Use of Need for Closure Scale to Predict and Affect Individual Tendency for Design Fixation

by

Song-Liang Lai

A thesis submitted in conformity with the requirements for the degree of Master of Applied Science
Department of Mechanical and Industrial Engineering
University of Toronto

© Copyright by Song-Liang Lai 2016
Use of Need for Closure Scale to Predict and Affect Individual Tendency for Design Fixation

Song-Liang Lai

Master of Applied Science

Department of Mechanical and Industrial Engineering
University of Toronto

2016

Abstract

Not all individuals may be equally susceptible to design fixation. I sought to identify characteristics that could predict and affect individual tendency for design fixation, and explored the use of Kruglanski’s Need for Closure Scale for this purpose. The Need for Closure Scale is an individual/dispositional difference variable, as well as an environmental variable that can be manipulated in the laboratory. I devised two experiments that involved developing concepts for which an example solution was provided. The first experiment investigated whether correlations exist between participants’ score on the Need for Closure Scale and the degree of fixation in concepts elicited. A pilot experiment and second experiment investigated whether environmentally-induced Need for Closure could be used to change participants’ likelihood to fixate. Significant results were found in the first and second experiments, supporting that Need for Closure can be used to predict and affect individual tendency for design fixation.
Acknowledgments

Thank you to my supervisor, Professor L.H. Shu for her vision and guidance, all experiment participants for their contributions, and Professor Birsen Donmez for her assistance in the statistical analyses.

I am thankful for the support of my parents, brother, and friends.

Thank you to my various lab mates for their support and company: Jayesh Srivastava, Michael Koerber, Amanda Chou, Greg Fukakusa, Mayukh Chakraborty, and Thomas Schaepsmeyer.

Finally, I gratefully acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada (NSERC).
# Table of contents

Acknowledgments ................................................................................................................................. iii

Table of contents ................................................................................................................................ iv

List of tables ........................................................................................................................................ vii

List of figures ........................................................................................................................................ viii

List of appendices ................................................................................................................................. x

Introduction ........................................................................................................................................... 1

1 Literature review ................................................................................................................................. 2

1.1 Design fixation ................................................................................................................................. 2

1.1.1 Classification and definition ..................................................................................................... 2

1.1.2 Contributing factors ................................................................................................................. 2

1.1.3 Overcoming fixation .................................................................................................................... 3

1.2 Need for Closure ............................................................................................................................. 4

1.2.1 Dispositional Need for Closure measurement and validation ................................................ 4

1.2.2 Environmental Need for Closure manipulation ....................................................................... 4

1.2.3 Short form of the Need for Closure Scale ............................................................................... 5

1.3 Relationship between design fixation and Need for Closure ....................................................... 5

2 Terminology and research goals ......................................................................................................... 6

2.1 Terminology and nomenclature ..................................................................................................... 6

2.2 Research goals ................................................................................................................................ 7

3 Experiment 1: NFC<sub>d</sub> and tendency for design fixation ............................................................ 8

3.1 Participants ....................................................................................................................................... 8

3.2 Method ............................................................................................................................................. 8

3.2.1 Location ..................................................................................................................................... 8

3.2.2 Part 1: NFC<sub>d</sub> Scale .......................................................................................................... 9

3.2.3 Part 2: Problem for concept generation .................................................................................... 9
3.2.4 Part 2: Concept generation and submission ......................................................... 12
3.2.5 Experiment 1 debriefing ......................................................................................... 12
3.3 Evaluation of design fixation ..................................................................................... 12
3.4 Results .......................................................................................................................... 16
  3.4.1 Observations .......................................................................................................... 16
  3.4.2 Group comparison: low NFC_d vs. high NFC_d ................................................. 18
  3.4.3 Comparing individual differences (NFC_d) ...................................................... 19
3.5 Discussion: effect of NFC_d on design fixation ....................................................... 22

4 Pilot Experiment: NFC_e and tendency for design fixation ........................................ 24
  4.1 Participants ................................................................................................................ 24
  4.2 Method ....................................................................................................................... 24
    4.2.1 Formation of groups prior to experiment ....................................................... 24
    4.2.2 Game of Taboo ................................................................................................... 25
    4.2.3 Experimental conditions and relation to NFC_e ............................................. 25
    4.2.4 Implementation .................................................................................................. 26
    4.2.5 Part 1: Round 1 - Peanut-sheller problem .................................................... 27
    4.2.6 Part 2: Round 2 - Alarm-clock problem ........................................................ 28
    4.2.7 Part 3: 15-item NFC_d Scale ............................................................................ 30
    4.2.8 Debrief ................................................................................................................ 31
  4.3 Results .......................................................................................................................... 31
    4.3.1 Fixation evaluation .............................................................................................. 31
    4.3.2 Observations ....................................................................................................... 34
    4.3.3 Condition differences ......................................................................................... 37
    4.3.4 Comparing individual differences ...................................................................... 38
  4.4 Discussion: effect of NFC_e on design fixation ....................................................... 39

5 Experiment 2: NFC_d and NFC_e and tendency for design fixation ............................ 42
5.1 Participants ........................................................................................................................................... 42

5.2 Method ............................................................................................................................................... 42
   5.2.1 Human Intelligence Task (HIT) set-up and recruitment ......................................................... 43
   5.2.2 Experiment registration and group assignment ................................................................. 43
   5.2.3 Part 1: Design problem .................................................................................................. 43
   5.2.4 Part 1: Preparation for concept generation ................................................................ 44
   5.2.5 Part 1: Manipulation of NFC during concept generation ........................................... 44
   5.2.6 Part 2: NFC Scale ........................................................................................................... 44
   5.2.7 Part 3: Exit survey ........................................................................................................... 44
   5.2.8 Experiment debriefing .................................................................................................. 46

5.3 Evaluation of design fixation ........................................................................................................... 46

5.4 Results and discussion ......................................................................................................................... 46
   5.4.1 Observations .................................................................................................................... 47
   5.4.2 Experimental group differences .................................................................................... 49

5.5 Discussion: effect of NFC on design fixation .................................................................................. 55

6 Conclusion and Next Steps .................................................................................................................... 56

References ................................................................................................................................................. 58

Appendices ............................................................................................................................................... 62
List of tables

Table 1: Terminology and nomenclature ........................................................................................................... 7
Table 2: Fixation score rating scheme .................................................................................................................. 13
Table 3: Evaluations of sample concepts ............................................................................................................ 13
Table 4: Wilcoxon Rank-Sum test groups ............................................................................................................ 18
Table 5: Log odds ratio for increase by one in NFC\textsubscript{d} score on fixation score ................................. 20
Table 6: Odds ratio and corresponding probabilities of a more fixated concept for various increases in NFC\textsubscript{d} score ...................................................................................................................... 21
Table 7: Allotted Taboo game-round time per number of submitted concepts ................................................. 26
Table 8: Peanut-sheller fixation rating scheme (Linsey et al., (2010) (2013)) ................................................. 31
Table 9: Alarm-clock fixation rating scheme (Genco et al., (2012)) ............................................................ 32
Table 10: Peanut-sheller problem - uncertain cases ......................................................................................... 33
Table 11: Alarm-clock problem - uncertain cases ............................................................................................ 33
Table 12: Intervention vs. control condition differences .................................................................................... 38
Table 13: Log odds ratio for increase by one in NFC\textsubscript{d} Score on fixation score ..................................... 39
Table 14: NFC\textsubscript{e} manipulation during concept generation (experiment 2) .......................................... 45
Table 15: General comparisons between experimental groups ............................................................................ 52
Table 16: Wilcoxon Rank-Sum test results between experimental groups ..................................................... 53
Table 17: Experiment 2 gender differences ...................................................................................................... 54
List of figures

Figure 1: Design problem description provided to participants ........................................... 10

Figure 2: Examples provided to participants ........................................................................ 11

Figure 3: Participant D, concept 2 of 3, similarity ratings: form 5/5, function 5/5; fixation score 10/10 ........................................................................................................................................ 14

Figure 4: Participant I, concept 3 of 8: similarity ratings: form 3/5, function 5/5; fixation score 8/10 ........................................................................................................................................ 14

Figure 5: Participant J, concept 1 of 8, similarity ratings: form 1/5, function 1/5; fixation score 2/10 ........................................................................................................................................ 15

Figure 6: Histogram of NFC_d scores of participants .............................................................. 16

Figure 7: Fixation scores of accepted concepts vs. NFC_d ..................................................... 17

Figure 8: Mean fixation scores of accepted concepts per participant vs. NFC_d .................... 17

Figure 9: Histogram of participants' concept fixation scores ............................................... 19

Figure 10: Histogram of participants’ concept fixation score categories ............................. 19

Figure 11: Experiment 1 probability distribution plots for fixation scores by NFC_d ............ 22

Figure 12: Problem 1 – Device to shell peanuts (Linsey et al., (2010) (2013)) ...................... 28

Figure 13: Problem 2 - Alarm clock (Adapted, (Genco, et al., 2012)) ................................. 29

Figure 14: 15-item NFC_d Scale vs 42-item NFC_d Scale (24 observations) ......................... 30

Figure 15: Histogram of NFC_d scores of participants ......................................................... 34

Figure 16: Mean fixation scores of peanut-sheller concepts (Full – PS7) ............................ 35

Figure 17: Mean fixation scores of peanut-sheller concepts (Adjusted – PS4) ...................... 35

Figure 18: Mean fixation scores of alarm-clock concepts (Full – AC9) .............................. 36
Figure 19: Mean fixation scores of alarm-clock concepts (Adjusted – AC7) .......................... 36

Figure 20: Histogram of NFC_d of participants................................................................. 47

Figure 21: Fixation scores of accepted concepts vs. NFC_d ........................................... 48

Figure 22: Mean fixation scores of accepted concepts per participant vs NFC_d ............ 48

Figure 23: Histogram of NFC_d of participants in Group 1: Low NFC_e......................... 49

Figure 24: Histogram of NFC_d of participants in Group 2: Control .............................. 49

Figure 25: Histogram of NFC_d of participants in Group 3: High NFC_e ..................... 50

Figure 26: Mean fixation scores of accepted concepts per participant vs NFC_d in Group1: Low NFC_e .......................................................... 50

Figure 27: Mean fixation scores of accepted concepts per participant vs NFC_d in Group2: Control ............................................................. 51

Figure 28: Mean fixation scores of accepted concepts per participant vs NFC_d in Group3: High NFC_e ............................................................. 51
List of appendices

Appendix 1: 42-item NFC<sub>d</sub> Scale (as administered in Experiment 1 and Experiment 2) .......... 62

Appendix 2: Profile screen (Experiment 2) ................................................................................. 68

Appendix 3: Drawing canvas (Experiment 2: Group 3, high NFC<sub>e</sub> condition) ..................... 70

Appendix 4: Sample timestamp output ....................................................................................... 71

Appendix 5: Inquisition of NFC estimate (Experiment 2) ............................................................ 73
Introduction

Jansson and Smith (1991) described design fixation as “a blind adherence to a set of ideas or concepts limiting the output of conceptual design”. This limitation or constraint may be viewed as a detriment during conceptual design.

My research goal was to predict and affect individuals’ tendency for design fixation. I explored the use of the Need for Closure Scale, an individual difference variable developed by Arie Kruglanski, for this purpose. Kruglanski (1990) described Need for Closure as a desire for “an answer on a given topic, any answer”, which I believed to be relevant to design fixation.

I anticipated that individuals with high Need for Closure would be more likely to fixate, and devised three experiments (including a pilot experiment) to study the effect of both dispositional and environmental Need for Closure on design fixation.

This thesis is structured into the following chapters:

Chapter 1: A review of relevant research in design fixation and Need for Closure.

Chapter 2: An outline of the research goals and terminology used in this thesis.

Chapter 3: Reporting of Experiment 1, which focused on the effect of dispositional Need for Closure on design fixation.

Chapter 4: Reporting of pilot experiment for Experiment 2, which studied the effects of environmental Need for Closure on design fixation.

Chapter 5: Reporting of Experiment 2, which studied the effects of both dispositional and environmental Need for Closure (i.e. total Need for Closure) on design fixation.

