (1984). Cognitive science and survey methods. *Cognitive aspects of survey methodology: Building a bridge between disciplines*, 73-100.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 1039-1061

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453-458.

Tversky, Amos, and Daniel Kahneman (1974), "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185, 1124-1131

van Doorn, C., 1999. A qualitative approach to studying health optimism, realism and pessimism. *Research on Aging* 21 (3), 440-457

Van Doorslaer, E., & Gerdtham, U. G. (2003). Does inequality in self-assessed health predict inequality in survival by income? Evidence from Swedish data. *Social science & medicine*, 57(9), 1621-1629.

Verplanken, B. (2006). Beyond frequency: Habit as mental construct. *British Journal of Social Psychology*, 45(3), 639-656

Viswanathan, M. (2005). *Measurement error and research design*. Sage.

Wagstaff, A. (2002). Poverty and health sector inequalities. *Bulletin of the world health organization*, 80(2), 97-105

- Ware Jr, J. E. (1999). Future directions in health status assessment. *JCOM*, 6(3)
- Weber, E. U., & Hsee, C. (1998). Cross-cultural differences in risk perception, but cross-cultural similarities in attitudes towards perceived risk. *Management Science*, 44(9), 1205-121
- Weems, G. H., Onwuegbuzie, A. J., Schreiber, J. B., & Eggers, S. J. (2003). Characteristics of respondents who respond differently to positively and negatively worded items on rating scales. *Assessment & Evaluation in Higher Education*, 28(6), 587-606
- Weng, L. J. (2004). Impact of the number of response categories and anchor labels on coefficient alpha and test-retest reliability. *Educational and Psychological Measurement*, 64(6), 956-972.
- Weinstein, N. D. (1983). Reducing unrealistic optimism about illness susceptibility. *Health psychology*, 2(1), 11.
- Weinstein, N. D. (1988). The precaution adoption process. *Health psychology*, 7(4), 355.
- Wilcox, V. L., Kasl, S. V., & Idler, E. L. (1996). Self-rated health and physical disability in elderly survivors of a major medical event. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 51(2), S96-S104.
- Wildt A.R. and Mazis M.B. (1978). 'Determinants of scale response: label versus position', *Journal of Marketing Research*, 15: 261-267.
- World Health Organization and Calouste Gulbenkian Foundation. *Social determinants of mental health*. Geneva, World Health Organization, 2014.
- Worsley, A. (1990). Laypersons' evaluation of health: an exploratory study of an Australian population. *Journal of epidemiology and community health*, 44(1), 7-11
- Wood, W., & Neal, D. T. (2007). A new look at habits and the habit-goal interface. *Psychological review*, 114(4), 843.
- Zaller, J., & Feldman, S. (1992). A simple theory of the survey response: Answering questions versus revealing preferences. *American journal of political science*, 579-616.
- Zamir, E. (2012). Loss Aversion and the Law. *Vand. L. Rev.*, 65, 829.
- Zeelenberg, M., & Beattie, J. (1997). Consequences of regret aversion 2: Additional evidence for effects of feedback on decision making. *Organizational Behavior and Human Decision Processes*, 72(1), 63-78
- Zimmer, Z., Natividad, J., Lin, H.-S., Chayovan, N., 2000. A cross-national examination of the determinants of self-assessed health. *Journal of Health and Social Behavior* 41, 465-481.

[https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/60524/403936\\_BehaviouralInsight\\_acc.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/60524/403936_BehaviouralInsight_acc.pdf)

[www.economist.com/node/13350892](http://www.economist.com/node/13350892)