Inter-Test Reliability and External Influences on nAch Testing Methods

Heather McCaig
This study looks at the inter-test reliability between Smith’s Quick Measure of Motivation Achievement, the sentence stem method by Sorrentino, and a self-report on Need for Achievement. As well, it looks at the influences of an external stimulus, one of two written stories in an arousal condition and a control condition. The results for the inter-test reliability were not significant. The results of the influence of the conditions were not significant in all groups except in the case of one observer’s scores of the sentence stem test. Future research suggests a larger sample size, and different methods for the conditions.

David C. McClelland’s trait, Need for Achievement (nAch), which defines people who seek out success based purely on the need for success and achievement has been the topic of much study throughout the past 50 years. According to his theory, these people are driven not by a desire for material success, nor to surpass their peers. McClelland devised the now well-known test, the Thematic Apperception Test (TAT), in which participants are presented with a picture that they must interpret through a paragraph-long story. According to McClelland, themes are unconsciously written into the story, which can be interpreted through the researcher’s analysis. If the story involves “competition with a standard of excellence,” it is eligible for nAch analysis, and can be further analysed for more nAch characteristics.

McClelland believed this to be the most reliable method of analysis, and that a more obvious method would not yield as accurate results. A self-
Comparison of nAch Tests With Varying Arousal States

report or a questionnaire might bias the answers of the participant. Though many researchers today still ascribe to his method of testing, many more researchers have contested that it is an inefficient and cumbersome test. Of the various methods created to test nAch to date, two were chosen to be used in the current study, as well as the addition of a self-report method. Smith’s Quick Measure of Motivational Achievement (Smith, 1973) claims to be a very brief and efficient method of testing nAch, without the time required for the TAT. It was created with the assumption that people are, indeed, able to be tested in a more overt and open way; they do not have to be completely unaware of what is being tested. As well, a revised version of the TAT has been devised, obtained from the lab of Dr. Richard Sorrentino of the University of Western Ontario, in which a sentence stem is presented in place of McClelland’s traditionally used pictures. The sentence stem allows for flexibility of the stimulus in the imagination of the participant, as well as preventing it from becoming outdated as a picture inevitably would. Most importantly, though, it is not necessarily limited by the image presented in a picture, such as the photograph of the young Caucasian boy from the 1950s sitting in front of a violin. The ability of the participant to visualize the stimulus person in a way that comes naturally to them may facilitate more natural results from the test.

The Sentence Completion Test, by Dr. Jules Holzberg (refer to McClelland, 1953), was one of the original methods of analysing achievement
motivation. It involved 40 sentence stems that the participant was to complete, in one sentence. As a result of the poor results it yielded, it was deemed as not as effective as the TAT.

Despite all the new tests developed and the revisions of old tests, no specific test has been universally proclaimed as the most effective model. Much debate centres on this, and it is the goal of the current study to find if the results of the two above-mentioned methods can be correlated.

McClelland also believes that nAch is a trait that is independent of the influence of external events.

The basis for McClelland’s test was that in order to reveal one’s nAch score, a less than obvious means was necessary, such as the TAT method of reading into the hidden meanings of the stories. McClelland believed, as do many researchers in this area, that people are unable to accurately self-report this trait. Of the tests available, the most popular rely on a lack of awareness of the goal of the tests. The focus of prior research says that, because this is seen as a desirable trait, most people would not give an accurate report on how much they have it. Whether this is because they do not have accurate self-knowledge or because they do not want to be cast in a negative light in the eyes of the researcher is unknown. However, it is possible that with a proper definition of nAch to work from, it may be that a self-report could yield the same results as the two aforementioned methods of testing.
McClelland also claims that this trait is not influenced or biased by external sources or situations, such as a competitive atmosphere.

Blackenship, Vega, Ramos, Romero, Warren, Keenan, et al. (2006), created a new and updated version of the TAT by introducing 6 new pictures, attempting to modernize it. Many researchers no longer use the TAT by itself, preferring some more convenient method of scoring, such as the Achievement Motivation Questionnaire.