Chapter 6: Conclusion and discussion of the implications of the results.
1 Literature review

1.1 Design fixation

1.1.1 Classification and definition

Youmans and Arciszewski (2014) highlighted the challenge to researchers of inconsistent definitions of design fixation. They classified design fixation phenomena as: unconscious adherence, conscious blocking, and intentional resistance. Unconscious adherence deals with a designer unaware of being influenced by prior designs, conscious blocking occurs when a designer becomes aware of his or her fixation, and intentional resistance is anecdotally, not “fixing what isn’t broken” (Youmans & Arciszewski, 2014). Youmans and Arciszewski (2014) proposed an updated definition of design fixation as “limitations in the inventive design process that occur when designers are biased towards, or are consciously or unconsciously influenced by, a set of conceptual ideas or a previous body of knowledge”.

1.1.2 Contributing factors

Many researchers have worked to identify factors that contribute to design fixation. Purcell and Gero (1996) conducted a study using Jansson and Smith (1991)’s car bike-rack problem and found no evidence of fixation, except for one feature of a pictorial fixating example, noting that fixation may depend on familiarity with an example. Tseng et al. (2008) manipulated information type and timing to study resulting effects on fixation. Viswanathan and Linsey discussed the roles in fixation of physical modeling as a sunk cost (2011) and of expertise (2013). Crilly (2015) interviewed professional designers, who despite being aware of design fixation and ways to avoid it, conceded that fixation is difficult to overcome. Vasconcelos and Crilly (2016) reviewed 25 studies on inspiration and fixation, in which 14 experimental variables were manipulated. Other researchers have examined the effects of representation of fixation as well as the fixating example. For example, Zahner et al. (2010) and Dong and Sarker (2011) studied the differing levels of representation/abstraction involved in fixation. Cheng et al. (2014) presented photographs of full versus partial examples for a design problem, and found that those presented with partial photographs fixated less. Toh et al. (2015) reported the effect of ownership bias during concept
selection, and specifically a gender difference where more males than females tended to select their own ideas during concept selection.

Atilola et al. (2016) explored the effects of differing representations of a fixating example. The fixating design example took on different forms: a sketch, function tree, and a combination of both. Atilola et al. (2016) found among their results that the mean number of repeated example features (where higher means more fixated) were highest in the combination group, followed by sketch, function tree, and control groups. Atilola et al. (2016) concluded that the use of function tree representation was not fixation-inducing in a design problem application.

### 1.1.3 Overcoming fixation

Efforts towards overcoming design fixation have been reported by several researchers including: Chrysikou and Weisberg (2005) instructing participants to avoid problematic elements in an example; Kohn and Smith (2009) using incubation and distraction; Youmans (2011) and Kershaw et al. (2011) using prototyping and critical feedback; Smith and Linsey (2011) adding hints and clues; Moreno et al. (2014) using analogies; Tsenn et al. (2014) using incubation; and Toh et al. (2014) using product dissection, further discussed below.

Toh et al. (2014) studied the relationship between product dissection activities on design fixation as well as individual personality attributes on the product dissection process, in a team-based, first-year engineering design classroom setting. Tasked with “redesigning an electric tooth-brush for increased portability”, participants were allotted 90 minutes to dissect various Oral-B electric toothbrushes (Toh, et al., 2014). A week later, participants were asked to individually generate toothbrush concepts. Toh et al. (2014) found that those with exposure to more parts in the dissection activity generated more ideas and noted that extraverted individuals dissected more brush head parts. Toh et al. (2014) thus postulated that certain personality traits may play a role in the amount of fixation experienced and the number of ideas generated. Overall, undergraduate students exposed to more parts in the product-dissection activity generated less fixated and more ideas when subsequently redesigning the product.
1.2 Need for Closure

1.2.1 Dispositional Need for Closure measurement and validation

Kruglanski and Webster (1993) devised a formal measure of dispositional Need for Closure. Known as the Need for Closure Scale, it includes five different subscales: 1) order and structure, 2) ambiguity, 3) decisiveness, 4) predictability, and 5) close-mindedness. The Need for Closure Scale consists of 42 questions scored on a 6-point scale, for a total Need for Closure score range of 42 (low) – 252 (high). Those with high Need for Closure generally have: preference for order and structure, discomfort with ambiguity, high decisiveness of judgments, desire for predictability, and high close-mindedness, while those with low Need for Closure have the opposite characteristics.

The Need for Closure Scale was validated by Webster and Kruglanski (1994), receiving high internal consistency in groups of n = 281 (Cronbach's α = .84) and n = 172 (Cronbach's α = .84). Test-retest reliability was also high (r = .86).

High Need for Closure corresponds to a desire to attain closure quickly, which may manifest as frantic “seizing” upon relevant cues to come to a quick evaluation, followed by “freezing” or protecting that evaluation (Kruglanski & Webster, 1996). Plaks (2011) described those with high Need for Closure as more likely to make snap judgments, while those with low Need for Closure (or high need to avoid closure) may have enormous difficulty in forming judgments (e.g., deciding what to order from a restaurant menu). Need for Closure is an individual difference (dispositional) variable, as well as an environmental variable that can be manipulated in the laboratory.

1.2.2 Environmental Need for Closure manipulation

Kruglanski and Fishman (2009) reported that Need for Closure may be environmentally manipulated, at least temporarily. For example, Need for Closure can be heightened in time-pressure situations that require immediate decisions. Mayseless and Kruglanski (1987) studied environmental Need for Closure. They asked participants to list as many as possible hypotheses for the identity of everyday objects presented in photographs. These photographs were enlarged to a point where the identity of the objects was unclear. Participants in the high Need for Closure condition were told that reaching firm decisions is an indication of general intelligence. Those in
the low Need for Closure condition were told that correct visual identification is an indication of general intelligence. The number of hypotheses produced, from highest to lowest, came from the low Need for Closure group, a control group, and the high Need for Closure group.

1.2.3 Short form of the Need for Closure Scale

Researchers have developed other versions of the NFC Scale. Roets and Van Hiel (2007)’s revised 41-item NFC Scale used an alternative set of six questions for the decisiveness subscale; the original scale has seven questions. This revised 41-item NFC Scale’s decisiveness subscale had a correlation to the original decisiveness subscale of \( r = .21 \) (\( p < .05 \), \( n = 164 \)). Roets and Van Hiel (2011) then developed a 15-item version of the NFC Scale, which was significantly correlated with their revised 41-item NFC Scale (\( r = .95 \), \( p < .001 \), \( n = 1584 \)). Test-rest reliability was \( r = 0.79 \) (\( n = 93 \)) for the 15-item scale, and \( r = .87 \) (\( n = 93 \)) for the 41-item scale.

1.3 Relationship between design fixation and Need for Closure

There were many parallels to design fixation found in Kruglanski and Webster (1993)’s Need for Closure Scale statements such as: “When trying to solve a problem I often see so many possible options that it's confusing”, “I do not usually consult many different options before forming my own view”, and “When faced with a problem I usually see the one best solution very quickly”.

Hallihan et al. (2012) (2013) found that confirmation bias (which can contribute to design fixation) can lead designers to “ignore or discount factual contradictory information” and “fixate on initial (confirmed) ideas”. Confirmation bias is similar to Kruglanski and Webster’s (1996) explanation of “seizing” and “freezing” in high Need for Closure.
2 Terminology and research goals

2.1 Terminology and nomenclature

The 42-item Need for Closure Scale as devised by Kruglanski and Webster (1993) measures dispositional Need for Closure, and will be referred to as the NFC_d Scale in this thesis. Need for Closure can also be manipulated environmentally, i.e., increased using time pressure (Kruglanski & Freund, 1983), environmental noise (Kruglanski & Webster, 1993), and dullness of the task (Webster, 1993).

Hereafter, I use NFC with the subscripts d, and e to denote dispositional, and environmental Need for Closure, respectively.

In subsequent sections, I use the term “participant” to refer to actual participants of the experiments, and the term “individual” to refer to any given individual of a population, e.g., when interpreting model results. In Experiment 2, I introduce the term “worker” (i.e. Amazon Mechanical Turk worker), which is distinguished from “participant” in that participants are workers who completed the experiment as intended. This is to account for the fact that data from Amazon Mechanical Turk workers can be rejected when they submit poor work, complete a task faster than is reasonably possible, etc. (Mason & Suri, 2012).

Reoccurring terminology and nomenclature are outlined in Table 1.
Table 1: Terminology and nomenclature

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFC</td>
<td>Need for Closure</td>
</tr>
<tr>
<td>NFC&lt;sub&gt;d&lt;/sub&gt;</td>
<td>dispositional Need for Closure</td>
</tr>
<tr>
<td>NFC&lt;sub&gt;e&lt;/sub&gt;</td>
<td>environmental Need for Closure</td>
</tr>
<tr>
<td>participant</td>
<td>actual participants of experiments</td>
</tr>
<tr>
<td>individual</td>
<td>any given individual of a population, e.g., when interpreting model results</td>
</tr>
<tr>
<td>worker</td>
<td>Amazon Mechanical Turk worker, only participant if work is accepted</td>
</tr>
</tbody>
</table>

2.2 Research goals

In Experiment 1, my aim was to investigate if there exists a relation between NFC<sub>d</sub> as measured by the NFC<sub>d</sub> Scale and individual tendency for design fixation. I sought to answer the following research question:

1) Is an individual’s tendency to fixate during concept generation affected by NFC<sub>d</sub>?

In the two subsequent studies, the pilot experiment and Experiment 2, my aim was to investigate whether NFC<sub>e</sub> has an effect on individual tendency for design fixation. I sought to answer the following two research questions:

2) Is an individual’s tendency to fixate during concept generation affected by NFC<sub>e</sub>? (Pilot experiment and Experiment 2)

Answering these questions would assist in the determination of the validity of NFC as a tool to predict and affect individual tendency for design fixation.
3 Experiment 1: NFC\textsubscript{d} and tendency for design fixation

The first experiment, previously reported in Lai and Shu (2016), aimed to study the relationship between NFC\textsubscript{d} and tendency for design fixation.

3.1 Participants

Following approval by the University of Toronto Research Ethics Board, 29 participants were recruited from a fourth-year undergraduate engineering design course. This course covers topics, e.g., design fixation and design for resource conservation, which may affect the results of the study. Therefore, Experiment 1 was scheduled before these topics were introduced in the course. Participants included students studying mechanical engineering, industrial engineering, and engineering science. Of the 29 participants, 22 were male and 7 were female.

3.2 Method

Experiment 1 was completed during a regularly scheduled 3-hour laboratory session, entitled “Creativity exercises” on the course syllabus. The experiment consisted of two parts: Participants 1) completed Kruglanski et al. (1993)’s NFC\textsubscript{d} Scale, and 2) developed concepts intended to reduce the amount of time that users spent running water in a shower.

3.2.1 Location

Instead of the regular laboratory room, the experiment took place in a computer classroom, with computer stations fixed in five rows, all facing the front of the room. Participants were seated with at least one empty computer station between them to reduce their ability to see others’ computer monitors in the same row.

Upon entering the computer classroom, participants were given personal access codes in their information packets. The information packet included instructions to visit the online website and use their personal access codes to complete the prescribed tasks. The personal access codes were non-identifying and enabled participants to anonymously complete the two parts of the experiment tasks online.
3.2.2 Part 1: NFC\textsubscript{d} Scale

The NFC\textsubscript{d} Scale was implemented as an online questionnaire, where participants answered each question by clicking on radio buttons corresponding to a Likert scale (see Appendix 1). Written instructions included, "Read each of the following statements and decide how much you agree with each according to your beliefs and experiences." Participants clicked a submit button at the end of the questionnaire to move on to the next task. If they neglected to answer any questions, they were instructed to return to those questions.

3.2.3 Part 2: Problem for concept generation

Participants were provided the shower design problem (Figure 1) and the Intatec shower (Figure 2) as an example solution. This problem was selected as participants were likely to be familiar with the problem domain. I wrote the specific wording of the problem and added a background to serve as framing for the intended home-use application. The Intatec shower was selected as an example solution because of its differing appearance versus a traditional shower head and time-based mode of operation. I believed that the replication of the time-based mode of operation was a sign of design fixation or more specifically, intentional resistance, or not “fixing what isn’t broken” (Youmans & Arciszewski, 2014).

Potential pitfalls of soliciting concepts for this problem were known due to the laboratory’s other research on resource conservation (Srivastava & Shu, 2013). To address these known pitfalls, I imposed restrictions on acceptable concepts, by excluding water aeration, and the use of information, feedback, or automation as behavior-change strategies (see Figure 1 and Figure 2).
Problem: Reducing water running time while showering

Background: California has declared a state of emergency due to severe drought conditions. Governor Jerry Brown issued Executive Order B-29-15 on April 1, 2015, imposing a statewide 25% reduction in urban water usage.

