University of Cape Town

School of Management Studies

THE ANCHORING AND ADJUSTMENT HEURISTIC IN UNSTRUCTURED INTERVIEWS: AN EXPERIMENTAL STUDY

CHELSEY ELLEN PIENAAR
(PNRCHE002)

A dissertation submitted in partial fulfilment of the requirements for the award of the Degree of Master of Commerce in Organisational Psychology

Faculty of Commerce

University of Cape Town

2014

COMPULSORY DECLARATION:

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works of other people has been attributed, cited and referenced.

Signature: Signed by candidate
Signature removed

Date: …16/05/2014…..

Acknowledgment

The author wishes to thank several people. I would like to thank my supervisor, François de Kock, for his incredible patience, support and guidance as I wrote this dissertation. I would also like to thank all of the staff in the Section of Organisational Psychology at the University of Cape Town for their assistance. Finally, I would like to acknowledge the National Research Foundation for the financial assistance that enabled me to study this year.

Abstract

The presence of anchoring and adjustment within structured interviews has been widely observed in research over the years. However, the unstructured interview is a more extensively used selection tool than the structured interview, making it important to understand anchoring and adjustment in the unstructured interview setting. The present study investigated the presence of the anchoring and adjustment heuristic in unstructured interviews using a between-subjects, post-test only experimental design. The sample consisted of 78 managers and human resource personnel from various organisations. Subjects were required to watch a short video of an unstructured interview, and then provide an overall rating of the candidate's communication skills. Subjects were divided into high, low, and control conditions. A high, low, or no anchor was provided in the question asking subjects to rate the video candidate. Results indicated that anchoring and adjustment did not occur in our simulated unstructured interviews. Limitations of the study and implications for future research are discussed.

Keywords: anchoring and adjustment, unstructured interviews, experiment, ratings

Table of Contents

List of Tables

List of Figures

The Anchoring and Adjustment Heuristic in Unstructured Interviews: An Experimental Study

The employment interview is the most commonly used and accepted selection tool in organisations across the world (Dipboye, Macan, & Shahani-Denning, 2012). Unstructured interviews are still employed widely, despite evidence that structured interviews are superior in validity and predictive accuracy (Dana, Dawes, & Peterson, 2013; Van der Zee, Bakker, & Bakker, 2002). For example, in a recent study of 959 companies from 20 different nations, it was found that only 34.7% of companies used structured interviews instead of unstructured interviews (Ryan, McFarland, Baron, & Page, 2004).

The unstructured interview frequently takes the form of a casual conversation where questions and model answers are not predetermined (Kataoka, Latham, & Whyte, 1997). However, extensive research has shown that interviews may be prone to subjective rating bias, especially when they are unstructured (Brtek & Motowidlo, 2002; Moore & Lee, 1974). In unstructured interviews there exist few normative constraints to guide the interviewer's evaluation and rating of the candidate. For this reason, interviewers may rely more on heuristics, distinctive mental operations that draw a relationship between occurrence and outcome (Moutier & Houde, 2003), to simplify complex information and make quicker judgements when faced with uncertainty (Jones & Welsh, 1971; Powell & Goffin, 2009). These mental shortcuts can have a biasing effect on candidate ratings. For example, research on *primacy*, (Steiner & Rain, 1989), *recency* (Farr & York, 1975) and *contrast effects* (Landy & Bates, 1973) within the recruitment and selection context have illustrated how judgement heuristics used by interviewers affect the judgement of candidates.

In the domain of judgement heuristics, *anchoring and adjustment* (Tversky & Kahneman, 1974) is receiving growing attention as a process of simplifying complex judgemental tasks in situations of uncertainty. This heuristic occurs when a decision must be made on the probability or value of events, when the true value is unknown or uncertain (Tversky & Kahneman, 1974). People start their estimate of an unknown value by fixing on a value that they do know (the anchor) and then adjusting until a satisfactory value is achieved (Gehlbach & Barge, 2012; Kataoka, et al., 1997). It is insufficient adjustment from the anchor that leads to incorrect value judgements.

The anchoring and adjustment heuristic has also been researched within structured interviews (Kataoka, et al., 1997). Using an experimental design and a sample of graduate students of business, Kataoka and colleagues provided a high anchor in one group, and a low anchor in another. Anchors were provided in the form of a numerical rating that differed in the low and high conditions; a rating of 1 and a rating of 5 respectively. Results revealed interviewers to rate the candidate's performance lower when provided with an anchor of 1, and higher when provided with an anchor of 5. This indicates that interviewers are susceptible to anchoring and adjustment effects when rating candidates in structured interviews. Candidate ratings skewed in the direction of the provided anchor, suggesting insufficient adjustment from the anchor. As such, the anchoring and adjustment heuristic may be important to consider in personnel selection applications, such as interviews, where uncertainty exists regarding the 'unknown' quantity, e.g., candidate competence or potential.

Anchoring and adjustment may be particularly prevalent in situations of uncertainty and ambiguity, such as the unstructured interview. In interviews, interviewers are uncertain about the characteristics (e.g., level of competence) of a particular candidate. This uncertainty leads

interviewers to use heuristics and mental shortcuts in order to make a judgement of a candidate and award them a rating on some competency (Simmons, LeBeouf, & Nelson, 2010). As such, the anchoring and adjustment heuristic may be utilised, and would thus be expected to operate specifically in unstructured interviews. While anchoring and adjustment has been demonstrated in high structure interviews, it has not yet been investigated whether the heuristic would also exist in low structure interviews.

The importance of understanding the presence and effects of anchoring and adjustment in unstructured interviews lies in the fact that they are far more widely used than structured interviews (Dana, et al., 2013; Ryan, et al., 2004). The impact of anchoring and adjustment on selection decisions in organisations globally is thus larger and more prevalent than that of structured interviews. Common use of the unstructured interview makes organisations vulnerable to anchoring and (insufficient) adjustment, which can have important implications for the organisation. When anchoring and (insufficient) adjustment occurs, inaccurate judgements are made and potentially harmful decisions affecting effort and performance become consequential for the organisation (Gehlbag & Barge, 2012; Rutledge, 1993). For example, during unstructured interviews, managers and human resource personnel may anchor on a particular interview instruction and insufficiently adjust their rating from that anchor, leading them to under or over estimate the performance or ability of candidates. This would result in the recruitment of new employees who are poor performers, at the expense of potentially high-performing employees. By attempting to better understand anchoring and adjustment, the prevalence of the heuristic in unstructured interviews can be reduced through training and clear interview guidelines (Furnham & Boo, 2011), leading to better quality selection decisions.

Understanding the anchoring and adjustment heuristic is thus an important issue that needs to be understood. The present study will attempt to address this issue by investigating whether or not the anchoring and adjustment heuristic occurs in the unstructured interview setting. Thus, the research question is "*What is the effect of the anchoring and adjustment heuristic on unstructured interview ratings?*". To this end we will test for the presence of anchoring and adjustment in unstructured interviews. In a highly controlled setting, we will show subjects controlled performance information in the form of a video interview, and provide low, high and no anchors in order to determine the presence of the anchoring and adjustment heuristic.

The Anchoring and Adjustment Heuristic

Anchoring and adjustment is a particularly ubiquitous heuristic that demonstrates the relativity of human judgement. This heuristic occurs when a decision must be made on the probability or value of events when the true value is unknown or uncertain, such as a judgement of comparisons (Gehlbach & Barge, 2012; Kataoka, et al., 1997) (Figure 1 is a visual representation of the process of anchoring and adjustment). A comparative nature is one of the most remarkable features of human judgement (Mussweiler & Englich, 2005). People will use a comparison to some context, norm or standard when judging a specific object or act. From the basic estimate of the distance of a river, to the more complex, intangible attributes of the self, judgement essentially relies on comparison. People start their estimate of an unknown value by fixing on a value that they do know (the anchor) and then adjusting until a satisfactory value is achieved. Thus, anchoring and (insufficient) adjustment can be defined as fixing on an irrelevant value that is easily accessible (anchoring) and that influences the process of adjusting, to arrive at

a final estimate of an unknown value (adjustment). It is insufficient adjustment from the anchor that leads to incorrect value judgements (Epley & Gilovich, 2001).

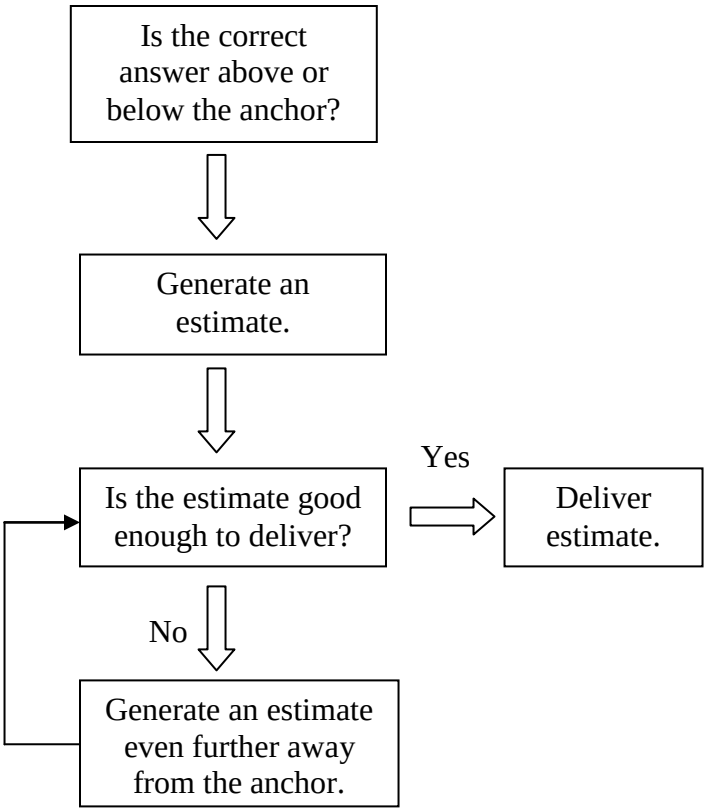


Figure 1: Graphic representation of the anchoring and adjustment paradigm by Simmons, et al. (2010).