Sinha & Prasad (1978) used both TAT cards as well as the Achievement Motivation Questionnaire in conjunction to test the inter-test reliability. They found a fairly high correlation of 0.61. The researchers suggested that the two tests be used together in an effort to provide a more accurate measure. A later study by Hundal and Singh (1979) recommended the Rogers and Niell Sentence Completion Achievement Motivation Scale as a substitute for the TAT.

**Method**

**Subject**

There were 36 participants, the majority of who were drawn from the university population at both Huron University College and the University of Western Ontario. All the participants had at least one year of university-level education. Although no data were collected regarding age, the age ranged from about 18 years to 40 years old. They were usually approached in public settings, such as public sitting areas, and asked to volunteer their time.
Comparison of nAch Tests With Varying Arousal States

Materials

Smith’s Quick Measure of Motivational Achievement was one of the three tests used to test nAch. It consists of a list of 17 true or false questions, 10 of which test motivational characteristics, and seven of which test carelessness, which were disregarded in the present study. A version of McClelland’s Thematic Apperception Test, which was taken from the lab of Dr Richard Sorrentino of the University of Western Ontario, was used to compare to the Quick Measure test. He used four sentence stems as a prompter for a story, in place of McClelland’s photographs. Of the four lines only one was used for the current study (“They were sitting, wondering what will happen...”), however, it was felt that “they” was too general a pronoun, and might inhibit the flow of the story. It was therefore replaced with the non-gender specific name “Sam” (“Sam was sitting, wondering what would happen...”). The last of the three tests used was a self report using a scale of 1 – 10 accompanied by a definition of nAch, which asked participants to circle the number that they felt represented how they would score when tested for nAch. This test, including the definition, has been included in the Appendix A.

As well, the researcher wrote the two stories that defined the two conditions of the study. The story in the arousal condition included competition between two individuals for a highly sought after placement in a school. The story of the second condition included no element of competition,
rather detailed the progression of a childhood relationship. The stories were both in between 350-390 words long (approximately one page), and have been included in Appendix B.

Procedure

The participants, after filling out a consent form, were told to read the story on the first page, and then to follow the instructions given on each proceeding page. It was found that many people were confused with the instructions for the sentence-writing test, so further clarification was given at the outset of the study, mainly saying to write any story that came to mind, and that it was not related to any previous material in the test. Clarification was also needed by about one third of the sample of participants during the Quick Measure test. Many of the participants were unclear on some of the coined terms (such as “trier,” meaning “one who tries”), as well as some unfamiliar terms (such as “red tape”). The three tests were given in one of six orders (abc, acb, bca, etc), in an attempt to prevent the results from a bias based on this. The debriefing statement that was given upon completion of the test is also included in the appendix.

Results

All results were calculated using SPSS software. Dependent samples t-tests were done between each of the tests to determine whether statistically significant differences existed between the scores. The self-report measure (M = 7.31) differed significantly from the Quick Measure (M = 6.26), t = 3.448, p <
Comparison of nAch Tests With Varying Arousal States

0.05, and differed significantly from observer 1’s scores of the sentence stem test \((M = 2.771), t = 6.948, p < 0.05\), as well as observer 2’s scores of the sentence stem test \((M = 2.200), t = 8.454, p < 0.05\). The Quick Measure differed significantly from the sentence stem group, (observer 1) \(t = 5.417, p < 0.05\), (observer 2) \(t = 7.106, p = 0.05\).

Due to the differences between measures found in this study, the differences between the arousal and control groups were calculated for each measure using independent samples t-tests. The self-report in the arousal group \((M = 7.176)\) did not differ significantly from the control group \((M = 7.444), t = 0.541, n.s.\). Using the Quick Measure the arousal group \((M = 6.235)\) did not differ from the control group \((M = 6.277), t = 0.103, n.s.\). Using the sentence stem group from observer 1, the arousal group \((M = 1.823)\) did not differ from the control group \((M = 3.666), t = 1.462, n.s.\). The sentence stem group scores from observer 2, however, did differ significantly between the arousal group \((M = 0.882)\) and the control group \((M = 3.444), t = 2.510, p < 0.05\).

The inter-observer reliability of the two measures of the sentence stem test was relatively high, with \(r = 0.738, p < 0.05\).