Description: Your task is to design products that reduce the amount of time spent running water while showering. Your solutions must not incorporate information, feedback or automation. Water aeration is also not an acceptable solution. Information and feedback, such as the information sign shown, can easily be ignored. Automated showers, such as the motion sensor activated model shown below, take away user control and are susceptible to sensor failures, malfunctions, and unintended activation.

A product that does not use information, feedback or automation is the Intatec shower panel shown below. The Intatec shower panel is operated with a push button that activates a timed water flow of approximately 15s.

Definitions:

Information: Tell the user about the problem and the benefits of change.

Feedback: The effect of the actions of the user is conveyed to the user.

Automation: Device performs actions for the user.

Figure 1: Design problem description provided to participants
Figure 2: Examples provided to participants
3.2.4 Part 2: Concept generation and submission

The concept generation activity was also performed online, using canvases of 1000 by 600 pixels. To supplement the written instructions, “Please draw your designs on the canvas below,” verbal instructions provided additional details. For example, participants were told to click and hold the left button of their workstation mouse to draw on the canvas, as well as to only draw one concept per canvas. In addition, clicking the right mouse button opened a text box into which they could type and save annotations for their concepts. A choice of 6 colors (black, blue, red, green, yellow, gray) was available for both drawing and text input. Participants could also undo and redo drawing and text-input actions before submitting their concepts.

While no explicit time limit was given for specific parts of the experiment, participants had up to 1 hour 15 minutes to complete both tasks (the usual duration of the laboratory until a mid-lab break). Participants spent 50 – 75 minutes on the activities, during which they were instructed to not communicate with other participants, and to raise their hands if they had any questions, which I answered individually.

3.2.5 Experiment 1 debriefing

Participants were debriefed in a 15-minute session at the end of the 3-hour laboratory, following activities unrelated to the experiment. Consistent with the ethics protocol, the debriefing consisted of a short explanation of the NFC_d Scale and my research question of whether NFC_d is related to design fixation. I also clarified that the activity did not affect their course grade, and then distributed information letters requesting permission to use their data. All participants consented to the use of their data in Experiment 1.

3.3 Evaluation of design fixation

I measured fixation relative to the provided Intatec shower example through degree of similarity in 1) form and 2) function. This evaluation is similar to the work of Linsey et al., (2010) (2013) (see Table 8) and Genco et al., (2012) (see Table 9), which compared participant concepts to features of provided examples. Table 2 shows my fixation rating scheme, whose application I demonstrate in Table 3 using three example concepts (Figure 3, Figure 4, and Figure 5).
Table 2: Fixation score rating scheme

<table>
<thead>
<tr>
<th>Score</th>
<th>Form Similarity</th>
<th>Function Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Concept borrows no features from Intatec and has novel form</td>
<td>Concept uses strategy other than discretized flow (e.g., require work)</td>
</tr>
<tr>
<td>2</td>
<td>Concept borrows no features from Intatec but looks like a conventional shower</td>
<td>Concept uses resource-based discretized water flow (e.g., packetization of water as in a camping shower)</td>
</tr>
<tr>
<td>3</td>
<td>Concept uses features similar, but not identical, to Intatec (e.g., lever or other timer instead of timed button)</td>
<td>Concept implements other time-based flow i.e. “only on when needed” (e.g., pressure-sensitive pad)</td>
</tr>
<tr>
<td>4</td>
<td>Concept uses at least one identical feature to that used in Intatec</td>
<td>Concept uses an explicit countdown timer (e.g., kitchen timer)</td>
</tr>
<tr>
<td>5</td>
<td>Concept solely uses (multiple) panel(s), push button(s) as in Intatec</td>
<td>Concept uses timed device to provide water flow for set time</td>
</tr>
</tbody>
</table>

Table 3: Evaluations of sample concepts

<table>
<thead>
<tr>
<th>Participant ID concept #/total</th>
<th>Form Similarity Score and Reasoning</th>
<th>Function Similarity Score and Reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td>D Concept 2 of 3</td>
<td>5/5, identical panel and push button elements</td>
<td>5/5, water flow controlled by timed button</td>
</tr>
<tr>
<td>I Concept 3 of 8</td>
<td>3/5, pulling handle instead of pushing button</td>
<td>5/5, water flow controlled by timed button</td>
</tr>
<tr>
<td>J Concept 1 of 8</td>
<td>1/5, no features from Intatec, novel form</td>
<td>1/5, used neither timed nor discretized flow as strategy</td>
</tr>
</tbody>
</table>

Figure 3 shows the second of three concepts, as submitted by participant D. For form similarity, the concept was rated 5/5 because it used an identical panel housing and multiple push buttons. Although the concept used multiple push buttons versus Intatec’s single button, it was still highly similar in form. For function similarity, the concept was rated 5/5, as the buttons had to be functionally identical (timed) push buttons for the concept to work. I treated the absence of explicit notes to the contrary as the participant intending the same function for the same form on the Intatec. In this case, while the participant did not explicitly state that the push button activates a timed water flow, it was rated as such because of the absence of any alternative information.
Figure 3: Participant D, concept 2 of 3, similarity ratings: form 5/5, function 5/5; fixation score 10/10

Figure 4 shows the third of eight concepts, as submitted by participant I. The concept rated 3/5 on Form Similarity because of the similarity between pulling a handle and pushing a button, especially given the explicit transfer of the 15-s timed water flow from the Intatec. The concept rated 5/5 on Function Similarity as the timed-water flow uses the same strategy as the Intatec.

Figure 4: Participant I, concept 3 of 8: similarity ratings: form 3/5, function 5/5; fixation score 8/10

Figure 5 shows the first of eight concepts, as submitted by participant J. The concept rated 1/5 on form similarity because it does not borrow any features from the Intatec, and the “balloon spikes”
constitute a novel form. The repeated use of a showerhead was not counted in form similarity. The concept rated 1/5 on function similarity as it does not use discretized (interrupted) water flow as a resource-conserving strategy. Rather, it aimed to decrease user comfort via “balloon spikes,” which increasingly inflate as the shower runs.

![Figure 5: Participant J, concept 1 of 8, similarity ratings: form 1/5, function 1/5; fixation score 2/10](image)

In total, the 29 participants generated 110 concepts. Two raters, myself and a second rater, both rated all 110 concepts. Concepts were rejected if they did not follow the instructions stipulated in the problem description. Examples of rejected concepts include showers using aeration, and only information and/or automation as resource-conserving strategies. I accepted 88 concepts while the second rater accepted 84 concepts, which were a subset of the 88 concepts accepted by myself.

The two components of rating for each concept, form similarity and function similarity, were summed into a combined similarity, or fixation score. The inter-rater reliability of the fixation scores of the 84 concepts accepted by both raters was calculated as Krippendorff’s α (ordinal) = 0.798, marginally below α ≥ 0.8, which corresponds to perfect agreement (Krippendorff, 2004). Given this high level of agreement, further statistical analyses were conducted using my ratings for my 88 accepted concepts.
3.4 Results

I analyzed the relationship between participants’ NFC\textsubscript{d} and fixation scores at both group and individual levels. At the group level, I assigned participants to two groups based on their NFC\textsubscript{d} scores and investigated whether a difference in fixation scores existed between the groups using a Wilcoxon Rank-Sum test. At the individual level, I used ordinal logistic regression to identify trends between participants’ NFC\textsubscript{d} and fixation scores.

3.4.1 Observations

**Figure 6** shows a histogram of participants’ NFC\textsubscript{d} scores. The range was 107 – 196, the mean was 157.24, and the standard deviation was 21.31.

![Figure 6: Histogram of NFC\textsubscript{d} scores of participants](image)

**Figure 6** and **Figure 8** plot as functions of participants’ NFC\textsubscript{d} scores, 1) fixation scores for each accepted concept, and 2) mean fixation score of accepted concepts per participant. A preliminary assessment was conducted using the Kendall Tau-b test, which assesses the similarity of ranking order between pairs, adjusting for ties. The Kendall Tau-b coefficient $\tau_b$ ranges from -1 (perfect
negative association) to +1 (perfect positive association). I obtained a significant Kendall’s rank correlation of $\tau_b = 0.26$ ($Z = 3.27$, $p = 0.001$) for Figure 7, which indicates a weak positive correlation between participants’ NFC$_d$ and their accepted-concepts’ fixation scores. Kendall Tau-b $\tau_b = 0.21$ ($Z = 1.60$, $p = 0.11$) for Figure 8 indicates a non-significant weak positive association between participants’ NFC$_d$ and their accepted-concepts’ mean fixation scores.

Figure 7: Fixation scores of accepted concepts vs. NFC$_d$

Figure 8: Mean fixation scores of accepted concepts per participant vs. NFC$_d$
3.4.2 Group comparison: low NFC<sub>d</sub> vs. high NFC<sub>d</sub>

Table 4 shows participants placed in low versus high NFC<sub>d</sub> groups based on their NFC<sub>d</sub> score, using a mean split. Participants who scored below the mean of 157.24 were placed in the low NFC<sub>d</sub> group, and the remaining were placed in the high NFC<sub>d</sub> group. I used the mean fixation scores of participants’ concepts to meet the Wilcoxon Rank-Sum test assumption of independent samples. In addition, the ordinal nature of fixation scores did not violate Wilcoxon Rank-Sum test assumption of non-nominal values. The mean fixation scores of participants’ concepts in the low NFC<sub>d</sub> and high NFC<sub>d</sub> groups were 5.90 and 7.56, respectively.

I used a mean instead of a median split as there were 29 participants/data points and I wanted to avoid removing 2 data points (two participants had a NFC<sub>d</sub> score of 159 which was the median). As a non-parametric technique, the Wilcoxon Rank-Sum test is also inherently more conservative compared to parametric techniques. Previously published analyses involving the NFC<sub>d</sub> Scale have used both mean (Holbert & Hansen, 2006) and median splits (Kardes, et al., 2007).

### Table 4: Wilcoxon Rank-Sum test groups

<table>
<thead>
<tr>
<th></th>
<th>Low NFC&lt;sub&gt;d&lt;/sub&gt;</th>
<th>High NFC&lt;sub&gt;d&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants, n</td>
<td>16</td>
<td>13</td>
</tr>
<tr>
<td>Mean NFC&lt;sub&gt;d&lt;/sub&gt; score</td>
<td>142.88</td>
<td>174.92</td>
</tr>
<tr>
<td>Mean participants’ concept fixation scores</td>
<td>5.90</td>
<td>7.56</td>
</tr>
</tbody>
</table>

The Wilcoxon Rank-Sum test revealed significant differences in the mean fixation scores of concepts between participants in the low NFC<sub>d</sub> and high NFC<sub>d</sub> groups (Wilcoxon $W = 243.50, Z = -2.11$, $m_{\text{fixationscore}_{1(\text{high})}} = 7.56$, $m_{\text{fixationscore}_{2(\text{low})}} = 5.90$, $n_1 = 13$, $n_2 = 16$, $p < 0.05$, $r = 0.39$, two-tailed). The effect size was $r = 0.39$, corresponding to a medium effect.

I also tested for, but did not find, significant differences in neither NFC<sub>d</sub> scores ($Z = 0.48, p = 0.63$, $r = 0.09$) nor fixation scores ($Z = -0.92, p = 0.36, r = 0.17$) by gender in the 29 participants (7 female and 22 male). The female group had mean NFC score of 160.29 and mean fixation score of 6.13/10. The male group had mean NFC score of 156.27 and mean fixation score of 6.80/10.
3.4.3 Comparing individual differences (NFC<sub>d</sub>)

Figure 9: Histogram of participants’ concept fixation scores

Figure 10: Histogram of participants’ concept fixation score categories
As the fixation metric is an ordinal variable, i.e., has two or more ordered/ranked categories, I used ordinal logistic regression analysis. This regression is used to predict an ordinal dependent variable given one or more independent variables, and can also use interactions between independent variables to predict the dependent variable (McCullagh, 1980). I selected participants’ NFCd scores as the predictor (independent) variable and fixation score of participants’ concepts as the outcome (dependent) variable. Fixation scores ranged from 2 (least fixated) to 10 (most fixated). Figure 9 shows a histogram of the obtained fixation scores. Because of this large range of possible fixation scores, I condensed it into three categories for ease of interpretation: low, medium, and high fixation. Low fixation comprised of concepts with fixation scores of 2-4, while medium fixation comprised of concepts with fixation scores of 5-7, and high fixation comprised of concepts with fixation scores of 8-10. Figure 10 shows the histogram of the concepts using the three fixation score categories.