Figure 1 depicts the process by which raters settle on an estimate, considering it to be close enough to the correct value. Without investing sufficient time and effort into adjustment, they tend to under or overestimate the correct value. The original anchor tends to exert undue influence on subsequent adjustments, resulting in a final estimate that is inaccurately close to the anchor value (Epley & Gilovich, 2006). In this way insufficient adjustment occurs and leads to bias in judgement and thus inaccurate decisions.

In a typical demonstration of anchoring and adjustment in its purest form, participants are asked to estimate some unknown value (e.g., the number of U.S. states in 1880) after being provided with an anchor (e.g., 13 – the number of original colonies; or 50 – the current number of states). Because the anchor is not the correct value, respondents will adjust their estimates up or down from the provided anchor. However, adjustments tend to be insufficient, with respondents resolving on final estimates that are conservatively close to the original anchor (Tamir & Mitchell, 2013). In the U.S. states example, respondents were found to underestimate the correct number of states when provided with an anchor of 13, and overestimate the correct number of states when provided with an anchor of 50. There are a number of underlying mechanisms that serve to bring about this effect.

Underlying Mechanisms of the Anchoring and Adjustment Heuristic

In order to determine the consequences of anchoring and adjustment in specific settings, such as the unstructured interview, the mechanisms that contribute to the anchoring and adjustment heuristic must be clearly understood. These include the dual-processing perspective, accessibility, and confirmatory information search.

First, the *dual-processing perspective* of human judgement posits that there are two independent but parallel systems for processing information (Allen, 2011). System 1 is a more primitive form of information-processing that is performed unconsciously, automatically, and requires little cognitive effort. Alternatively, System 2 is a more evolved, analytical cognitive process that requires conscious effort. System 1 has been found to be linked to heuristics. These are rapid, automatic processes of reducing the complexity of tasks in order to predict or judge unknown values (Furnham & Boo, 2011). This type of mental shortcut can often result in biases or errors in judgement. Anchoring and adjustment is a heuristic related to system 1 information-

processing. This means that it takes mental effort to work against the automaticity of anchoring on a given value, and to adjust sufficiently from that anchor in order to arrive at a more accurate judgement (Allen, 2011).

Accessibility is the second mechanism contributing to anchoring and adjustment. Zaller (1992) argues that opinion formation is deeply affected by the most recently occurring events. More plainly, those events that have been seen, heard, or even indirectly implied most recently will be the most *accessible* and easily recalled when making judgements. In the U.S. states example, the provided anchor became the most accessible reference for respondents, and was thus recalled and reused every time they estimated the number of U.S. states in 1880; a question to which the correct answer was unknown.

Finally, *confirmatory information search* refers to the tendency to overestimate the value of information that is perceived to support one's estimate or viewpoint (Fischer, Fischer, Englich, Aydin, & Frey, 2011). This mechanism can obstruct the revision of incorrect initial judgements, resulting in overconfidence in the anchor value, and thus insufficient adjustment (Fischer, Lea, Kastenmuller, Guitemeyer, Fischer, & Frey, 2011).

The three underlying mechanisms, the dual-processing perspective, accessibility, and confirmatory information search, work concurrently to produce the anchoring and (insufficient adjustment) heuristic, which can be observed in various, every-day decision-making events.

Anchoring and Adjustment in "Real Life"

Anchoring and adjustment has been demonstrated within a wide variety of practical decision-making contexts, including consumer buying behaviour (Epley & Gilovich, 2009; Naylor, Lamberton, & Norton, 2011), property pricing (Northcraft & Neale, 1987), auditing decisions (Joyce & Biddle, 1981), and voting behaviour (Zuckerman, Kotler-Berkowiz, &

Swaine, 1998). Anchoring and adjustment is evident in consumer buying behaviour, where consumers have been willing to pay more at Bistro 97 than at Bistro 17 (Epley & Gilovich, 2009). Property pricing is also susceptible to anchoring and adjustment, with property buyers being unduly influenced by the initial asking or listing price. The ensuing negotiations are insufficiently adjusted from the anchor – the listing price (Northcraft & Neale, 1987). Other than the evident prevalence of anchoring and adjustment in practical decision-making contexts, the heuristic as also been frequently observed within the organisational setting.

Anchoring and Adjustment in the Organisational Context

The heuristic has been well-established within the organisational context, being frequently observed in literature on performance appraisals (Heslin, Latham, & Vande Walle, 2005), survey responses (Gehlbag & Barge, 2012) and group decisions (Rutledge, 1993). In their study on implicit attitudes during performance appraisals, Heslin, et al. (2005) found that raters anchored on initial impressions, and thus provided lower ratings of good performance when initial performance was poor. In 2012, Gehlbag and Barge discovered that anchoring and adjustment influences respondents' scores on psychometric scales. Respondents were found to anchor on the first item relating to a concept, and then adjust their attitudes and scores on subsequent items of the same concept. Rutledge (1993) found evidence for the anchoring and adjustment heuristic in group decision-making, similar to individual decision-making. However, group members were found to adjust further away from the anchor than individuals who had not participated in group discussions. This further adjustment could perhaps be attributed to the exposure to new information that is brought about by discussions with others. This in turn may have increased thought effort, thereby widening the anchor-estimate gap.

Awareness of group effects on anchoring and adjustment is important in the organisational context, as it encourages the use of group discussions in uncertain decision-making situations when more accurate judgements are desired. An example would be to use an interview panel rather than a single interviewer when recruiting.

Anchoring and adjustment is thus clearly a prevalent heuristic within organisations. However, its prevalence and intensity is not always constant, but rather is dependent upon certain limiting factors called boundary conditions.

Boundary Conditions

Evidence for anchoring and adjustment in organisational decision-making processes has led researchers to seek out those conditions that may serve to limit the prevalence of this heuristic. After all, when anchoring and (insufficient) adjustment occurs, inaccurate judgements are made and potentially harmful decisions affecting effort and performance become consequential for the organisation (Gehlbag & Barge, 2012; Rutledge, 1993). However, the anchoring and adjustment heuristic has proven to be exceptionally robust (Mussweiler & Englich, 2005). Anchoring and adjustment persists independently of the relevance of the anchor to the estimated value. Even when the anchor is as random as the throwing of a dice, judgement is influenced. Furthermore, extreme anchors affect the judgements of values, despite their implausibility (Mussweiler & Englich, 2005). For example, when provided with the anchor of 140 for Mahatma Ghandi's age, respondents significantly overestimated the true age, despite the improbability of the anchor value (Mussweiler & Strack, 2000). However robust and ubiquitous, there are factors that can serve to greaten or lessen the effects of anchoring and adjustment.

Over the past five years, research has revealed certain settings and conditions that influence the existence and extremity of anchoring and adjustment. A study on the effects of time

on anchoring and adjustment found that, because adjusting from a given anchor requires effort, increasing the amount of time provided to estimate an unknown value yields estimates further away from the anchor (Shafir & LeBoeuf, 2009). Respondents distanced their estimates further and further from the anchor as time increased (hours to days to weeks) because they had more time to invest effort in the adjustment process. This is evidence that increasing response time decreases the effects of anchoring and (insufficient) adjustment.

Furthermore, studies on accuracy motivation have found that incentives, such as monetary rewards, can widen the anchor-estimate gap (Simmons, et al., 2010). When respondents are offered incentives in exchange for accurate estimates, they will concentrate more effort into sufficiently adjusting from the provided anchor, thereby decreasing the effects of anchoring and (insufficient) adjustment.

Ambiguity and uncertainty are also well-known to encourage anchoring and adjustment. Findings from a recent study on consumer behaviour (Naylor et al., 2011) further support the relationship between uncertainty and the anchoring and adjustment heuristic. Findings revealed that priming participants to access other related thoughts through external cues resulted in further adjustment from the provided anchor. Similarly, Tamir and Mitchell (2013) discovered that silent cues, such as shaking the head in disagreement or nodding the head in agreement, were related to anchoring and adjustment. When a respondent's estimate is met with a shaking head, they will continue to correct and adjust their estimate. However, a nod of the head was found to cause respondents to cease adjustment and settle on a final estimate. The results of the last two studies indicate that external cues can either broaden or narrow the anchor-estimate gap through manipulation of respondents' thought processes.