Discussion

The results of the dependent samples t-tests, which looked at the differences between the scores of each of the tests, do not support the claim that the Quick Measure and sentence stem tests, as well as the self-report test,
are testing the same trait. The results of the independent t-tests, which looked at the influence of the two conditions on each of the tests, were mixed. As the results between the two conditions were significant for observer 2's scores for the sentence stem test and the results for observer 1 was close to significant, it is possible that, given a larger sample size, both observer's would have rated the difference between the arousal states as significant. It is possible that this method of testing is more susceptible to external situations than McClelland believed.

The results of the self-report and the Quick Measure tests were seemingly not influenced by the conditions, pointing to a possibility that these two tests are less susceptible to external situations. It may be that people, when questioned directly, have a more concrete and stable view of themselves, compared to the less concrete results of the story-telling technique in the sentence stem test. The creative licence given within the context of the sentence stem test may be too open to outside influence - if the participant was unable to generate an idea to write on, they may have used the story as a cue. It is interesting to note that, many participants, after filling out the tests, told the researcher that they intentionally used the arousal and control stories read at the beginning of the study as a basis for the sentence stem test. In the control group, there were two direct references and one indirect reference to the control story; in the arousal group, there were four direct references to the arousal story, and three stories that mirrored the
Comparison of nAch Tests With Varying Arousal States

theme (acceptance into a specialized or graduate school). It may be that the competitive nature of the arousal story exerted more influence on the attention of the participants, causing them to use it as material for their own stories, whereas the control group story may have exerted influence to a lesser degree. It was found, upon questioning these people that the references stemmed from a lack of ideas for the short story. Future studies might look at trying to create conditions based not on stories but using some other more subtle method. A photograph or a dramatized video may be less prone to being copied by the participants. A distracter task between the Condition Stories and the tests would perhaps solve this dilemma, although through time passage the effects of the conditions may lessen. Ultimately, however, a lack of creativity is unavoidable when sampling using the sentence stem completion task.

It may be that with a full and accurate description of nAch, people may be able to use the self-report method more effectively. It may be that, upon hearing the original definition of the term, they believed it to encompass more a general desire for achievement. By rating themselves low, they would not only be identifying themselves in their eyes as having low ambition, but also feel that it is not accurate. What they may fail to see is that it is actually measuring something far more specific.
Perhaps, though, how easily people are influenced by the Condition Stories is not a failing of the study, rather it shows just how susceptible people really are to external influences when being tested in this manner.
a) Write a story to accompany the following sentence:

_Sam was sitting, wondering what would happen..._
c)

nAch Defined:

nAch is a term coined by David C. McClelland which stands for Need for Achievement. Essentially, it means the extent to which you are driven to achieve based on simply the reward to achievement itself. It does not take into account peer pressure or competition, or external goals such as passing or failing. Basically, it's success for success' sake, and is seen when one knows that performance will be evaluated by oneself or by another based on some standard of excellence, and the result of one's actions will lead to either a favourable or unfavourable report.

On the scale below (1 is very low and 10 very high) circle the number that you feel represents how you would score when tested for nAch:

1   2   3   4   5   6   7   8   9   10
b) Read each of the following statements, and circle the TRUE or FALSE as it applies to you.

I am not clear about the instructions for this test  TRUE  FALSE
When I was young I enjoyed the feeling of accomplishment after I had done something well  TRUE  FALSE
The satisfaction of a job well done is a great satisfaction  TRUE  FALSE
I don’t think I’m a good trier  TRUE  FALSE
I would sooner admire a winner than win myself  TRUE  FALSE
If it’s worth doing, it’s worth doing well  TRUE  FALSE
Failure is no sin  TRUE  FALSE
Incentives do more harm than good  TRUE  FALSE
In an unknown situation it doesn’t pay to be pessimistic  TRUE  FALSE
I dislike red tape  TRUE  FALSE
I work best when I have a job that I like  TRUE  FALSE
It’s never best to set one’s own challenges  TRUE  FALSE
I don’t care what others do, I go my own way  TRUE  FALSE
Even a good poker player can’t do much with a poor hand  TRUE  FALSE
Modern life isn’t too competitive  TRUE  FALSE
You can try too hard sometimes, it’s best to let the world drift by  TRUE  FALSE
Most people want success because it brings respect  TRUE  FALSE
1.