Because participants produced multiple concepts, the data consisted of repeated measures. I therefore used generalized estimating equations (GEE) for the model, which was fitted using PROC GENMOD in SAS 9.1, with further specifications of cumlogit link function and multinomial distribution. Table 5 shows significant effects for the predictor variable NFCd (log odds ratio = 0.0243, $\chi^2(1) = 3.90, p < 0.05$), which is equivalent to an odds ratio of 1.025 (exp (0.0243)). This means that if an individual’s NFCd increases by 1 point, there is 1.025 increased odds that individual’s concept will have a higher fixation score relative to the contiguous lower fixation score. Table 6 further expands upon this by showing the corresponding probabilities for an individual’s concept to have a higher fixation score category relative to the contiguous lower fixation score category for a 1, 5, 10, and 20-unit increase in NFCd score.

I assessed the validity of the model using a likelihood-ratio test, which was significant at the 0.05 level ($\chi^2(1) = 6.24, p = 0.0125$).

Table 5: Log odds ratio for increase by one in NFCd score on fixation score

<table>
<thead>
<tr>
<th>Log Odds Ratio</th>
<th>95% Confidence Interval</th>
<th>Wald Test (Type 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0243</td>
<td>0.0002 – 0.0483</td>
<td>$\chi^2(1) = 3.90, p = 0.048^*$</td>
</tr>
</tbody>
</table>

* $p < .05$
Table 6: Odds ratio and corresponding probabilities of a more fixated concept for various increases in NFCd score

<table>
<thead>
<tr>
<th>Increase in NFCd score</th>
<th>Odds Ratio</th>
<th>95% Confidence Interval</th>
<th>Probability**</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.025</td>
<td>1.0002 – 1.0495</td>
<td>50.6%</td>
</tr>
<tr>
<td>5</td>
<td>1.129</td>
<td>1.01 – 1.273</td>
<td>53.0%</td>
</tr>
<tr>
<td>10</td>
<td>1.275</td>
<td>1.02 – 1.62</td>
<td>56.0%</td>
</tr>
<tr>
<td>20</td>
<td>1.640</td>
<td>1.04 – 2.63</td>
<td>62.1%</td>
</tr>
</tbody>
</table>

**that individual’s concept will have a higher fixation score category relative to the contiguous lower fixation score category

Figure 11 shows curves that estimate the probability of a concept by an individual with a specific NFC_d of achieving a particular fixation score category for the obtained participant NFC_d range of 107 - 196. Three fixation-score-category probability curves represent the probabilities of a concept achieving each respective fixation score category outcome. By definition, for each NFC_d, the total of all probabilities sums to 1.

To illustrate the probability increase for a more fixated concept, Figure 11 shows that for an individual with NFC_d = 140, the probability of a concept by that individual to have High Fixation rather than Medium Fixation is approximately 50% (Probability of Medium Fixation = 0.285, High Fixation = 0.29). If the NFC_d score is increased by 20, the probability of a concept by that individual with High Fixation rather than Medium Fixation is now approximately 58% (Probability of Medium Fixation = 0.287, High Fixation = 0.40).

The curve for Low Fixation shows that it is more probable for a lower NFC_d individual’s concept to obtain fixation scores of 2-4. Conversely, the curve for High Fixation shows that it is more probable for a higher NFC_d individual’s concept to attain fixation scores of 8-10.
Figure 11: Experiment 1 probability distribution plots for fixation scores by NFC_d

Note: Low Fixation (Fixation Scores = 2-4), Medium Fixation (Fixation Scores = 5-7), High Fixation (Fixation Score = 8-10)

3.5 Discussion: effect of NFC_d on design fixation

In the first statistical analysis, participants were divided into two groups, low NFC_d and high NFC_d based on a mean split of their NFC_d scores. A Wilcoxon Rank-Sum test confirmed significant differences between the mean fixation scores of participants’ concepts in the low versus high NFC_d groups. Specifically, the high NFC_d group had a higher mean fixation score than the low NFC_d group.

In the second statistical analysis, ordinal logistic regression was used to create a model of participants’ NFC_d and the degree of fixation of their concepts. This model showed that an individual’s tendency to fixate on the provided example for the concept generation problem increased with increased NFC_d. Specifically, an increase by 20 in an individual’s NFC_d score is
associated with an odds ratio of 1.64 or 62.1% probability that a concept by that individual to be rated a higher fixation score category relative to the contiguous lower fixation score category.

The results support my hypothesis of a relationship between NFC$_d$ and design fixation, and relate to differences in design fixation observed by other researchers. For example, Purcell and Gero (1996) asked mechanical engineering and industrial design students to develop concepts that assist elderly people in and out of a bathtub, and provided an Autofit fixating example. Of the Autofit’s 12 detail features, the mechanical engineering students fixated on 8, while the industrial design group fixated on 3. Purcell and Gero (1996) offered a potential explanation as “differences in educational processes in the two disciplines” where industrial design “emphasizes creativity and difference”. This difference may correspond to educations that attract and/or encourage higher NFC$_d$ in mechanical engineers and lower NFC$_d$ in industrial designers.

Toh et al. (2014) inversely related fixation with the number of parts examined during product dissection, and noted that extraverts examined more parts. Combined with the results, there may also be a relationship between NFC$_d$ and introversion versus extraversion.

Overall, the results were sufficiently promising towards identifying a link between NFC$_d$ and design fixation. This warranted further studies in NFC, particularly in NFC$_e$. 
4 Pilot Experiment: NFC_e and tendency for design fixation

Experiment 1’s results showed that NFC_d may be potentially used to predict an individual’s tendency for design fixation. In the study, participants’ NFC_d scores were significantly correlated with their concepts’ design fixation scores in an intuitive manner – higher NFC_d scores corresponded to higher fixation scores, and vice versa.

I wanted to study the effects of NFC_e in a pilot experiment. The rationale behind this was that NFC_e manipulation could be a potential mitigation tool for fixation.

4.1 Participants

Following approval by the University of Toronto Research Ethics Board, 27 participants (19 male and 8 female) were recruited from a fourth-year undergraduate engineering design course. These participants were students in mechanical engineering, industrial engineering, or engineering science.

Of the 27 participants, 24 had participated in Experiment 1. The 3 new participants (2 male and 1 female) were absent from the previous experiment but received identical instruction on design fixation in lecture.

4.2 Method

4.2.1 Formation of groups prior to experiment

In the pilot experiment, I organized the participants into two groups – Taboo-round1 and Taboo-round2. The groups were similar in size and mean NFC_d score (measured during the previous experiment using the 42-item NFC_d Scale) to better detect possible effects of NFC_e. The Taboo-round1 group had 14 members (10 male and 4 female) and a mean NFC_d score of 154.9 (42-item
NFC_d Scale, for 13 previously tested members). The Taboo-round2 group had 13 members (9 male and 4 female), with a mean NFC_d score of 149.9 (42-item NFC_d Scale, for 11 previously tested members). A small difference in mean NFC_d score remained between the two groups because some students were unexpectedly absent from the scheduled course laboratory during which the experiment was conducted.

4.2.2 Game of Taboo

I used the group-based game of Taboo in the pilot experiment to induce NFC_e. In Taboo, a player is given a card with a target word and five forbidden words, and within a time limit, must give clues (verbal, visual, etc.) without using any of the forbidden words and target word, to other players who then attempt to guess the target word. To reduce potential ignorance of the Taboo target words, I developed Taboo cards that used material from the undergraduate engineering design course from which the participants were recruited.

4.2.3 Experimental conditions and relation to NFC_e

The pilot experiment had two conditions – intervention and control – that both groups underwent. In the control condition, participants were given written instructions to individually, without consulting their group, generate as many concepts as possible with each concept on a separate sheet of paper. The group in the intervention condition had the same written instructions with the addition of the Taboo game. The participants in the intervention condition were allowed to communicate only for the purposes of organizing strategies for the Taboo game. They were informed that they could submit concepts at any point to obtain a Taboo card for the group. In practice, a group/game leader solicited completed concepts from their group in return for a Taboo card. Concepts from the group could be submitted in quantities of 1-6 at a time, with more simultaneously submitted concepts earning more time during the corresponding Taboo game round.
(see Table 7). Following concept submission, the group leader would have 15 seconds to first inspect the Taboo card before using the allotted time to give clues to the group. If the group guessed the target word within the allotted time and rules, they would receive 1 point and 0 otherwise. Overall, the group with more points would win.

From a NFC viewpoint, I anticipated that the Taboo game would be engaging to the participants because of the novelty and competitive nature of the activity. However, the time pressure to constantly submit concepts to keep the game going would potentially induce heightened NFC in the intervention group.

Table 7: Allotted Taboo game-round time per number of submitted concepts

<table>
<thead>
<tr>
<th># Concepts submitted at same time</th>
<th>Allotted time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 concept</td>
<td>10 seconds</td>
</tr>
<tr>
<td>2 concepts</td>
<td>12 seconds</td>
</tr>
<tr>
<td>3 concepts</td>
<td>16 seconds</td>
</tr>
<tr>
<td>4 concepts</td>
<td>20 seconds</td>
</tr>
<tr>
<td>5 concepts</td>
<td>25 seconds</td>
</tr>
<tr>
<td>6 concepts (MAX)</td>
<td>30 seconds</td>
</tr>
</tbody>
</table>

4.2.4 Implementation

The experiment took place in the last hour of a regularly scheduled three-hour laboratory in the course previously described. The laboratory room had a black curtain that could be used to divide the room in half, allowing for visual but not audio separation.
I first instructed all participants on how to play the game of Taboo through an example. The participants were then given individualized sheets containing their group number and secret identifier code to enable anonymity. I asked participants to separate into the two previously organized groups on opposite sides of the laboratory room. The room’s curtain system was then drawn to visually separate the two groups. I then handed participants written instructions and blank concept submission sheets.

The pilot experiment consists of three parts: Participants 1) generated concepts for a peanut-sheller problem, 2) generated concepts for an alarm clock problem, and 3) completed the 15-item NFC\(_d\) Scale (Roets & Van Hiel, 2011).

Parts 1 and 2 took 20-minutes each. In part 1, Taboo-round1 had the intervention condition and Taboo-round2 had control condition, while in part 2, Taboo-round2 had the intervention condition and Taboo-round1 had the control condition. I had to use different design problems than Experiment 1 because 24 of the 27 participants were already aware of the problem from completing Experiment 1 and the experiment design required two problems.

### 4.2.5 Part 1: Round 1 - Peanut-sheller problem

The peanut-sheller problem (Linsey, et al., 2010) (Viswanathan & Linsey, 2013) is shown in **Figure 12**.
Problem Description:

In places like Haiti and certain West African countries, peanuts are a significant crop. Most peanut farmers shell their peanuts by hand, an inefficient and labor-intensive process. The goal of this project is to design and build a low-cost, easy to manufacture peanut shelling machine that will increase the productivity of the African peanut farmers. The target throughput is approximately 50 kg (110 lbs) per hour.

Customer Needs:

- Must remove the shell with minimal damage to the peanuts.
- Electrical outlets are not available as a power source.
- A large quantity of peanuts must be quickly shelled.
- Low cost.
- Easy to manufacture.

Example: This system uses a gas-powered press to crush the peanut’s shell. The shell and the peanut then fall into a collection bin.

![Diagram of peanut shelling machine](image)

Figure 12: Problem 1 – Device to shell peanuts (Linsey et al., (2010) (2013))

4.2.6 Part 2: Round 2 - Alarm-clock problem

The alarm-clock problem was adapted from Genco, et al. (2012) and shown in Figure 13. The problem originally used physical examples and I provided pictorial representations based on the
authors’ descriptions of the clocks: "The first alarm clock had a beeping alarm, a digital lit display, buttons for input, a back-up battery in addition to an electric plug, and a snooze button. The second alarm clock was a simple analog clock, with a beeping alarm, a wall plug, dials for input and no snooze feature. Neither clock had the capability of playing music or performing any other functions." (Genco, et al., 2012).

**Problem Description:** Assume that you are working for a design company. A client has asked your design company to design a next-generation alarm clock.