Because of the relationship between uncertainty and anchoring and adjustment, it might be suggested that interview technique would influence effects of anchoring and (insufficient) adjustment. It seems logical to propose that the more structured the interview, and thus the less uncertainty, the less anchoring and (insufficient) adjustment would occur.

Anchoring and Adjustment in Interview Ratings

Uncertainty and Interview Structure

Expansive research has been dedicated towards finding the most effective interview techniques (Kataoka, et al., 1997; Lin, et al., 1992). Typically in this research, the structured interview has superior reliability and validity to the unstructured interview (McDaniel, Whetzel, Schmidt, & Maurer, 1994). A structured interview is defined as a standardised method of questioning that guides the interviewer in determining what questions to ask and how to interpret interviewee responses (Van der Zee, et al., 2002). The increased structure and standardisation of interviews can make them less susceptible to heuristics, such as anchoring and adjustment. This can lead to better selection decisions and consequently more productive employees and improved organisational profitability. This is an advantage over the unstructured interview, which, according to Cronshaw and Weisner (1989), is so unsound an assessment technique that no further research should be invested in them.

However, despite the convincing evidence against the use of the unstructured interview, it is still the dominant selection tool among organisations today (Blackman, 2002). There are a number of possible reasons for the commonality of the unstructured interview. Firstly, employers may be unaware of the vast amount of research promoting the use of the structured interview, and thus the potential benefits of the structured interview to the organisation. Second, the

unstructured interview may be more intuitively appealing to interviewers. Unstructured interviews allow more autonomy, power and challenge, and are often less monotonous than the standardised procedures of the structured interview (Van der Zee, et al., 2002). Finally, structured interviews may not be used because they conflict with the general ethos of the organisation. When an organisation has limited resources, the structured interview is often seen as senselessly costly and time-consuming as compared to the unstructured interview (Van der Zee, et al., 2002). Thus, despite their poor psychometric properties, unstructured interviews still hold much importance as a selection tool, and their effectiveness as such should continue to be researched as long as they are being widely used.

Anchoring and Adjustment in the Unstructured Interview

Interviewers are raters in an uncertain decision-making situation where the candidate's competence is objectively indeterminable, and thus they too will be susceptible to anchoring on irrelevant information. Research on anchoring and adjustment suggests that raters may be susceptible to influence from irrelevant information that serves as an anchor in forming judgements (Chapman & Johnson, 2002). Comprehensive revision of the interview literature suggests that future research directed towards analysing the cognitive processes involved in different types of interviews is necessary for the progression of effective selection decisions (Harris, 1989). By understanding the value of different interview techniques, organisations will be able to employ the most valid and reliable of these in selection procedures.

Anchoring and adjustment may occur in low structure interviews for various reasons. First, the heuristic has already been established as being rifer in situations of uncertainty. The structure of the interview influences how much uncertainty exists for the interviewer during the interview process. By providing more standardisation and guidance on how to interpret

interviewee responses, uncertainty is reduced and anchoring and (insufficient) adjustment is less likely to impact on the interviewer's judgement and decisions. Thus, the unstructured interview is a situation of uncertainty in which anchoring and adjustment can be expected to operate.

Second, anchoring and (insufficient) adjustment is an automatic, unconscious heuristic that requires effort to work against (Allen, 2011). When interviewers find themselves in situations of uncertainty, such as the unstructured interview, and have little time or incentive to exercise cognitive effort, they may default to an unconscious, automatic cognitive process and experience anchoring and (insufficient) adjustment. Anchoring and adjustment is thus a likely heuristic to be utilised by raters during an unstructured interview.

In summary, it is not being disputed that anchoring and adjustment does occur in structured interviews. However, when there is more structure in the interview, interviewers may be better guided in their unconscious thought processes. They would have better direction as to what aspects to look for in judging the quality of the interviewee, and what responses constitute desirable. Thus, because anchoring and adjustment was found in high structure interviews, we also expect to find the anchoring and adjustment in unstructured interviews where conditions are more conducive for the heuristic.

The Present Study

The present study will seek to determine the presence of the anchoring and adjustment heuristic in unstructured interviews. The importance of this study is in gaining an improved understanding of the heuristics that weaken the accuracy of our judgements. The growing complexity of information has resulted in mental shortcuts that deteriorate our ability to make accurate and effective judgements (Allen, 2011). It is important that these shortcuts are well understood so that we are able to reduce their effects, and thus the biases and erroneous decisions

that are detrimental to the organisation. The present study will aim to understand one specific such mental shortcut, the anchoring and adjustment heuristic.

Prior research has demonstrated the presence of anchoring and (insufficient) adjustment in different types of structured interviews (Kataoka, et al., 1997). Respondents were provided with written interview transcripts and were asked to rate the interviewee's responses to each question on a rating scale. The interview transcripts were provided randomly in three different interview techniques. Before the respondent began rating the interviewee, they were asked whether the applicant's rating deserved a particular score (e.g. score of 5).

Following a similar design to Kataoka et al. (1997), the presence of anchoring and adjustment will be evaluated in unstructured interviews. Drawing on social cognition theory, we argue that when interviewers are provided with a low verbal anchor ("very poor"), it would bias subject ratings, leading them to insufficiently adjust and provide lower ratings. On the other hand, when a high verbal anchor ("excellent") is provided, subjects' ratings will be positively biased, leading them to assign higher ratings to the candidate. Although no work in the recruitment and selection literature is available to propose the following hypotheses exactly, the social psychology literature is instructive:

H_1: Interviewer ratings of identical responses to interview questions will be significantly more favourable when interviewers are provided with a high anchor rather than no anchor.

H_2: Interviewer ratings of identical responses to interview questions will be significantly less favourable when interviewers are provided with a low anchor rather than no anchor.

Method

Research Approach

To determine the effect of the anchoring and adjustment heuristic on interview judgements, a between-subjects design was used (Millsap & Maydeu-Olivares, 2009). The research approach was quantitative with a post-test only control group experimental design. In this design, two treatment groups and a control group are used with all groups completing the treatment only once after the intervention (Highhouse, 2009). Experimental research is defined as the objective, systematic investigation into cause-and-effect through the manipulation of an independent variable in a strictly controlled environment (Highhouse, 2009). The current study followed the max-min-con principle (Kerlinger, 1986) in order to allow the isolation of the anchoring and adjustment heuristic as the sole explanation for differences in ratings.

Subjects

Non-probability sampling was used, including purposive sampling, followed by additional snowball sampling, in order to maximise sample size. A sufficiently sized sample was sourced from the broader population of managers and human resource (HR) personnel across various industries. Because interviews are mostly conducted by managers and HR personnel, the sampling frame consisted of a list of all managers and HR personnel in various organisations. A required minimum sample size of 78 subjects was determined a priori to achieve sufficient statistical power (.80) (Cohen, 1988; Cohen, 1992); using the statistical software G*Power 3 (Cunningham & McCrum-Gardner, 2007). Statistical power of .80 is likely to lead to results that correctly detect whether the hypothesis is true or not (Cohen, 1992).

Senior management or HR personnel in various organisations, across various industries were contacted via email or phone to obtain permission for participation of company employees. Once written or verbal permission was granted, an email including the survey link and study information was sent to the company HR manager, who then distributed the survey to all company management and HR personnel.

Responses (N = 77) were recorded and, because the online survey captured only complete responses, each participant completed 100% of the items in the survey. Subjects' ages in this sample ranged from 21 to 62 (M = 38, SD = 11) years of age. The sample was marginally skewed towards females in terms of gender, with 54 females (69%) and 24 males (31%). The majority of subjects were white (80%) and worked in a middle or senior level management position (41%). Subjects were relatively evenly spread across industry, with the largest percentage of subjects indicating jobs as managers and HR personnel in the accounting, banking and financial sectors (15%). The larger majority of subjects indicated that they hold a degree or diploma (81%).

Subjects were randomly assigned to conditions (high anchor: n = 26; low anchor: n = 25; control anchor: n = 25), thereby splitting them into three roughly equal groups. The demographic information of subjects in each condition can be found in Table 1 below. The respective conditions did not differ significantly in terms of gender (χ^2 = 1.083, p = .582, n.s.), race (χ^2 = 3.589, p = .732, n.s.) and education level (χ^2 = 8.452, p = .585, n.s.). Therefore, random assignment effectively controlled for possible confounding effects of gender, race and education level.

Table 1

Descriptive Statistics of Subjects in Each Condition (*N = 78*)

Socio-demographic		High Condition (*n* = 26)		Low Condition (*n* = 25)		Control condition (*n* = 25)	
		f	%	*f*	%	*f*	%
Gender	Female	19	73	19	76	16	64
	Male	7	27	7	24	10	36
Race	Black	1	4	1	4	0	0
	White	22	84	21	84	19	76
	Coloured	2	8	3	8	6	20
	Prefer not to answer	1	4	1	4	1	4
Education	Matric	2	8	6	24	7	28
	Diploma	4	15	5	20	5	20
	Bachelor's Degree	7	27	6	24	6	24
	Postgraduate Degree (Honours)	8	31	5	20	6	24
	Postgraduate Degree (Masters)	3	11	4	12	2	4
	Postgraduate Degree (PhD)	2	8	0	0	0	0

Video Stimulus

We used a single video-taped interview as the stimulus, enabling the control of performance and other extraneous influences. The five minute videotape of an unstructured interview was chosen from a range of eight videos in an earlier unpublished study

(Hauptfleisch, 2012). In the video, the interviewee is seated at a desk in front of a camera, facing the interviewer. Only the candidate is visible; while the interviewer is hidden to control for extraneous variance due to demographic characteristics of the interviewer (McCarthy, Iddekinge, & Campion, 2010). Interviewer questions are displayed on the screen in the form of text, while audio is blanked out (see Appendix A). This is a precaution employed to prevent the influence of the interviewer's speech or voice from affecting ratings (Highhouse, 2009). Example questions from the video are *"Tell me about yourself and your background"* and *"How do you usually handle interpersonal conflict situations"*. These are typical questions used in unstructured interviews in practice. See Appendix B for the complete list of interview questions.