Paul and Anna were both in the auditorium, practicing their songs before they had to play tonight. Nervously anticipating the interview that would let one of them live out their dreams, they both concentrated on the task at hand. They had been rivals for years, and this night would finally clinch who was the better violin player, and this was the night they'd been working for for years. There had been weeks of interviews and applications, and the decision would be made tonight. They both declared out loud that they'd get it and they both secretly doubted it.

Their anxiety came to a climax, as first Anna played. She could see in the crowd her mom, dad, and sister watching. Was that doubt she read in their faces?

She had always known Paul had natural talent. As opposed to her learned skills, his playing always seemed to be almost beyond his control - like it came without effort. For her to even compete with him, she had to practice three times more.

As she finished her song, she thanked the judges, bowed, and walked quickly off-stage, nodding briefly to Paul as he walked on.

As Paul stood ready to play, he thought of all that was wrapped up in this moment - years of work behind him, and God willing, years of work ahead. This moment represented both his past and his future. He knew Anna was a very dedicated violinist, and it had been hard to compete with her determination.

As he started to play, it felt like he was doing what he was created for, and the judges seemed to feel this. When he finished, and walked off stage, there was the nerve-wracking few minutes where the judges made their decision. The decision would be a hard one, they knew but after much deliberation, they had an answer.

With much ceremony, they announced that the spot was to be awarded to Paul.
They met when they were in Miss Moore's class. That was grade 2. Terry had just moved to town; Chris had grown up here. As soon as they met, they knew they'd be friends, and ever since they've been inseperable. Even during the summer of grade 5, when Chris had to stay with his dad half an hour away, they spent every weekend hanging out somehow. They were the kind of friends that, if you mentioned the name of one without the other, no one would know who you were talking about. "Chris who? Oh, you mean Chris-n'-Terry, of course I know them." Always having fun, going on adventures, whether it was riding their bikes all around town or building forts in the back-forty of Terry's house, they were always up to something.

There were tests to their friendship, though. Like the time that Terry liked Cindy. That was trouble, and everyone thought this could be the end of it all. Terry spent all of his time with this girl, and it people started to refer to them without mentioning the other. This didn't last too long, though, because as it turned out, Cindy liked someone named Max, whose full name, you know, was Maximillian. Really, now, how could Terry crush on a girl who liked someone named Maximillian? A James, perhaps, but not a Maximillian - what was the girl thinking? They both agreed it just couldn't be done, so they left Cindy behind and the friendship was reinstated.

Then, there was that time that Chris got sick. It was pretty serious for a while, he was in the hospital for a few weeks, making him miss an entire semester of school. His mom even took off time from her work to take care of him. But after a while, the doctors said they'd fixed it, and soon he could go back to living a normal life, and he did.

Now a new chapter is opening up, with graduation looming ahead. Their ambitions were leading them to opposite parts of the country, providing a new test to their friendship. Time would tell what this would mean to their friendship, but who knows, maybe in a few years, they'll come back home and be up to the same old antics, the only difference being that they'd be all grown up. Sort of.
The study you just completed tested three ways of measuring one trait, David C. McClelland's Need for Achievement (nAch) (McClelland, 1961). The purpose of completing all three tests was to correlate the scores between the methods, in order to examine the inter-test reliability. As well, you were part of one of two groups, an arousal group or a control group - the first consisted of reading a story involving competition, and the second consisted of reading a story not involving competition. This was done to find any influences level of arousal would have on the scores of each of the three tests. It is predicted that the arousal condition will influence the outcome of the scores of the three tests.

For more information on nAch, you can look up McClelland’s books, *The Achievement Motive*, and *The Achieving Society* (1961), which can both be found in D.B. Weldon Library.

If you would like to know the results of this study, feel free to email me at hmccaig@uwo.ca, after April 20th.

Thank you for participating in my study!

Heather McCaig
References