**Examples:**

<table>
<thead>
<tr>
<th>This alarm clock has a beeping alarm, a digital lit display, buttons for input, a back-up battery in addition to an electric plug, and a snooze button.</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Digital Lit Display" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>This analog alarm clock has a beeping alarm, a wall plug, dials (on reverse of clock) for input, but does not have snooze feature.</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image2" alt="Dials on Reverse of Clock" /></td>
</tr>
</tbody>
</table>

**Figure 13: Problem 2 - Alarm clock (Adapted, (Genco, et al., 2012))**
4.2.7 Part 3: 15-item NFC$_d$ Scale

After the two rounds of concept generation, all participants completed the 15-item NFC$_d$. I chose the 15-item NFC$_d$ Scale because of time constraints, and used it to exclusively determine the NFC$_d$ score for those who did not participate in Experiment 1.

I compared NFC$_d$ scores measured using the 15-item NFC$_d$ Scale and the 42-item NFC$_d$ Scale for the 24 participants who took part in both Experiment 1 and the pilot (see Figure 14). I multiplied the scores from the 15-item NFC$_d$ Scale by 2.8 to facilitate comparison with the 42-item NFC$_d$ Scale scores. The Pearson product-moment correlation coefficient, which measures the strength of linear association from -1 to +1, was calculated to be $r = 0.55$. I concluded that since $r = 0.55$ is a large strength of association, the measure of NFC$_d$ using the 15-item NFC$_d$ Scale was valid.

![Figure 14: 15-item NFC$_d$ Scale vs 42-item NFC$_d$ Scale (24 observations)](image-url)
4.2.8 Debrief

After the experiment, participants were debriefed and provided information letters requesting permission to use their data. All participants consented to the use of their data in the pilot experiment.

4.3 Results

4.3.1 Fixation evaluation

I measured fixation relative to the provided examples using evaluation criteria developed by the problems’ respective authors. For the peanut-sheller problem, I used evaluation criteria developed by Linsey et al. (2010) (2013). For the alarm-clock problem, I used the conformity score metric developed by Genco et al. (2012). In both evaluation schemes, concepts received a score of 1 or 0 on each criterion, with 1 indicating a feature shared with the example solution, and 0 indicating a differing feature. Table 8 shows the peanut-sheller fixation-rating scheme, and Table 9 shows the alarm-clock fixation-rating scheme. I refer hereafter to the total scores obtained from using these evaluation schemes as fixation scores.

Table 8: Peanut-sheller fixation rating scheme (Linsey et al., (2010) (2013)).

<table>
<thead>
<tr>
<th>Function</th>
<th>Solution from Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(Material)</strong></td>
<td></td>
</tr>
<tr>
<td>Guide</td>
<td>Sloped surface, conveyor</td>
</tr>
<tr>
<td>Import</td>
<td>Hopper</td>
</tr>
<tr>
<td>Position</td>
<td>Table legs</td>
</tr>
<tr>
<td>Remove (Shell)</td>
<td>Crushing plates</td>
</tr>
<tr>
<td><strong>(Energy)</strong></td>
<td></td>
</tr>
<tr>
<td>Convert</td>
<td>Gas press</td>
</tr>
<tr>
<td>Store</td>
<td>Bin</td>
</tr>
<tr>
<td>Separate (nut and shell)</td>
<td>Grate</td>
</tr>
</tbody>
</table>

Table 9: Alarm clock fixation rating scheme (Genco et al., (2012)).

<table>
<thead>
<tr>
<th>Function</th>
<th>Solution from Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(System)</strong></td>
<td></td>
</tr>
<tr>
<td>Mode</td>
<td></td>
</tr>
<tr>
<td><strong>(Unit)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>(Function)</strong></td>
<td></td>
</tr>
<tr>
<td>Convert</td>
<td></td>
</tr>
<tr>
<td><strong>(Control)</strong></td>
<td></td>
</tr>
<tr>
<td>Position</td>
<td></td>
</tr>
<tr>
<td>Position</td>
<td></td>
</tr>
</tbody>
</table>
Table 9: Alarm-clock fixation rating scheme (Genco et al., (2012))

<table>
<thead>
<tr>
<th>Basic Features:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode of Alarm</td>
</tr>
<tr>
<td>Display Type</td>
</tr>
<tr>
<td>Information Shown</td>
</tr>
<tr>
<td>Mode of Input</td>
</tr>
<tr>
<td>Energy Source</td>
</tr>
<tr>
<td>Snooze</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Additional Features:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Music</td>
</tr>
<tr>
<td>Alternative Use</td>
</tr>
<tr>
<td>Shape/Layout</td>
</tr>
</tbody>
</table>

While evaluating the concepts, I found that some criteria from Table 8 and Table 9 were difficult to rate because the concepts generated did not explicitly address these criteria. In the absence of differentiating information, the concepts were rated as fixated with respect to these criteria, i.e., sharing the corresponding features. Table 10 and Table 11 show the number of uncertain cases for each criterion. I calculated two fixation scores for each problem: a full score including all original criteria, and an adjusted score omitting criteria with high (over 70%) uncertainty. In addition, the peanut-sheller “guide” criterion was removed from the adjusted fixation score, because whether a concept included a “sloped surface” was difficult to assess, as participants’ concept sketches usually displayed a single perspective.

A second rater was asked to rate 20 randomly selected peanut-sheller and alarm-clock concepts using the respective rating schemes. Inter-rater reliability tests compared the second rater’s fixation scores to my scores, omitting all uncertain criteria between the two ratings. For the 20 peanut-sheller concepts, inter-rater reliability was Krippendorff’s $\alpha$ (ordinal) = 0.937 and for the 20 alarm-clock concepts, was Krippendorff’s $\alpha$ (ordinal) = 0.651. Krippendorff (2004) states that there are no “magical numbers” to assess an acceptable level of agreement but suggests $\alpha \geq .800$ or $\alpha \geq .667$, respectively.
“where tentative conclusions are still acceptable”. The obtained $\alpha = 0.651$ in the alarm-clock concepts is marginally close to the suggested $\alpha \geq 0.667$ cut-off.

Table 10: Peanut-sheller problem - uncertain cases

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Full (PS7)</th>
<th>Adjusted (PS4)</th>
<th>Uncertain Cases (# concepts unclear on criterion/all concepts)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guide</td>
<td>X</td>
<td></td>
<td>54/106 (50.9%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>*unclear what constitutes “sloped surface” (definition and also from concepts)</td>
</tr>
<tr>
<td>Import</td>
<td>X</td>
<td></td>
<td>78/106 (75.6%)</td>
</tr>
<tr>
<td>Position</td>
<td>X</td>
<td></td>
<td>89/106 (84.0%)</td>
</tr>
<tr>
<td>Remove (Shell)</td>
<td>X X</td>
<td></td>
<td>1/106</td>
</tr>
<tr>
<td>Store</td>
<td>X X</td>
<td></td>
<td>54/106</td>
</tr>
<tr>
<td>Separate (Nut &amp; Shell)</td>
<td>X X</td>
<td></td>
<td>52/106</td>
</tr>
<tr>
<td>Convert</td>
<td>X X</td>
<td></td>
<td>42/106</td>
</tr>
</tbody>
</table>

Table 11: Alarm-clock problem - uncertain cases

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Full (AC9)</th>
<th>Adjusted (AC7)</th>
<th>Uncertain Cases (# concepts unclear on criterion/all concepts)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode of Alarm</td>
<td>X X</td>
<td></td>
<td>39/135</td>
</tr>
<tr>
<td>Display Type</td>
<td>X X</td>
<td></td>
<td>49/135</td>
</tr>
<tr>
<td>Information Shown</td>
<td>X X</td>
<td></td>
<td>54/135</td>
</tr>
<tr>
<td>Mode of Input</td>
<td>X X</td>
<td></td>
<td>119/135 (88.1%)</td>
</tr>
<tr>
<td>Energy Source</td>
<td>X</td>
<td></td>
<td>109/135 (80.7%)</td>
</tr>
<tr>
<td>Snooze</td>
<td>X X</td>
<td></td>
<td>0/135</td>
</tr>
<tr>
<td>Music</td>
<td>X X</td>
<td></td>
<td>0/135</td>
</tr>
<tr>
<td>Alternative Use</td>
<td>X X</td>
<td></td>
<td>0/135</td>
</tr>
<tr>
<td>Shape/Layout</td>
<td>X X</td>
<td></td>
<td>10/135</td>
</tr>
</tbody>
</table>
4.3.2 Observations

Figure 15 shows a histogram of participants’ NFC\textsubscript{d} scores as measure by the 15-item scale. The range was 98 – 198.8, the mean was 154.62, and the standard deviation was 25.89.

![Histogram of Participant's NFCd Scores](image)

**Figure 15**: Histogram of NFC\textsubscript{d} scores of participants

Figure 16, Figure 17, Figure 18, and Figure 19 plot the mean, full and adjusted, mean fixation scores of each participants’ peanut-sheller and alarm-clock concepts as a function of NFC\textsubscript{d} for both experimental groups.

For the peanut-sheller problem, it appeared that there were no differences in the full or adjusted mean fixation scores between the two groups for. For the alarm-clock problem, the full and adjusted mean fixation scores appeared to be higher for the intervention group compared to the control group.
Figure 16: Mean fixation scores of peanut-sheller concepts (Full – PS7)

Figure 17: Mean fixation scores of peanut-sheller concepts (Adjusted – PS4)
Figure 18: Mean fixation scores of alarm-clock concepts (Full – AC9)

Figure 19: Mean fixation scores of alarm-clock concepts (Adjusted – AC7)
4.3.3 Condition differences

Table 12 shows the differences between the 2 groups for the two problems. The Mann-Whitney U non-parametric statistical test was used to compare (ordinal) fixation scores of concepts by participants of the intervention versus control groups.

For the peanut-sheller problem, the Mann-Whitney U test did not find significant differences in the mean fixation scores of peanut-sheller concepts between participants in the intervention and control groups (Full: Mann–Whitney $U = 81.5$, $Z = -0.44$, $m_{Taboo-round2\text{group(control)}} = 5.20$, $m_{Taboo-round1\text{group(intervention)}} = 5.10$, $n_1 = 13$, $n_2 = 14$, $p = 0.66$, $r = 0.016$, two-tailed; Adjusted: Mann–Whitney $U = 90.5$, $Z = 0$, $m_{Taboo-round2\text{group(control)}} = 2.59$, $m_{Taboo-round1\text{group(intervention)}} = 2.58$, $n_1 = 13$, $n_2 = 14$, $p = 1$, $r = 0$, two-tailed). Effect sizes of $r = 0.016$ and $r = 0$ were below the $r = 0.10$ cutoff for a small effect.

For the alarm-clock problem however, the Mann-Whitney U test revealed marginally significant differences in the mean fixation scores of alarm-clock concepts between participants in the intervention and control groups (Full: Mann–Whitney $U = 53.5$, $Z = -1.79$, $m_{Taboo-round2\text{group(intervention)}} = 6.95$, $m_{Taboo-round1\text{group(control)}} = 6.49$, $n_1 = 13$, $n_2 = 14$, $p = 0.07$, $r = 0.34$, two-tailed; Adjusted: Mann–Whitney $U = 55$, $Z = -1.72$, $m_{Taboo-round2\text{group(intervention)}} = 5.20$, $m_{Taboo-round1\text{group(control)}} = 4.65$, $n_1 = 13$, $n_2 = 14$, $p = 0.085$, $r = 0.33$, two-tailed). Effect sizes of $r = 0.34$ and $r = 0.33$ corresponded to medium effects.
Table 12: Intervention vs. control condition differences

<table>
<thead>
<tr>
<th></th>
<th>Taboo-round2 group</th>
<th>Taboo-round1 group</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Mean NFC(_d) Score (NFC(_d) C 42)</td>
<td>149.9 (missing 2 values)</td>
<td>154.9 (missing 1 value)</td>
</tr>
<tr>
<td>Mean NFC(_d) Score (NFC(_d) 15)</td>
<td>153.2 (missing 2 values)</td>
<td>154.0 (missing 1 value)</td>
</tr>
<tr>
<td>Mean NFC(_d) Score (NFC(_d) 15)</td>
<td>151.8</td>
<td>157.2</td>
</tr>
<tr>
<td>Total Number of Peanut-Sheller Concepts</td>
<td>55 (no Taboo)</td>
<td>51 (Taboo)</td>
</tr>
<tr>
<td>Mean Fixation Scores – Full (PS7)</td>
<td>5.20 (no Taboo)</td>
<td>5.10 (Taboo)</td>
</tr>
<tr>
<td>Mean Fixation Scores – Adjusted (PS4)</td>
<td>2.59 (no Taboo)</td>
<td>2.58 (Taboo)</td>
</tr>
<tr>
<td>Total Number of Alarm-Clock Concepts</td>
<td>48 (Taboo)</td>
<td>87 (no Taboo)</td>
</tr>
<tr>
<td>Mean Fixation Scores – Full (AC9)</td>
<td>6.95 (Taboo)</td>
<td>6.49 (no Taboo)</td>
</tr>
<tr>
<td>Mean Fixation Scores – Adjusted (AC7)</td>
<td>5.20 (Taboo)</td>
<td>4.65 (no Taboo)</td>
</tr>
</tbody>
</table>

4.3.4 Comparing individual differences

Similarly to Experiment 1, I wanted to compare whether there was any relationship between participants’ NFC\(_d\) scores and the fixation scores of participants’. The 15-item NFC\(_d\) Scale used in the pilot experiment would only measure NFC\(_d\) and not NFC\(_e\). Because of this, I anticipated that an ordinal logistic regression analysis would result in no significant log odds ratios because NFC\(_e\) would be absent in the model.