We chose the video stimulus by inspecting the ratings (means and standard deviations) of 196 subjects across all eight interview videotapes in the earlier unpublished study (Hauptfleisch, 2012). From this, a single videotape of the most regular interviewee with the most average rating (M = 51.94) and lowest standard deviation (SD = 12.93) was chosen as the stimulus for the present study. This candidate was assumed to be the most consistently average candidate in terms of communication competence. For this reason, this candidate would allow anchoring and adjustment to be observed in both a high and low anchor direction.

Interview Measure

Subjects had to complete an unstructured interview rating sheet (see Appendix C). After viewing the video of the unstructured interview, subjects rated the candidate on a single dimension, *communication*. Communication is among the most commonly used competencies by managers in interviews (CIPD, 2008).

In the present study, the rating indicated a single, overall rating of the candidate's communication competence. The question asking subjects to rate the candidate's communication

was phrased differently depending on the condition (high, low or control) that they were assigned to. Ratings were provided on a 5-point Likert-type scale ranging from *1 = Very Poor* to *5 = Excellent.* In the study by Kataoka, et al. (1997), an average rating of many rating items was used in the measurement of the anchoring and adjustment effect. However, an average rating does not control for fluctuations in the effect of anchoring and adjustment across items, and may thus threaten internal validity. For this reason, we used a single, overall rating.

Procedure and Data Collection

Subjects were contacted via an email which was distributed to managers and HR personnel in various organisations. The email included a brief explanation of the study, the researcher's contact details, and informed subjects of the voluntary nature of the study. It was reiterated that participation was confidential, and that the study had been approved by the UCT Commerce Faculty Research in Ethics Committee. An online survey link was provided to direct them to Qualtrics.

Information on the purpose and nature of the study was provided on the opening page of the survey. Subjects were then asked to watch a video of an unstructured interview. The proceeding question asking the subject to rate the interview candidate's communication skills was phrased differently depending on whether a high, low, or no anchor was used. Qualtrics automatically randomly assigned subjects into these high, low, and control conditions. In other words, every participant had an equal probability of being assigned to the high, low or control condition. This served to enhance the internal validity of the results.

Finally, subjects were asked a series of questions pertaining to their demographic and occupational information. Before thanking subjects for their participation in the study, subjects were provided an opportunity to supply their email address for entrance into a lucky draw to win

a R500 spa treatment voucher.

Data was collected online via the Qualtrics online survey platform, from which the data was downloaded directly into the statistical programme SPSS version 20.

Treatment Conditions

Two treatment conditions were employed, namely a high anchor and low anchor condition. Similar to earlier studies (such as Kataoka, et al., 1997), a control condition was included in order to demonstrate whether the anchoring and adjustment heuristic had been successfully induced or not.

High-anchor condition. *Would you say that this candidate has excellent communication skills?* The word "excellent" was used as the high verbal anchor because of its positive connotation. The subject was then asked to rate the candidate's communication skills on a scale from 1 (very poor) to 5 (excellent).

Low-anchor condition. *Would you say that this candidate has very poor communication skills?* The word "poor" was used as the low verbal anchor because of its negative connotation. The subject was then asked to rate the candidate's communication skills on a scale from 1 (very poor) to 5 (excellent).

Control condition. In the control condition, subjects were asked to respond to "*How would you rate this candidate's communication skills?"* A control group was included to demonstrate the presence of the anchoring and adjustment heuristic and isolate its effect on ratings. In the control condition subjects were simply asked to rate the candidate's communication skills on a rating scale ranging from 1 (very poor) to 5 (excellent). No anchor was provided.

Anchors. We decided to use verbal anchors (high anchor = excellent; low anchor = poor), rather than numeric anchors. These were externally provided. Kataoka, et al. (1997) used a numeric anchor (rating from 1 to 5 on a Likert-type rating scale) as the anchor in their study. However, it can be argued that the use of a numeric anchor in the context of interviews is unrealistic. When people judge people, the most natural thought would be to assess them by some overall quality, or potential, rather than a number. Likewise, when interviewers judge the quality of candidates' responses, they are more likely to evaluate them in terms of high suitability or low suitability for the job. It has been found that highly applicable anchors have a greater effect on judgements of performance (Thornstein, Breier, Atwell, Hamilton, & Privette, 2008). Due to the intrinsic implication of performance judgement in interviews, the present study will attempt to elicit anchoring and adjustment in a realistic manner by using a more applicable suitability judgement (as opposed to a numeric rating scale) as the anchor. Verbal anchors are more realistic given the informal rating guidelines often used in unstructured interviews (Powell & Goffin, 2009; Van Der Zee, et al., 2002), and are expected to improve the external validity of the study.

In our study, verbal anchors were evident in the phrasing of the question that asked subjects to rate the candidate on his communication competence. This question was phrased differently according to the condition.

Experimental Control Measures

We employed a number of measures to enhance the internal and external validity of our study. First, the sample was preselected from a managerial sampling frame in order to collect a sample relevant to the research topic. Second, subjects were all given exactly the same detailed instructions prior to viewing the video, including information on the nature of the study, and

recommendations to find a quiet environment and ensure that they had access to sound. They were also informed that the candidate was being interviewed for a junior management position in an advertising firm. The junior management position was chosen due to the uncomplicated, generic nature of the job requirements. Prior familiarity with the job was thus unnecessary and prior knowledge of the job characteristics was unlikely to have biased results. All subjects watched the same video via Qualtrics, dismissing the possibility of differences in ratings being due to differences in candidate characteristics or features of the video. Next, all responses with a duration of less than three minutes were disregarded as careless responses that could negatively affect the results. Finally, to control for the possibility that previous interview experience of subjects could influence ratings, subjects were asked whether they had conducted an interview before or not. This information was used later on in the statistical analyses.

Statistical Analyses

In order to test for differences in the means of high-control and low-control conditions, two independent-samples t-tests were used (one for each hypothesis). The hypotheses were one-tailed and directional. Therefore, the t-test was used rather than using a non-directional test of differences, such as analysis of variance (ANOVA). All assumptions and hypotheses were tested using a 5% level of significance.

Results

Test of Assumptions

The six assumptions of the independent samples t-test (Howell, 2011) were investigated prior to analysing the data. These assumptions indicate whether the data is appropriate for the independent samples t-test or not.

Measurement level. The dependent variable must be measured on a continuous or interval scale (Howell, 2011). The dependent variable, rating, is measured on a 5-point Likert-type scale. This is a type of continuous scale, more specifically, an interval scale. This assumption was thus met.

Independent categories. The independent variable must consist of two independent categories or groups (Howell, 2011). In order to test the hypotheses specific to this study, the data was organised into high vs. control condition groups, and low vs. control condition groups, with the intention of a separate t-test being performed for each of the two hypotheses. Thus, the independent variable for each t-test consisted of only two groups. In addition, each of the high, low and control conditions was completely separate. Thus, the independent variable consisted of two independent groups, and assumption two was thus met.

Independence of observations. This assumption specifies independence of observations, meaning that there must be no relationship between observations or between groups (Howell, 2011). Random assignment of subjects into each group (condition) ensured that each of the high, low and control groups consisted of an independent sample. Assumption three was met.

Outliers. There must be no significant outliers, because these could negatively influence the validity of results (Aguinis, Gottfredson, & Joo, 2013). Outliers were identified by way of

box plots and Z-scores. Through these methods, two significant outliers were found with Z-scores exceeding an absolute value of 2.57 ($p = .01$) (Green & Salkind, 2008). These outliers were subjects who provided high ratings of the candidate's communication competence, despite having been assigned to the low condition, and thus contrary to the expected direction. While the deletion of outliers is not taken lightly, no pattern or plausible explanation could be found for the two outliers in question. They appeared to be spurious, and for this reason both of these outliers were deleted (Jacobs, 2001). Thus assumption four was met.

Homogeneity of variances. There must be homogeneity of variances (Howell, 2011). This assumption was tested using Levene's test for homogeneity of variances. Variances were assumed to be equal for the high vs. control comparison ($F = .06$, $p = .81$, n.s.) and the low vs. control comparison ($F = 2.33$, $p = .13$, n.s.). Therefore, homogeneity of variances was assumed and assumption five was met.

Normality. The dependent variable must be approximately normally distributed for each group of the independent variable (Howell, 2011). The Shapiro-Wilk test of normality was used to investigate this assumption. This test was chosen above that of Kolmogorov-Smirnov because it was more appropriate for the small size of the sample (Guo, 2012; Zimmerman, 2003).

None of the independent groups were found to be approximately normally distributed: high condition ($W = .63$, $p < .01$), low condition ($W = .83$, $p < .01$), control condition ($W = .63$, $p < .01$). Although the t-test is relatively robust to violations of normality (Green & Salkind, 2008), the results of the Shapiro-Wilks test indicated non-trivial violation of normality. For this reason, assumption six was violated.

The violation of one of the six assumptions reveals the data to be inappropriate for an independent samples t-test. An equivalent non-parametric test was thus more appropriate for the data. The most appropriate non-parametric test for this sample was the Mann-Whitney U test.

Four assumptions had to be investigated before the Mann-Whitney U test could be performed. The first three of these assumptions are the same as those of the independent samples t-test. All three of these assumptions were met. However, the Mann-Whitney U test does not require normally distributed data (Howell, 2011). Rather, the data in each group of the independent variable must resemble a similar shape. By analysing histograms for each of the high, low and control groups, they were found to have a similar shape of distribution. Thus, the fourth assumption of the Mann-Whitney U test was met.

Descriptive Statistics

The means and standard deviations for each condition are represented in Table 2 and Figure 2 below.

Table 2

Means and Standard Deviations for Each Condition

Condition	*n*	*M*	*SD*
High	26	2.58	.50
Low	25	2.64	.76
Control	25	2.44	.51
Totals	76	2.55	.60

The mean scores, based on the 5-point Likert scale with a midpoint of 3 used in the survey indicated that the respondents perceived the interview candidate to have relatively

average communication competence, regardless of the external anchor provided. A moderate distribution of ratings around the mean was indicated by the standard deviations and by the error bars displayed in Figure 2.

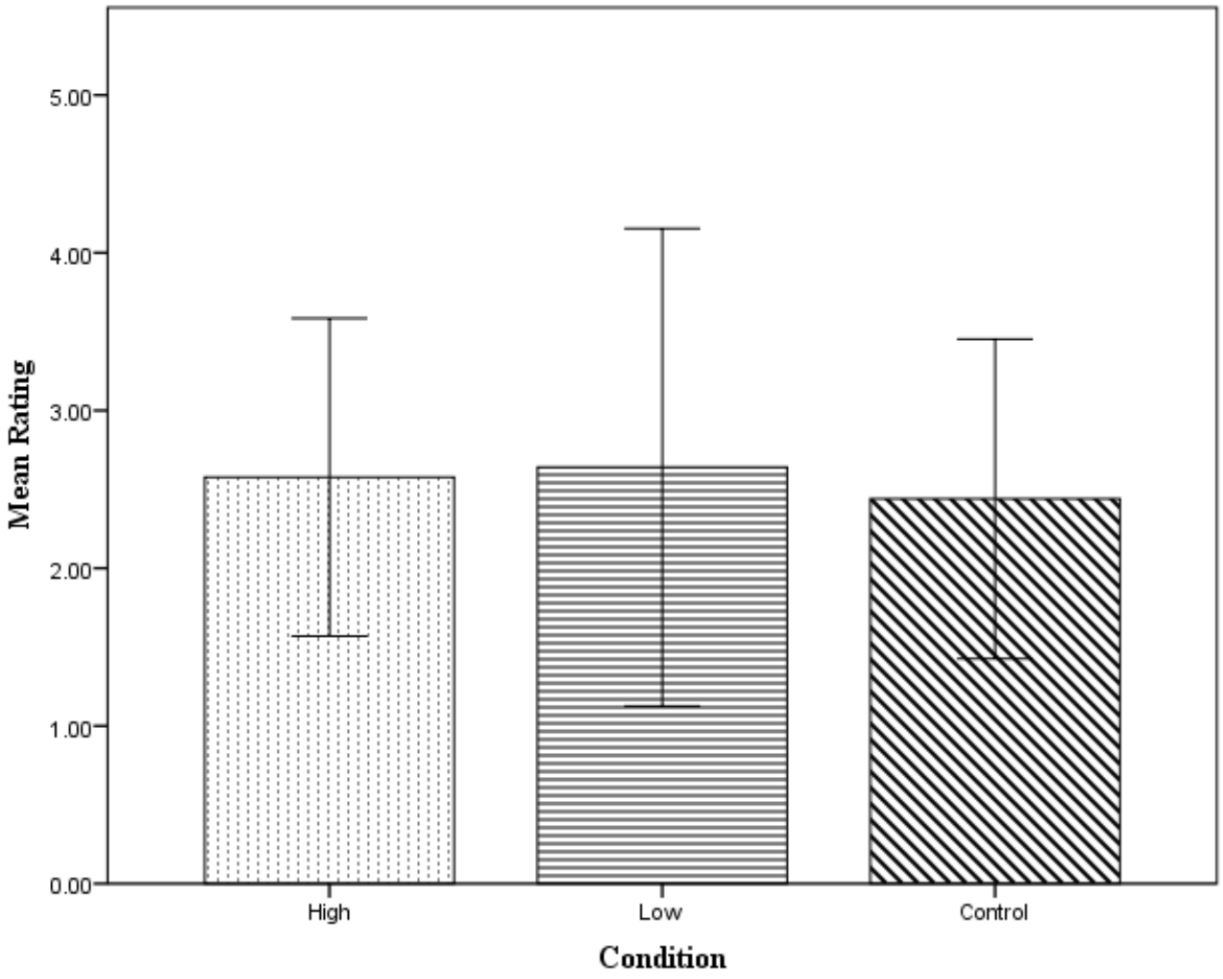


Figure 2: Graphic representation of mean ratings and error bars across the three experimental anchor conditions (high, low, control).

Possible moderating variables. Mean ratings across race, gender and educational level categories were calculated in order to identify any significant differences in ratings based on demographics. The mean ratings for White (M = 2.5), Black (M = 3.0), and Coloured (M = 2.64) subjects were observably similar. Subjects also did not differ significantly in terms of race, with

mean ratings of males (M = 2.50) and females (M = 2.58) being very similar. Finally, educational level did not seem to influence ratings, as mean ratings across all educational levels were calculated as approximately 2.5. Therefore, ratings were unaffected by demographic characteristics of the subjects. Mean ratings across educational levels are displayed in Figure 3 below.

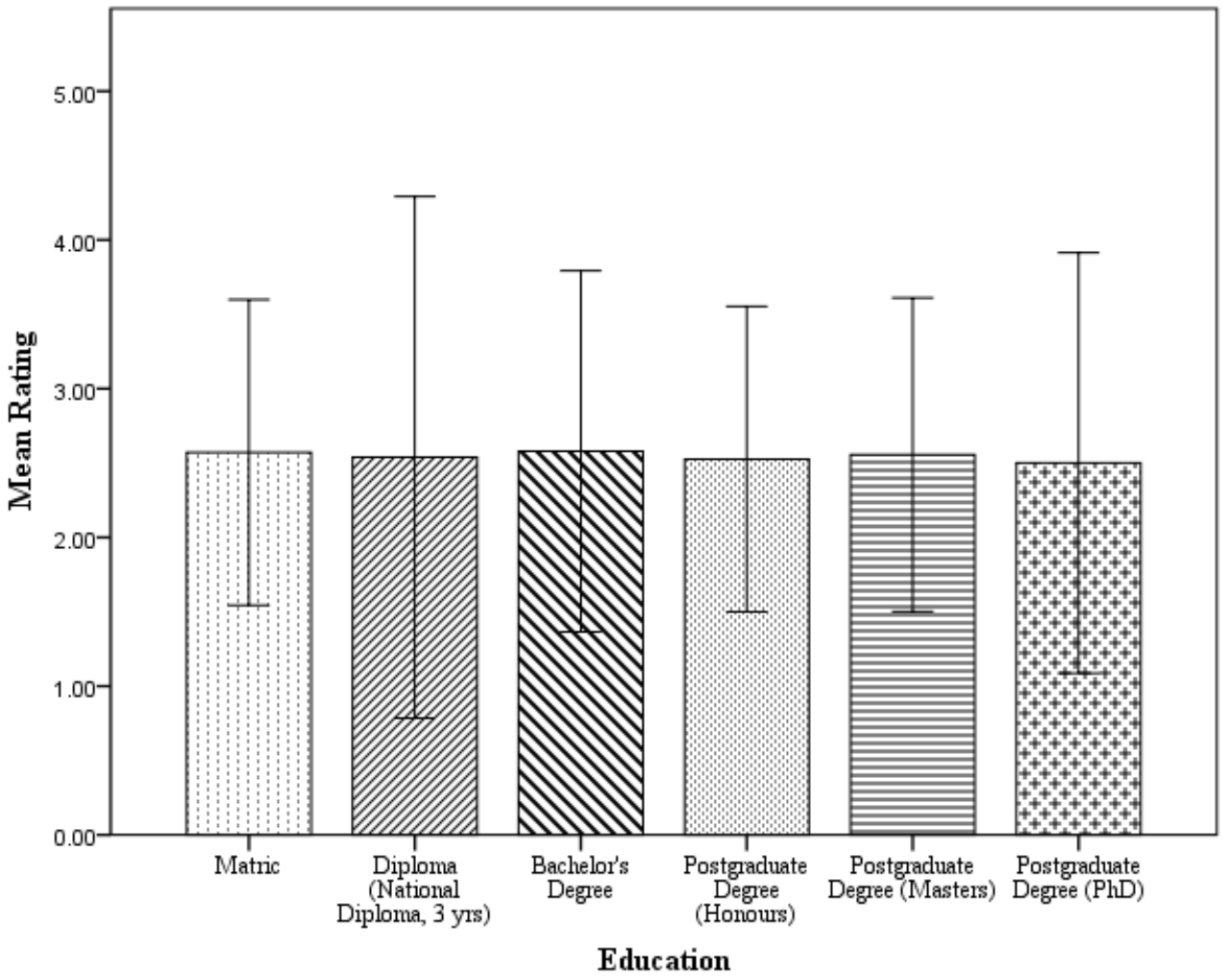


Figure 3: Graphic representation of mean ratings and error bars across educational levels.