Table 13 shows the log odds ratios obtained from each of the models using NFC\(_d\) score as a predictor for the four different fixation scoring metrics: peanut-sheller full (PS7), peanut-sheller adjusted (PS4), alarm-clock full (AC9), and alarm-clock adjusted (AC7) (one model per scoring metric).
Surprisingly, the log odds ratio obtained using the adjusted fixation score for the peanut-sheller problem yielded a significant result. I believe this is because the peanut-sheller problem’s round was executed ineffectively, further discussed in 4.4 Discussion.

Table 13: Log odds ratio for increase by one in NFC\textsubscript{d} Score on fixation score

<table>
<thead>
<tr>
<th>Problem</th>
<th>Rating Scheme Used</th>
<th>Log Odds Ratio Est.</th>
<th>95% Confidence Interval</th>
<th>Wald Test (Type 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peanut Sheller</td>
<td>PS7 – Full</td>
<td>0.0073</td>
<td>-0.0079 - 0.0226</td>
<td>$\chi^2(1) = 0.89, p = 0.35$</td>
</tr>
<tr>
<td></td>
<td>PS4 - Adjusted</td>
<td>0.0175</td>
<td>0.0016 - 0.0335</td>
<td>$\chi^2(1) = 4.67, p = 0.03^*$</td>
</tr>
<tr>
<td>Alarm Clock</td>
<td>AC9 - Full</td>
<td>0.0064</td>
<td>-0.0080 - 0.0208</td>
<td>$\chi^2(1) = 0.76, p = 0.38$</td>
</tr>
<tr>
<td></td>
<td>AC7 - Adjusted</td>
<td>0.0074</td>
<td>-0.0063 - 0.0211</td>
<td>$\chi^2(1) =1.12, p = 0.29$</td>
</tr>
</tbody>
</table>

*$p < .05$

4.4 Discussion: effect of NFC\textsubscript{e} on design fixation

The primary focus of the pilot experiment was to investigate research question 2) Is an individual’s tendency to fixate during concept generation affected by NFC\textsubscript{e}?

In round 1 (peanut-sheller problem), a Mann-Whitney U test did not find any significant differences in the mean fixation scores of concepts between the control and intervention groups. However, in round 2 (alarm-clock problem), there was a marginally significant difference.

I believe that in round 1, the manipulation of NFC\textsubscript{e} was not effectively executed (peanut-sheller problem). An absence of the anticipated NFC\textsubscript{e} effect would remove differences between the intervention and control conditions, and participants would perform according to their NFC\textsubscript{d} scores. Because the two groups had similar mean NFC\textsubscript{d} scores, there would be no anticipated
differences in mean fixation scores of the concepts generated. Evidence towards this is the fact that ordinal logistic regression of participants adjusted peanut-sheller fixation scores compared to their NFCₐ yielded a significant log odds ratio (similar to in Experiment 1) while the ordinal logistic regressions for the alarm-clock fixation scores did not.

The difference in the two conditions between rounds – intervention and control – occurred due to the following reasons: differences in explanation of the experimental procedure, differences in Taboo game engagement between groups, and learning effect. Although unplanned, there were differences in how the participants perceived the experimental procedure between the two rounds.

In round 1, the Taboo-round1 group played the Taboo game but they were not told that the Taboo-round2 group would not be playing the Taboo game in this round, or that there would even be a second round at all. If the Taboo-round1 group believed that the Taboo-round2 group was playing the Taboo game concurrently, the fact that Taboo-round2 group was instructed to be completely silent throughout round 1 would have had an effect. In fact, this may have manifested in Taboo-round1 group’s relaxed attitude while playing Taboo, where no concept submissions occurred until the halfway point. In effect, my failure to properly communicate the experimental protocol may have eliminated any sort of time pressure or urgency in the Taboo-round1 group, reducing the planned NFCₑ effect. Furthermore, the Taboo-round2 group in round 1 was only given the control instructions and may have spent some effort following the Taboo-round1 group, whom they could hear. This distracting effect may have increased their NFCₑ as well, further reducing differences in NFCₑ between the two groups.

In round 1, the Taboo-round1 group played the Taboo game 9 times, answering correctly 8 times, with an average of 5.67 concepts per submission. In round 2, the Taboo-round2 group played the Taboo game 23 times, answering correctly 13 times, with an average of 2.09 concepts per
submission. Taboo-round2 group also had the advantage of playing the game second, after having heard the first group play, and was substantially different in their engagement in the Taboo game activity. Additionally, they fully realized the competitive aspect of the experiment and emphasized the need to win this competition, by generating concepts to submit in order to be able to play Taboo games and earn points. Starting immediately at the beginning of the round, the Taboo-round2 group appointed one member of their group to be the almost exclusive Taboo game leader. There was a more consistent solicitation for more concepts by the group in general. This frantic system resulted in heightened time pressure in the members of the Taboo-round2 group to complete concepts for submission. The heightened time pressure was much more apparent in round 2 as evidenced by the more frequent submissions compared with Taboo-round1 group’s less frequent submissions in round 1. Given this, I believe that round 2 was a more successful manipulation of NFC_e in the Taboo-round1 group through the Taboo game.

While this pilot experiment did not provide definitive evidence towards the notion that NFC_e affects individual tendency for fixation, the marginally significant results for the alarm-clock problem and evidence towards an improper execution of the peanut-sheller problem warranted an additional experiment. In addition, the pilot experiment provided many learning experiences in the manipulation of NFC_e such as not having intervention and control conditions in the same environment, the use of the 15-item NFC_d Scale as adding another factor of uncertainty, the use of Taboo as perhaps too elaborate of a way to induce time pressure, and the need for another piece of information in order to estimate NFC_e.
5 Experiment 2: NFC\textsubscript{d} and NFC\textsubscript{e} and tendency for design fixation

Experiment 1 showed that individual differences in NFC\textsubscript{d} could be used to predict an individual’s tendency for design fixation. I found that participants’ NFC\textsubscript{d} scores were significantly correlated with their concepts’ design fixation scores, where higher NFC\textsubscript{d} scores corresponded to higher fixation scores, and vice versa.

The pilot experiment attempted to test whether an individual’s tendency to fixate during concept generation is affected by NFC\textsubscript{e} and produced mixed results. The pilot experiment provided many lessons to ensure that a subsequent experiment would be completed effectively.

Building on Experiment 1 and the pilot experiment, I conducted Experiment 2 to study the effect of NFC\textsubscript{e} on an individual’s tendency for design fixation. I hypothesized that the manipulation of NFC\textsubscript{e} would induce similar tendencies for design fixation as observed in Experiment 1. If shown to be true, strategies to overcome fixation could be tailored for individuals.

5.1 Participants

Sixty-one participants were recruited from Amazon Mechanical Turk. Of the 61 participants, 26 were male and 35 were female. The mean age of 60 participants was 32.3 (range 18 – 67, SD = 10.12). One participant incorrectly entered their age as “1” but submitted valid entries for all other questions.

5.2 Method

The experiment consisted of Mechanical Turk workers completing a three-part task: 1) Generating 3 (or more) solution concepts for a design problem, 2) Completing the NFC\textsubscript{d} Scale (without being shown their score initially), 3) Completing an exit survey with questions on their age, gender, and estimation of their NFC score. Upon completion of all three parts, workers were given their measured NFC\textsubscript{d} score as well as a unique confirmation code to enter in their HIT assignment. Each worker’s submission was reviewed and if approved, he or she would be paid and be counted as a
participant in this study. A timestamp log of each worker was obtained for the reviewing process (see Appendix 4).

5.2.1 Human Intelligence Task (HIT) set-up and recruitment

The HIT had the title “Drawing Activity and Questionnaire” and description “Use your mouse to draw at least 3 concepts for a design activity on a canvas on your browser and answer a few questions. Should take less than an hour to complete. Approval will be processed within 48 hours.” The HIT had the following keywords: survey, design, concept, drawing. Each HIT assignment paid approximately $1.60 USD. I did not require “Master Worker” status or any other additional qualifications.

5.2.2 Experiment registration and group assignment

Once Mechanical Turk workers were directed to the website, they were greeted with preliminary instructions, including statements stipulated by the University of Toronto Research Ethics Board (e.g. “there are no known or anticipated risks associated to participation in this study”) as well as instructions for registering on the website. To prevent multiple submissions from a single worker, I processed registrations using their Mechanical Turk unique worker IDs. Once workers were registered, they would log-in and be taken to their profile page. The three task modules were displayed on the profile page and each subsequent module would be “locked” until the previous one was completed. Workers would start on Part 1 (See Appendix 2).

Our experiment consisted of three experimental groups: Group 1: low NFC_e, Group 2: control, and Group 3, high NFC_e. Mechanical Turk workers were automatically assigned to the group with the fewest completed submissions. If all three groups had the same number of completed submissions, the system would assign the next registered worker to a group randomly.

5.2.3 Part 1: Design problem

I used the same design problem as in Experiment 1 (see Figure 1 and Figure 2). Workers viewed the problem on a separate webpage, and were not allowed to return to the design problem page upon exiting it. Once they were ready, and familiarized themselves on the use of the canvas tool, they were instructed to continue to concept generation.
5.2.4 Part 1: Preparation for concept generation

Workers used the same canvas of 1000 by 600 pixels as in Experiment 1 for drawings and text annotations. Before beginning with concept generation, workers were directed to a webpage entitled “Drawing Tool Familiarization.” Here, written instructions asked workers to: “draw a green circle on the canvas”, “type ‘I understand the textbox function.’ in red”, and “click the Submit button and follow the instructions on the popup to continue to the next part”.

5.2.5 Part 1: Manipulation of NFC\textsubscript{e} during concept generation

I chose to manipulate NFC\textsubscript{e} using time pressure during concept generation, as it has been shown to heighten or increase NFC (Kruglanski & Freund, 1983), and it is relatively straightforward to implement. I applied time pressure directly using a timer instead of indirectly (such as through the Taboo game in the pilot experiment).

Table 14 summarizes the three experimental groups, with differences in instruction only on the concept generation webpage. In Group 1, low NFC\textsubscript{e}, workers were allotted 30 minutes per concept while in Group 3, high NFC\textsubscript{e}, workers were allotted 5 minutes. For both Groups 1 and 3, the page would refresh after the timer expired, but Group 3 received additional instruction that this refresh would be documented to discourage them from circumventing the timer (all workers were instructed not to refresh the webpage) (see Appendix 3).

Once workers had submitted at least three concepts, they were given a link to continue to Part 2 of the experiment to complete the NFC\textsubscript{d} Scale.

5.2.6 Part 2: NFC\textsubscript{d} Scale

The NFC\textsubscript{d} Scale was administered in an identical manner to Experiment 1 (see Appendix 1).

5.2.7 Part 3: Exit survey

The exit survey consisted of 6 questions. The first question asked for the worker’s familiarity with the shower problem using a 10-point Likert scale from 1 (not at all familiar) to 10 (very familiar). The second question asked with which shower models the worker was familiar. The next three questions asked the worker’s gender, age and personal estimation of their NFC score. The final,
optional question provided a textbox for workers to submit any comments, suggestions, and concerns.