Tests of Hypotheses

The Mann-Whitney U test was used to test for differences in medians between two independent groups for each of the hypotheses. Hypothesis 1 stated that interviewer ratings of

identical responses to interview questions will be significantly more favourable when interviewers are provided with a high anchor rather than no anchor.The results were not significant for the first hypothesis (U = 280.50, z = -0.97, p = .33, n.s.), indicating that the distribution of ratings across the high and control conditions is the same (control condition mean rank = 28, high condition mean rank = 24). Thus, interviewer ratings are the same when interviewers are provided with a high anchor or no anchor. Hypothesis 1 was thus rejected and the null hypothesis was retained.

The second hypothesis stated that interviewer ratings of identical responses to interview questions will be significantly less favourable when interviewers are provided with a low anchor rather than no anchor. The results were not significant for Hypothesis 2 (U = 253.00, z = -1.29, p = .20, n.s.), indicating that the distribution of ratings across the low and control conditions is the same (control condition mean rank = 28, high condition mean rank = 23). Thus, interviewer ratings are the same when interviewers are provided with a low anchor or no anchor. Therefore, Hypothesis 2 was rejected and the null hypothesis was retained.

As our results may have been affected by the decision to delete outliers, both hypotheses were tested again including outliers in the data set. This led to the same conclusion as when the outliers were excluded. Hypothesis 1 (U = 306.50, z = -0.66, p = .51, n.s) and Hypothesis 2 (U = 277.00, z = -1.23, p = .22, n.s.).

Follow-up Tests

Independent samples t-test. Despite the lack of a normal distribution, an independent samples t-test was conducted for each hypothesis in order to satisfy the argument that the t-test is robust to deviations from normality (Green & Salkind, 2008). Non-significant results were found

for both Hypothesis 1 (t = 0.97, p = .34, n.s.) and Hypothesis 2 (t = 1.10, p = .28). Thus, the same results were found for the independent samples t-test and the non-parametric Mann-Whitney U test. Similar results were found for Hypothesis 1 (t = .51, p = .61, n.s.) and Hypothesis 2 (t = 1.12, p = .27, n.s.) when including the outliers in the analyses.

Power. As our results may have been affected by lack of power (Aberson, 2010), we calculated the achieved power using G*Power 3 (Cunningham & McCrum-Gardner, 2007). We calculated power for the current study as low (.28), using the small effect size (d = .31) and a conservative level of significance (p = .05). The achieved power was below .80, a level commonly accepted as a good statistical power (Aberson, 2010). Thus, there is a substantial possibility that the anchoring and adjustment effect was not detected, and that the hypotheses were incorrectly rejected.

Discussion

Main Findings

The purpose of this study was to detect whether the anchoring and adjustment heuristic is present in unstructured interviews or not. Despite past research that would support the hypotheses of the current study, neither of the hypotheses was supported by the results of the statistical analyses. In our study, anchoring and (insufficient) adjustment did not occur in simulated unstructured interviews. This is an unexpected result given the support for anchoring and adjustment in structured interviews (Furnham & Boo, 2011; Kataoka, et al., 1997). In addition, uncertainty is a factor known to foster the anchoring and adjustment heuristic, and the level of uncertainty increases as the structure of the interview lowers (Mussweiler, & Strack, 2000. It therefore follows that anchoring and adjustment should also occur in unstructured interviews if it occurs in structured interviews.

Furthermore, given the nature of the heuristic, anchoring and adjustment would be expected to have occurred in our study. Research has established anchoring and adjustment to be an effortful process by which the magnitude of adjustment is motivated by consequence (Epley & Gilovich, 2006; Simmons, et al., 2010). The higher the consequence of making an incorrect judgement, the more effort raters will invest in achieving what they perceive to be an accurate judgement. This in turn leads to greater (and more sufficient) adjustment from the anchor. There was no consequence for inaccurately rating the interview candidate in our study. It was a simulation, where the candidate's job and the performance of the organisation did not depend on the outcomes of the rating such as they would have in a real-world situation. The subjects, therefore, had no real investment in making an accurate judgement. Thus, subjects could have neglected to invest a significant amount of effort in the rating process, and as a result insufficiently adjusted from the anchor. The results of our study did not reflect support for the presence of anchoring and adjustment.

However, there are other mechanisms underlying the anchoring and adjustment heuristic that may have influenced the results, such as the source of the anchor. A study by Epley and Gilovich (2001) revealed the source of the anchor to have an important influence over the sufficiency of the adjustment. The results of this study suggest that anchoring and adjustment is less likely to take place when the anchor is provided by the experimenter, rather than self-generated. Given that in the current study the high and low anchors were provided by the experimenter, and not self-generated, it is possible that the situation was less conducive to the heuristic. The provision of the anchor by the experimenter points to the possibility that anchoring and adjustment actually failed to occur in the rating of the candidate in our study.

Another factor that may have affected the results is the precision of the anchor provided. Research on numerical anchors show that the more precise the anchor, the less adjustment will occur (Janiszewski & Uy, 2008). The same may apply for verbal anchors. Possibly, if a more general verbal anchor had been used (such as “above average”), and not a precise rating point on the scale (“excellent”), insufficient adjustment may have been observed.

All considered, the meanings of the results are clear. In this particular study, managers who were provided with a high anchor did not insufficiently adjust their ratings, and managers who were provided with a low anchor also did not insufficiently adjust their ratings. Instead, raters assigned a similar rating to the interview candidate, regardless of whether they received an anchor (high or low) or not. Therefore, the anchoring and adjustment heuristic did not seem to generalise from the structured (Kataoka, et al., 1997) to the unstructured interview setting that we simulated. This would imply that low structure in interviews is not a condition that promotes anchoring and adjustment. However, we suggest that our results be replicated before drawing conclusions.

Limitations

There are a number of explanations for our failure to replicate the results of earlier studies conducted in structured interviews (Kataoka, et al., 1997), in an unstructured interview setting, including the sample, stimulus, and possible confounding variables.

The sample size exerts considerable influence over the power of a study. Our study held a one year time limit for completion, and as such, there was insufficient time to source a larger sample from the population. In addition, resource constraints prevented a more valuable incentive being offered in return for participation in our study. For these reasons, only a relatively small sample (approx. 25 subjects per condition) was sourced from the population.

This sample size may have been too small to correctly detect an effect. The effect size was found to be small and the achieved power insufficient for correctly detecting whether the hypotheses should be accepted or rejected (Aberson, 2010). Given the statistical power of the study, it is possible that a significant effect could not be detected and that the hypotheses were rejected in error. However, subject background and previous experience may also have contributed to the unexpected results of this study.

The variety of subjects' backgrounds could also have confounded the results. The sample was unfocused in terms industry, organisation, and previous interview experience. Subjects possessed different levels of interview experience, and some subjects had no previous interview experience at all. This is a concern, because general interview experience has a positive correlation with the quality of interviewer ratings (Olson & Peytchev, 2007). Experienced interviewers are more sensitive to identifying key behaviours and skills that predict a candidate's competence.

However, we tried to limit the effects of differences in subject backgrounds and experiences on our results by providing clear and consistent instructions and candidate information to every subject before they watched the interview video. Furthermore, when we analysed the data excluding those participants who lacked formal interview experience, the results were the same as the initial results. This is probably due to the sample being a field sample consisting of management level employees, rather than a student sample. Given the management element of their jobs, it is likely that the majority of subjects had acquired some kind of ability and experience relevant to rating (e.g. through performance reviews), despite whether they had specifically conducted an interview before or not. Thus, previous interview

experience seems an unlikely explanation for the lack of significant results. The video stimulus, however, may have interfered with the results.

Some subjects may have responded carelessly in order to complete the survey as quickly as possible. Although the forced response setting on the survey programme, Qualtrics, forced subjects to answer all questions, we could not control the effort or intention of subjects when completing the survey. Thus, the quality of responses may have affected the results. However, the duration for survey completion was recorded, and those subjects who completed the study in less than three minutes (a reasonable time for completion of this survey) were excluded from the final data analyses. Careless response is therefore not a plausible explanation for the observed results, but, unintentional problems with viewing the video may have made it difficult for subjects to provide quality responses.

Despite the controls present in our study, the possibility of a third variable must be considered. Subjects did not all view the video in the same, consistent setting. Subjects viewed the video at their own convenience, meaning that they viewed the video at different times of the day, week or month, in different environments and on computers with differing quality in terms of visuals, sound and internet speed. This may have influenced the ratings provided by subjects.

In addition, personal characteristics of the interview candidate may have acted as a confounding variable to subject ratings. A recent study found that initial impressions of candidates based on ethnic and gender differences have a negative effect on the validity of interview ratings (Barrick, Dustin, Giluk, Stewart, Shaffer, & Swider, 2012). Initial impressions of the candidate in our study may have interfered with the ratings, thereby confounding the process of anchoring and adjustment.

However, controls implemented in the research process would have minimised the effects of the two above-mentioned third variables: viewing the video in different settings, and characteristics of the video candidate. The same clear and specific instructions were given to each and every subject regarding the study, information about the interview candidate and the rating process, all subjects watched the exact same video, and all subjects were advised to answer the survey in a quiet environment, ensure that their sound was functional, and attempt to minimise interruptions. These controls make the influence of a third variable improbable.

Implications for Theory and Future Research

The findings of this study contradict the existing research on anchoring and (insufficient) adjustment. In this study, anchoring and (insufficient) adjustment did not occur in unstructured interviews. It is possible that anchoring and adjustment is not as robust a heuristic as research postulates (Chapman & Johnson, 1994; Furnham & Boo, 2011). Anchoring and adjustment has failed to replicate in a number of settings. Brewer and Chapman (2002) attempted to replicate a study by Wilson, Houston, Eitling and Brekker (1996) on the basic numerical anchoring effect. However, the study failed to replicate the results of Wilson, et al. (1996), finding instead that anchoring only occurred under very specific circumstances. Another study on anchoring and adjustment in auctions had findings in contrast to the theory (Chapman & Johnson, 2002). This study found that the heuristic worked in reverse; the lower the anchor, the higher the bid. Thus, the findings of our study contribute to the theory by exposing another setting in which anchoring and adjustment does not replicate according to the theory.

However, if anchoring and adjustment does not occur in unstructured interviews, then it fails to contribute to the standing of the unstructured interview as being an unreliable selection tool. Research has shown that anchoring and adjustment does occur in structured interviews, so it

would seem unlikely that it would not occur in the unstructured interview. It is more plausible to consider that anchoring and adjustment does indeed occur in unstructured interviews, synonymous with the direction of the existing literature, but that the effect was simply not detected in this study due to limitations of the sample size. Thus, more research is required before we can conclude that anchoring and adjustment does not occur in unstructured interviews.

Future research on anchoring and adjustment in unstructured interviews can be enhanced in a number of ways. First, studies should replicate the current study using a larger sample of approximately 200 participants per condition (according to G*Power 3) (Cunningham & McCrum-Gardner, 2007). Next, it is suggested that the sample be more focused within a particular industry or company, using random selection of management and human resource level personnel in a particular industry with previous interview experience. These changes would not only improve the validity of the results, but would improve the generalisability of the results to the broader population. Finally, more control should be exercised in terms of viewing the video and completing the survey. It is suggested that all subjects watch the same video in the same environment (on a screen in a large boardroom or lecture hall) in order to control for any confounding variables relating to the environment. This would allow differences in computer settings, sound, noise, lighting, interruptions, and any other environmental characteristic to be discounted as interferences with the validity of results. More controlled experimental research on the anchoring and adjustment heuristic is thus required. It would be premature to draw implications for practice from our study before results are replicated.

Conclusion

This study investigated the presence of anchoring and adjustment in the unstructured interview setting. Anchoring and adjustment is a heuristic described as a process of simplifying

complex judgemental tasks in situations of uncertainty. Although anchoring and adjustment has been observed within structured interviews, no earlier studies have investigated the anchoring and adjustment heuristic in unstructured interviews. From our findings we conclude that anchoring and adjustment did not occur in our simulated unstructured interviews. These findings are in contradiction to previous research on anchoring and adjustment in highly structured interviews. Thus, we hope that this study will encourage future research into anchoring and adjustment in the unstructured interview setting.

References

Aberson, C. L. (2010). *Applied power analysis for the behavioral science*. New York: Psychology Press.

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods, 16*, 270-301.

Allen, D. (2011). Information and decision making in time-constrained practice: A dual processing perspective. *Journal of the American Society for Information Science and Technology, 62*, 2165-2181.

Barrick, M. R., Dustin, S. L., Giluk, T. L., Stewart, G. L., Shaffer, J. A., & Swider, B. W. (2012). Candidate characteristics driving initial impressions during rapport building: Implications for employment interview validity. *Journal of Occupational & Organizational Psychology, 85*, 330-352.

Blackman, M. C. (2002). Personality judgement and the utility of the unstructured employment interview. *Basic and Applied Social Psychology, 24*, 241-250.

Brewer, N. T., & Chapman, G. T. (2002). The fragile basic anchoring effect. *Journal of Behavioral Decision Making, 15*, 65-77.

Brtek, M. D., & Motowidlo, S. J. (2002). Effects of procedure and outcome accountability on interview validity. *Journal of Applied Psychology, 87*, 185-191.

Burns, R. B., & Burns, R. A. (2008). *Business research methods and statistics using SPSS*. London: Sage.

Chapman, G. B., & Johnson, E. J. (1994). The limits of anchoring. *Journal of Behavioral Decision Making, 7*, 223-242.

Chapman, G. B., & Johnson, E. J. (2002). Incorporating the irrelevant: Anchors in judgements of belief and value. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgement* (pp. 120–138). New York: Cambridge University Press.

Chartered Institute for Personnel Development (CIPD). (2008). *Competency and competency frameworks.* London: Author.

Cohen, J. B. (1988). *Statistical power analysis for the behavioural sciences*. Lawrence Erlbaum: New York.

Cohen, J. B. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Cohen, J. B., & Reed, A. (2006). A multiple pathway anchoring and adjustment (MPAA) model of attitude generation and recruitment. *Journal of Consumer Research, 33*, 1-15.

Critcher, C. R., & Gilovich, T. (2008). Incidental environmental anchors. *Journal of Behavioral Decision Making, 21*, 241−251.

Cronshaw, F. S., & Weisner, W. H. (1989). The validity of the employment interview: Models for research and practice. In R. W. Eder & F. R. Ferris (Eds.). *The employment interview: Theory, research, and practice* (pp. 269-281). Beverly Hills: Sage.

Cunningham, J. B., & McCrum-Gardner, E. (2007). Power, effect and sample size using GPower: Practical issues for researchers and members of research ethics committees. *Evidence Based Midwifery, 5,* 132-136.

Dana, J., Dawes, R., & Peterson, N. (2013). Belief in the unstructured interview: The persistance of an illusion. *Judgement and Decision Making, 8,* 512-520.

De Kock, F. S., Lievens, F., & Born, M. P. (2013). An in-depth look at dispositional reasoning and interviewer accuracy. Manuscript submitted for publication.

Dipboye, R. L., Macan, T., & Shahani-Denning, C. (2012). The selection interview from the interviewer and applicant perspectives: Can't have one without the other. In N. Schmitt (Ed.), *The Oxford handbook of personnel assessment and selection* (pp. 323-352). New York, NY US: Oxford University Press.

Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors. *Psychological Science, 12,* 391-396.

Epley, N., & Gilovich, T. (2006). The anchoring and adjustment heuristic: Why the adjustments are insufficient. *Psychological Science, 17,* 311-318.

Epley, N., & Gilovich, T. (2009). Anchoring unbound. *Journal of Consumer Psychology, 20,* 20-24.

Farr, J. L., & York, C. M. (1975). Amount of information and primacy-recency effects in recruitment decisions. *Personnel Psychology, 28,* 233-238.

Fischer, J., Fischer, P., Englich, B., Aydin, N., & Frey, D. (2011). Empower my decisions: The effects of power gestures on confirmatory information processing. *Journal of Experimental Social Psychology, 47,* 1146-1154.

Fischer, P., Lea, S., Kastenmuller, A., Greitemeyer, T., Fischer, J., & Frey, D. (2011). The process of selective exposure: Why confirmatory information search weakens over time. *Organizational Behavior and Human Decision Processes, 114,* 37-48.

Furnham, A., & Boo, H. C. (2011). A literature review of the anchoring effect. *The Journal of Socio-Economics, 40,* 35-42.

Gehlbach, H., & Barge, S. (2012). Anchoring and adjusting in questionnaire responses. *Basic and Applied Social Psychology, 34,* 417-433.

Green, S. B., & Salkind, N. J. (2008). *Using SPSS for Window and Macintosh: Analyzing and understanding data* (5th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Guo, J. (2012). Optimal sample size planning for the Wilcoxon–Mann–Whitney and van Elteren tests under cost constraints. *Journal of Applied Statistics, 39,* 2153-2164.

Harris, M. M. (1989). Reconsidering the employment interview: A review of recent literature and suggestions for future research. *Personnel Psychology, 42,* 691-726.

Hauptfleisch, D. B. (2012). The moderating effect of interview structure on race-group similarity effects in simulated interview ratings (Unpublished Master's thesis). Stellenbosch University, Stellenbosch.

Heslin, P. A., Latham, G. P., & Vande Walle, D. (2005). The effect of implicit person theory on performance appraisals. *Journal of Applied Psychology, 90,* 842-856.

Highhouse, S. (2009). Designing experiments that generalize. *Organizational Research Methods, 12*(3), 554-566.