Table 14: NFCe manipulation during concept generation (experiment 2)

<table>
<thead>
<tr>
<th>Group</th>
<th>Group 1: Low NFC&lt;sub&gt;e&lt;/sub&gt;</th>
<th>Group 2: Control</th>
<th>Group 3: High NFC&lt;sub&gt;e&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time allotted per concept</td>
<td>30 minutes</td>
<td>N/A</td>
<td>5 minutes</td>
</tr>
<tr>
<td>Special instructions in color shown to workers</td>
<td>“You have 30 minutes to draw and submit each concept. The amount of time remaining is indicated by the floating timer. Once time has expired, the page will reload and all current canvas progress will be lost.”</td>
<td>N/A</td>
<td>“You have 5 minutes to draw and submit each concept. The amount of time remaining is indicated by the floating timer. Once time has expired, the page will reload and all current canvas progress will be lost and this will be noted.”</td>
</tr>
<tr>
<td>Timer example in color shown to workers</td>
<td><img src="image" alt="Green Timer Example" /></td>
<td>N/A</td>
<td><img src="image" alt="Red Timer Example" /></td>
</tr>
</tbody>
</table>

(text in green in color)
5.2.8 Experiment debriefing

After completing the three parts of the experiment, workers were returned to their profile page where they were given both their measured $\text{NFC}_d$ score and their unique confirmation code to enter on Mechanical Turk. All submissions were reviewed within 48 hours of submission. If accepted, the worker was paid and became part of the participant pool. If rejected, the worker was notified of the reason, and was not paid or included in the participant pool. Approximately 38% of workers’ submissions were discarded due to reasons such as incompleteness, and random concept submissions (e.g. smiley faces or indiscernible scribbles). Participants, or workers with accepted submissions, spent an average of about half an hour to complete Experiment 2.

5.3 Evaluation of design fixation

The same Fixation Score Rating Scheme (Table 2) as Experiment 1 was used in Experiment 2. In total, the 61 participants generated 191 concepts, of which 120 were accepted. Concepts were rejected for the same reasons as for Experiment 1, i.e., for not following instructions provided in the problem description. Examples of rejected concepts include showers using aeration, and only information and/or automation as resource-conserving strategies. Experiment 2’s concepts used fewer annotations than Experiment 1’s concepts, which sometimes led to concepts that were not understandable, leading to their rejection. Participants sometimes submitted the same concept more than once, each of which were counted towards the 191 generated concepts.

A second rater (not the same rater from Experiment 1) was asked to rate a random subset of 30 of the accepted concepts. The second rater was provided with only: the design problem (Figure 1), solution examples (Figure 2), fixation score rating scheme (Table 2), the 30 concepts to be rated, and a sample concept rated by the first author with fixation score of 10. The inter-rater reliability of the fixation score on the subset of 30 concepts was calculated as Krippendorff’s $\alpha$ (ordinal) = 0.809 ($\alpha \geq 0.8$ corresponds to perfect agreement (Krippendorff, 2004)).

5.4 Results and discussion

I analyzed the relationship between participants’ $\text{NFC}_d$, $\text{NFC}_e$ and fixation scores at both group and individual levels. At the group level, I compared the three experimental groups with respect to their fixation scores using Wilcoxon Rank-Sum tests ($\text{NFC}_e$).
5.4.1 Observations

Figure 20 shows a histogram of participants’ NFC\textsubscript{d} scores. The range was 98 – 199, the mean was 160.63, and the standard deviation was 22.79.

![Histogram of NFC\textsubscript{d} of participants](image)

**Figure 20: Histogram of NFC\textsubscript{d} of participants**

Figure 21 and Figure 22 show plots of participants’ NFC\textsubscript{d} against 1) fixation scores for each accepted concept, and 2) mean fixation score of accepted concepts per participant.

A preliminary assessment of this data was performed using the Kendall Tau-b test. I obtained a significant Kendall’s rank correlation of $\tau_b = 0.16$ ($Z = 2.39, p = 0.017$) for Figure 21, which indicates a weak positive correlation between participants’ NFC\textsubscript{d} and their accepted-concepts’ fixation scores. Kendall’s rank correlation tau $\tau_b = 0.12$ ($Z = 1.31, p = 0.19$) for Figure 22 indicates a non-significant weak positive association between participants’ NFC\textsubscript{d} and their accepted-concepts’ mean fixation scores.
Figure 21: Fixation scores of accepted concepts vs. NFCd

Figure 22: Mean fixation scores of accepted concepts per participant vs NFCd
5.4.2 Experimental group differences

Figure 23, Figure 24, and Figure 25 show the histograms of participants’ NFC\textsubscript{d} scores in each group. Group 1, Group 2, and Group 3 all had similar mean NFC\textsubscript{d} of 162.11, 158.10, and 161.82 respectively.

![Histogram of NFC\textsubscript{d} of participants in Group 1: Low NFC\textsubscript{e}](image1)

Figure 23: Histogram of NFC\textsubscript{d} of participants in Group 1: Low NFC\textsubscript{e}

![Histogram of NFC\textsubscript{d} of participants in Group 2: Control](image2)

Figure 24: Histogram of NFC\textsubscript{d} of participants in Group 2: Control
Figure 25: Histogram of NFC_d of participants in Group 3: High NFC_e

Figure 26: Mean fixation scores of accepted concepts per participant vs NFC_d in Group1: Low NFC_e
Figure 27: Mean fixation scores of accepted concepts per participant vs NFC\(_d\) in Group 2: Control

Figure 28: Mean fixation scores of accepted concepts per participant vs NFC\(_d\) in Group 3: High NFC\(_e\)
Figure 26, Figure 27, and Figure 28 show the mean fixation scores of accepted concepts per participant for Group 1, Group 2, and Group 3 respectively. Group 1 and Group 2 had similar mean fixation scores of 4.57 and 4.59. Group 3 had a much higher mean fixation score at 6.33.

Table 15 outlines the differences between the three experimental groups. Interestingly, the participants in Group 3: High NFC\textsubscript{e} gave the highest mean NFC estimate. I had asked participants to estimate their NFC at the end of the experiment, after they had completed the NFC\textsubscript{d} Scale but before they were informed of their NFC\textsubscript{d} score (see Appendix 5). I believe that their estimation of NFC may have been affect by their NFC\textsubscript{e} condition because this estimation was requested right after they completed the design activity under their assigned condition. My belief is supported by the fact that the lowest mean estimation of NFC came from the low NFC\textsubscript{e} group, who were given 30 minutes to submit each concept. This was more than enough time, and likely more time than they expected, as the experiment description noted that it should take less than an hour in total to complete. Conversely, the high NFC\textsubscript{e} group had the highest estimate, possibly because they did not expect 5-minute time limits given the overall time estimate in the experiment description.

Table 15: General comparisons between experimental groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Group 1: Low NFC\textsubscript{e}</th>
<th>Group 2: Control</th>
<th>Group 3: High NFC\textsubscript{e}</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>18</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>Fixation Score (Mean)</td>
<td>4.57</td>
<td>4.579</td>
<td>6.33</td>
</tr>
<tr>
<td>Time spent per accepted concept (seconds) (Mean)</td>
<td>189.26</td>
<td>231.20</td>
<td>164.88</td>
</tr>
<tr>
<td>NFC\textsubscript{d} (Mean)</td>
<td>162.11</td>
<td>158.10</td>
<td>161.82</td>
</tr>
<tr>
<td>NFC estimated by participant (Mean)</td>
<td>127.89</td>
<td>147.10</td>
<td>159</td>
</tr>
</tbody>
</table>
The Wilcoxon Rank-Sum non-parametric statistical test was selected to compare mean (ordinal) fixation scores of participants’ concepts of the different groups.

**Table 16** shows the results of the Wilcoxon Rank-Sum tests. There were no significant differences in mean fixation scores between participants in the low NFC<sub>e</sub> and control groups (Wilcoxon \( W = 357.50 \), \( Z = -0.06 \), \( m_1 = 4.57 \), \( m_2 = 4.57 \), \( n_1 = 18 \), \( n_2 = 21 \), \( p = 0.96 \), \( r = 0.01 \), two-tailed). The effect size was below the \( r = 0.10 \) cutoff for a small effect. Therefore, no significant differences were observed between the low NFC<sub>e</sub> and control groups.

**Table 16: Wilcoxon Rank-Sum test results between experimental groups**

<table>
<thead>
<tr>
<th></th>
<th>Low NFC&lt;sub&gt;e&lt;/sub&gt; (n = 18)</th>
<th>Control (n = 21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Fixation Score:</td>
<td><strong>4.57</strong></td>
<td><strong>4.579</strong></td>
</tr>
<tr>
<td>Mean NFC&lt;sub&gt;d&lt;/sub&gt;:</td>
<td><strong>162.11</strong></td>
<td><strong>158.10</strong></td>
</tr>
<tr>
<td>Control (n = 21)</td>
<td>( W = 357.50 )</td>
<td></td>
</tr>
<tr>
<td>Mean Fixation Score:</td>
<td><strong>4.579</strong></td>
<td></td>
</tr>
<tr>
<td>Mean NFC&lt;sub&gt;d&lt;/sub&gt;:</td>
<td><strong>158.10</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( Z = -0.06 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( p = 0.96 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( r = 0.01 ) (two tailed)</td>
<td></td>
</tr>
<tr>
<td>High NFC&lt;sub&gt;e&lt;/sub&gt; (n = 22)</td>
<td>( W = 287.50 )</td>
<td>( W = 369.50 )</td>
</tr>
<tr>
<td>Mean Fixation Score:</td>
<td><strong>6.33</strong></td>
<td>( Z = -2.25 )</td>
</tr>
<tr>
<td>Mean NFC&lt;sub&gt;d&lt;/sub&gt;:</td>
<td><strong>161.82</strong></td>
<td>( p = 0.030* )</td>
</tr>
<tr>
<td></td>
<td>( p = 0.033* )</td>
<td>( r = 0.34 ) (two tailed)</td>
</tr>
<tr>
<td></td>
<td>( r = 0.35 ) (two tailed)</td>
<td></td>
</tr>
</tbody>
</table>

\*\( p < .05 \)

I did obtain significant results when comparing low NFC<sub>e</sub> versus high NFC<sub>e</sub> (Wilcoxon \( W = 287.50 \), \( Z = -2.21 \), \( m_1 = 4.57 \), \( m_2 = 6.33 \), \( n_1 = 18 \), \( n_2 = 22 \), \( p = 0.033 \), \( r = 0.35 \), two-tailed; and control versus high NFC<sub>e</sub> groups (Wilcoxon \( W = 369.50 \), \( Z = -2.25 \), \( m_1 = 4.57 \), \( m_2 = 6.33 \), \( n_1 = 21 \), \( n_2 = 22 \),
\[ p = 0.030, \ r = 0.34, \text{ two-tailed} \]. Effect sizes of \( r = 0.35 \) and \( r = 0.34 \) corresponded to medium effects.

The average time spent per accepted concept was below 5 minutes (high NFC_e group’s cutoff) for all three groups, however, participants in the high NFC_e group produced significantly more fixated concepts than those in the low NFC_e and control groups. Referring to Table 15, the three groups had similar mean NFC_d scores, such that their fixation scores should also be similar in the absence of different conditions. I found particularly interesting that the high NFC_e group had a higher mean self-estimated NFC compared to the low NFC_e and control groups.

There were no significant gender differences in the mean fixation scores within the three experimental groups (see Table 17). The largest difference was observed in the control group (Wilcoxon \( W = 125, Z = 1.03, m_{1(\text{female})} = 161.10, m_{2(\text{male})} = 155.36, n_1 = 10, n_2 = 11, p = 0.30, r = 0.22, \text{ two-tailed} \)) but this can also be attributed to the small difference in mean NFC_d scores. The effect size \( r = 0.22 \) corresponds to a medium effect.