Howell, D. C. (2011). *Fundamental statistics for the behavioral sciences* (7th ed.). Belmont: Wadsworth.

Jacobs, R. (2001). *Outliers in statistical analysis: Basic methods of detection and accommodation.* ERIC, *EBSCOhost,* accessed 4 December 2013.

Janiszewski, C., & Uy, D. (2008). Precision of the anchor influences the amount of adjustment. *Psychological Science, 19,* 121-127.

Jones, R. G., & Welsh, J. B. (1971). Ability attribution and impression formation in a strategic game: A limiting case of the primacy effect. *Journal of Personality and Social Psychology, 20,* 166-175.

Joyce, E. J., & Biddle, G. C. (1981). Anchoring and adjustment in probabilistic inference in auditing. *Journal of Accounting Research, 19,* 120-145.

Kataoka, H. C., Latham, G. P., & Whyte, G. (1997). The relative resistance of the situational, patterned behaviour, and conventional structured interviews to anchoring effects. *Human Performance, 10,* 47-63.

Kerlinger, F. N. (1986). *Foundations of behavioral research (3rd ed.).* Orlando, FL: Harcourt Brace & Company.

Landy, F. J. & Bates, F. (1973). Another look at contrast effects in the employment interview. *Journal of Applied Psychology, 58,* 141-144.

Lin, T., Dobbins, G. H., & Farh, J. (1992). A field study of race and age similarity effects on interview ratings in conventional and situational interviews. *Journal of Applied Psychology, 77,* 363-371.

McCarthy, J. M., Iddekinge, C., & Campion, M. A. (2010). Are highly structured job interviews resistant to demographic similarity effects? *Personnel Psychology, 63,* 325-359.

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, *79*, 599-616.

Millsap, R. E., & Maydeu-Olivares, A. (2009). *The SAGE handbook of quantitative methods in psychology*. London: SAGE Publications Ltd.

Moore, L. F., & Lee, A. J. (1974). Comparability of interviewer, group, and individual interview ratings. *Journal of Applied Psychology, 59,* 163-167.

Moutier, S., & Houde, O. (2003). Judgement under uncertainty and conjunction fallacy inhibition training. *Thinking & Reasoning, 9*, 185.

Mussweiler, T., & Englich, B. (2005). Subliminal anchoring: Judgemental consequences and underlying mechanisms. *Organizational Behavior and Human Decision Processes, 98,* 133-143.

Mussweiler, T., & Strack, F. (2000). Numeric judgement under uncertainty: The role of knowledge in anchoring. *Journal of Experimental Social Psychology, 36*, 495–518.

Naylor, R. W., Lamberton, C. P., & Norton, D. A. (2011). Seeing ourselves in others: Reviewer ambiguity, egocentric anchoring, and persuasion. *Journal of Marketing Research, 48,* 617-631.

Northcraft, G. B., & Neale, M. A. (1987). Experts, amateurs, and real estate: An anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes, 39*, 84-97.

Olson, K., & Peytchev, A. (2007). Effect of interviewer experience on interview pace and interviewer attitudes. *Public Opinion Quarterly, 71*, 273-286.

Powell, D. M., & Goffin, R. D. (2009). Assessing personality in the employment interview: The impact of training on rater accuracy. *Human Performance, 22*, 450-465.

Rutledge, R. W. (1993). The effects of group decisions and group-shifts on use of the anchoring and adjustment heuristic. *Social Behavior and Personality, 21*, 215-226.

Ryan, A. M., McFarland, L., Baron, H., & Page, R. (2004). An international look at selection practices: Nation and culture as explanations of variability in practice. *Personnel Psychology, 52*, 359–392.

Shafir, E., & LeBoeuf, R. A. (2009). Anchoring on the “here” and “now” in time and distance judgments. *Journal of Experimental Psychology: Learning, Memory and Cognition, 35*(1), 81-93.

Simmons, J. P., LeBoeuf, R. A., & Nelson, L. D. (2010). The effect of accuracy motivation on anchoring and adjustment: Do people adjust from provided anchors. *Journal of Personality and Social Psychology, 99*, 917-932.

Steiner, D. D., & Rain, J. S. (1989). Immediate and delayed primacy and recency effects in performance evaluation. *Journal of Applied Psychology, 74*, 136-142.

Tamir, D. I., & Mitchell, J. P. (2013). Anchoring and adjustment during social inferences. *Journal of Experimental Psychology: General, 142,* 151-162.

Thornstein, T. J., Breier, J., Atwell, A., Hamilton, C., & Privette, M. (2008). Anchoring effects on performance judgements. *Organizational Behavior and Human Decision Processes, 107*, 29-40.

Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science, 185*, 1124-1130.

Van der Zee, K. I., Bakker, P., & Bakker, A. B. (2002). Why are structured interviews so rarely used in personnel selection? *Journal of Applied Psychology, 87,* 176-184.

Wilson, T. D., Houston, C. E., Eitling, K. M., & Brekker, M. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General, 125,* 387-402.

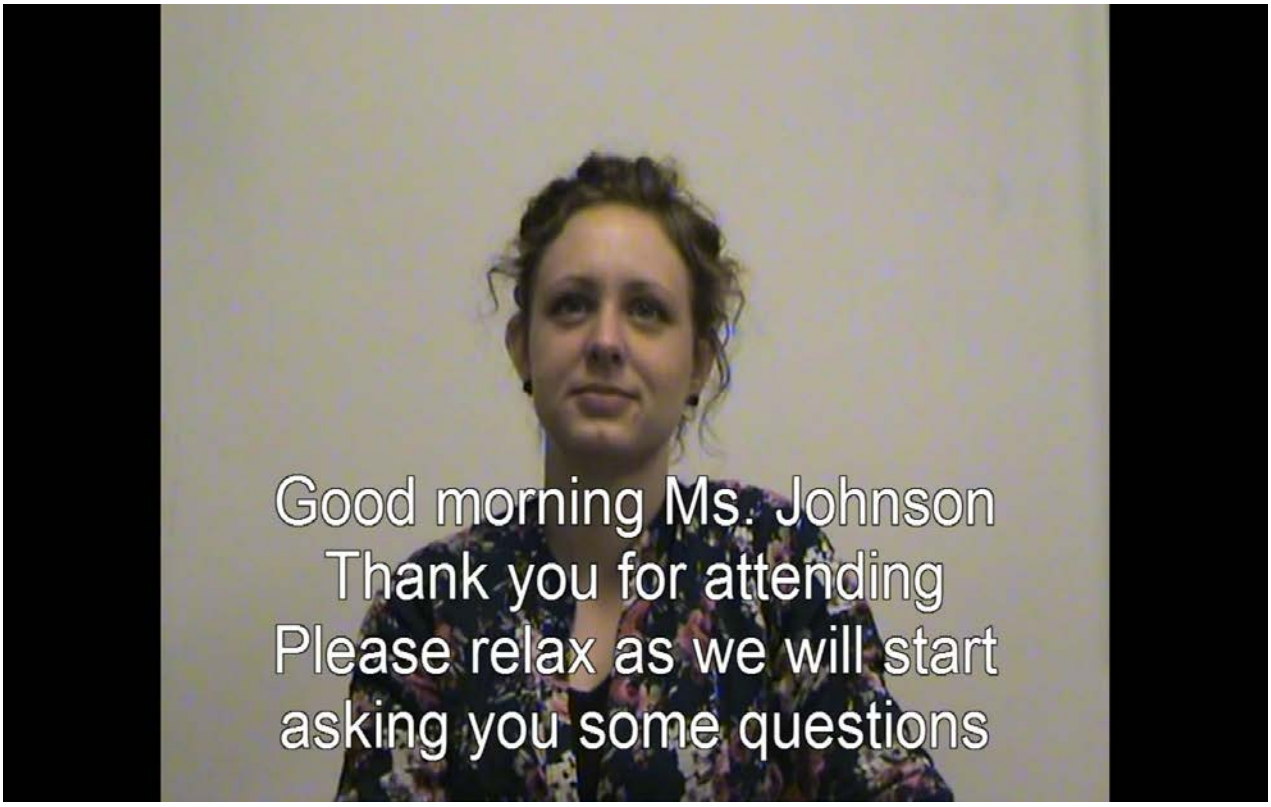
Zaller, J. R. (1992). *The nature and origins of mass opinion*. New York: Cambridge University Press.

Zimmerman, D. W. (2003). A warning about the large-sample Wilcoxon-Mann-Whitney test: Earlier studies conducted in structured interviews. *Understanding Statistics, 2,* 267-280.

Zuckerman, A. S., Kotler-Berkowitz, L. A., & Swaine, L. A. (1998). Anchoring political preferences: The structural bases of stable electoral decisions and political attitudes in Britain. *European Journal of Political Research, 33*, 285-321.

Appendix A

Screenshot of the Unstructured Interview Video Stimulus



Appendix B

Unstructured Interview Questions

1) *Tell me about yourself and your background. Who are you and where do you come from?*
2) *What are your hobbies and interests?*
3) *Where do you see yourself in five years' time?*
4) *How do you usually handle interpersonal conflict situations?*
5) *What are some of your personal strengths and weaknesses?*

Appendix C

Example Interview Rating Item – Control Condition