**Table 17: Experiment 2 gender differences**

<table>
<thead>
<tr>
<th></th>
<th>Low NFC_e</th>
<th>Control</th>
<th>High NFC_e</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>6</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>Fixation Score (Mean, Males)</td>
<td>4.67</td>
<td>4.09</td>
<td>6.46</td>
</tr>
<tr>
<td>NFC_d (Mean, Males)</td>
<td>160.17</td>
<td>155.36</td>
<td>161.88</td>
</tr>
<tr>
<td>Females</td>
<td>12</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>Fixation Score (Mean, Females)</td>
<td>4.53</td>
<td>5.1</td>
<td>6.23</td>
</tr>
<tr>
<td>NFC_d (Mean, Females)</td>
<td>163.08</td>
<td>161.10</td>
<td>161.80</td>
</tr>
</tbody>
</table>
5.5 Discussion: effect of NFC\(_e\) on design fixation

In Experiment 2, I studied the effect of NFC\(_e\) on design fixation. Using the same design problem that solicited concepts to reduce shower water-run time, I split participants into three experimental groups: Group 1: low NFC\(_e\), Group 2: control, and Group 3: high NFC\(_e\). Low NFC\(_e\) participants were given 30 minutes to submit each concept, high NFC\(_e\) participants were given only 5 minutes, and control-group participants were not given a time limit.

In the statistical analysis, participants were divided into three groups, low NFC\(_e\), control, and high NFC\(_e\), corresponding with their experimental group condition. Participants in the high NFC\(_e\) group produced significantly more fixated concepts than those in the low NFC\(_e\) and control groups, and results were statistically significant in Wilcoxon Rank-Sum tests. The three groups had similar mean NFC\(_d\) scores as measured by the NFC\(_d\) Scale, but there were differences in their NFC estimates, where the high NFC\(_e\) group had a higher mean NFC estimate than the other two groups. I believe that their estimated NFC scores may have been impacted by the experimental condition they had just experienced.

Possible contributions to the difference between Experiments 1 and 2 include the following. I switched the order of the design activity and NFC\(_d\) Scale as to prevent the NFC\(_d\) Scale items from affecting participants’ behavior in Experiment 2, where I also had a larger sample size. Experiment 1 used engineering students while Experiment 2 used Mechanical Turk workers.
6 Conclusion and Next Steps

This thesis aimed to answer the following research questions:

1) Is an individual’s tendency to fixate during concept generation affected by NFC\(_d\)?

2) Is an individual’s tendency to fixate during concept generation affected by NFC\(_e\)?

Experiment 1 aimed to answer research question 1) and showed that individual differences in NFC\(_d\) may potentially be used to predict an individual’s tendency for design fixation. When participants were divided into two groups, low NFC\(_d\) and high NFC\(_d\) based on a mean split of their NFC\(_d\) scores, a Wilcoxon Rank-Sum test showed that the high NFC\(_d\) group had significantly higher mean fixation scores.

An ordinal logistic regression analysis revealed that higher NFC\(_d\) scores corresponded to higher probability of higher fixation scores, and vice versa. Specifically, an increase by 20 in an individual’s NFC\(_d\) score is associated with an odds ratio of 1.64 or 62.1% probability that a concept by that individual to be rated a higher fixation score category relative to the contiguous lower fixation score category. An example was that from NFC\(_d\) = 140 to NFC\(_d\) = 160, the probability that a concept by an individual is of High Fixation (Fixation Score = 8-10) rather than Medium Fixation (Fixation Score = 5-7) is increased from approximately 50% (Probability of Medium Fixation = 0.285, High Fixation = 0.29) to 58% (Probability of Medium Fixation = 0.287, High Fixation = 0.40).

The pilot experiment provided some initial evidence towards research question 2) in that NFC\(_e\) affects individual tendency for design fixation. It also provided lessons learned that were implemented in Experiment 2.

Experiment 2 studied the role of NFC\(_e\) in design fixation and aimed to answer research questions 2) by manipulating NFC\(_e\) through time pressure. Three experimental groups were formed (low NFC\(_e\), control, and high NFC\(_e\) groups). The concepts produced by the high NFC\(_e\) group were shown to be significantly more fixated than both the low NFC\(_e\) and control groups through Wilcoxon Rank-Sum tests. This result provided evidence towards answering research question 2), that individual tendency for fixation was affected by NFC\(_e\). Additional evidence was provided in the results of participant’s estimations of their NFC scores. They were asked to estimate their NFC
score in the exit survey, after completing the experimental condition concept generation task and NFC\textsubscript{d} scale and given a short description of NFC\textsubscript{d} but before they were shown their measure NFC score. The high NFC\textsubscript{e} group had the highest mean NFC score estimate which may be an indication that the estimates were affected by participants’ respective NFC\textsubscript{e} conditions.

This work supports the existence of a relationship between Need for Closure and design fixation. Specifically, NFC\textsubscript{d} can be used to predict individual tendency for design fixation and NFC\textsubscript{e} can affect individual tendency for design fixation.

Further studies are required to ensure the validity of these results. Other design fixation problems should be used to ensure that the effect of NFC\textsubscript{d} is repeatable. Additionally, more NFC\textsubscript{e} manipulations beyond time pressure should be explored to ensure that the observed NFC\textsubscript{e} effect is not exclusive to just time pressure.

Shortcuts to the NFC\textsubscript{d} questionnaire beyond Roets and Van Hiel’s (2011) 15-item version of the NFC\textsubscript{d} Scale could be developed, resulting in quick predictions of individuals’ tendencies to fixate, and applying interventions as required. Finally, manipulations known to affect NFC\textsubscript{e} could be used to develop novel interventions to complement existing interventions to reduce design fixation.
References


Appendices

Appendix 1: 42-item NFCᵋ Scale (as administered in Experiment 1 and Experiment 2)

Instructions:

Follow the provided instructions.

Once you have completed the Questionnaire, please click the Submit button at the bottom of the page.

Please ensure that you have answered ALL questions before submitting!

IMPORTANT: Scroll to the bottom of this page after you have pressed the Submit button! The instructions to continue will be shown on the bottom of this page!

Read each of the following statements and decide how much you agree with each according to your beliefs and experiences.

1. I think that having clear rules and order at work is essential for success.

2. Even after I've made up my mind about something, I am always eager to consider a different opinion.

3. I don't like situations that are uncertain.

4. I dislike questions which could be answered in many different ways.
5. I like to have friends who are unpredictable.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Moderately Disagree</th>
<th>Slightly Disagree</th>
<th>Slightly Agree</th>
<th>Moderately Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

6. I find that a well ordered life with regular hours suits my temperament.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Moderately Disagree</th>
<th>Slightly Disagree</th>
<th>Slightly Agree</th>
<th>Moderately Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

7. I enjoy the uncertainty of going into a new situation without knowing what might happen.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Moderately Disagree</th>
<th>Slightly Disagree</th>
<th>Slightly Agree</th>
<th>Moderately Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

8. When dining out, I like to go to places where I have been before so that I know what to expect.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Moderately Disagree</th>
<th>Slightly Disagree</th>
<th>Slightly Agree</th>
<th>Moderately Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

9. I feel uncomfortable when I don't understand the reason why an event occurred in my life.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Moderately Disagree</th>
<th>Slightly Disagree</th>
<th>Slightly Agree</th>
<th>Moderately Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

10. I feel irritated when one person disagrees with what everyone else in a group believes.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Moderately Disagree</th>
<th>Slightly Disagree</th>
<th>Slightly Agree</th>
<th>Moderately Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

11. I hate to change my plans at the last minute.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Moderately Disagree</th>
<th>Slightly Disagree</th>
<th>Slightly Agree</th>
<th>Moderately Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

12. I would describe myself as indecisive.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Moderately Disagree</th>
<th>Slightly Disagree</th>
<th>Slightly Agree</th>
<th>Moderately Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

13. When I go shopping, I have difficulty deciding exactly what it is I want.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Moderately Disagree</th>
<th>Slightly Disagree</th>
<th>Slightly Agree</th>
<th>Moderately Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>
14. When faced with a problem I usually see the one best solution very quickly.

| Disagree strongly | Disagree moderately | Disagree slightly | Agree slightly | Agree moderately | Agree strongly |

15. When I am confused about an important issue, I feel very upset.

| Disagree strongly | Disagree moderately | Disagree slightly | Agree slightly | Agree moderately | Agree strongly |

16. I tend to put off making important decisions until the last possible moment.

| Disagree strongly | Disagree moderately | Disagree slightly | Agree slightly | Agree moderately | Agree strongly |

17. I usually make important decisions quickly and confidently.

| Disagree strongly | Disagree moderately | Disagree slightly | Agree slightly | Agree moderately | Agree strongly |

18. I think it is fun to change my plans at the last moment.

| Disagree strongly | Disagree moderately | Disagree slightly | Agree slightly | Agree moderately | Agree strongly |

19. My personal space is usually messy and disorganized.

| Disagree strongly | Disagree moderately | Disagree slightly | Agree slightly | Agree moderately | Agree strongly |

20. In most social conflicts, I can easily see which side is right and which is wrong.

| Disagree strongly | Disagree moderately | Disagree slightly | Agree slightly | Agree moderately | Agree strongly |

21. I tend to struggle with most decisions.

| Disagree strongly | Disagree moderately | Disagree slightly | Agree slightly | Agree moderately | Agree strongly |

22. I believe orderliness and organization are among the most important characteristics of a good student.

| Disagree strongly | Disagree moderately | Disagree slightly | Agree slightly | Agree moderately | Agree strongly |
23. When considering most conflict situations, I can usually see how both sides could be right.

| □ Disagree strongly | □ Disagree moderately | □ Disagree slightly | □ Agree slightly | □ Agree moderately | □ Agree strongly |

24. I don't like to be with people who are capable of unexpected actions.

| □ Disagree strongly | □ Disagree moderately | □ Disagree slightly | □ Agree slightly | □ Agree moderately | □ Agree strongly |

25. I prefer to socialize with familiar friends because I know what to expect from them.

| □ Disagree strongly | □ Disagree moderately | □ Disagree slightly | □ Agree slightly | □ Agree moderately | □ Agree strongly |

26. I think that I would learn best in a class that lacks clearly stated objectives and requirements.

| □ Disagree strongly | □ Disagree moderately | □ Disagree slightly | □ Agree slightly | □ Agree moderately | □ Agree strongly |

27. When thinking about a problem, I consider as many different opinions on the issue as possible.

| □ Disagree strongly | □ Disagree moderately | □ Disagree slightly | □ Agree slightly | □ Agree moderately | □ Agree strongly |

28. I don't like to go into a situation without knowing what I can expect from it.

| □ Disagree strongly | □ Disagree moderately | □ Disagree slightly | □ Agree slightly | □ Agree moderately | □ Agree strongly |

29. I like to know what people are thinking all the time.

| □ Disagree strongly | □ Disagree moderately | □ Disagree slightly | □ Agree slightly | □ Agree moderately | □ Agree strongly |

30. I dislike it when a person's statement could mean many different things.

| □ Disagree strongly | □ Disagree moderately | □ Disagree slightly | □ Agree slightly | □ Agree moderately | □ Agree strongly |

31. It's annoying to listen to someone who cannot seem to make up his or her mind.

| □ Disagree strongly | □ Disagree moderately | □ Disagree slightly | □ Agree slightly | □ Agree moderately | □ Agree strongly |
32. I find that establishing a consistent routine enables me to enjoy life more.

33. I enjoy having a clear and structured mode of life.

34. I prefer interacting with people whose opinions are very different from my own.

35. I like to have a plan for everything and a place for everything.

36. I feel uncomfortable when someone's meaning or intention is unclear to me.

37. When trying to solve a problem I often see so many possible options that it's confusing.

38. I always see many possible solutions to problems I face.

39. I'd rather know bad news than stay in a state of uncertainty.

40. I do not usually consult many different options before forming my own view.
41. I dislike unpredictable situations.

- Disagree strongly
- Disagree moderately
- Disagree slightly
- Agree slightly
- Agree moderately
- Agree strongly

42. I dislike the routine aspects of my work (studies).

- Disagree strongly
- Disagree moderately
- Disagree slightly
- Agree slightly
- Agree moderately
- Agree strongly

(The below was not shown to participants)

Scoring:

For all questions except 2, 5, 7, 12, 13, 16, 18, 19, 21, 23, 26, 27, 34, 37, 38, and 42.

- Disagree strongly 1
- Disagree moderately 2
- Disagree slightly 3
- Agree slightly 4
- Agree moderately 5
- Agree strongly 6

For questions 2, 5, 7, 12, 13, 16, 18, 19, 21, 23, 26, 27, 34, 37, 38, and 42.

- Disagree strongly 6
- Disagree moderately 5
- Disagree slightly 4
- Agree slightly 3
- Agree moderately 2
- Agree strongly 1
Appendix 2: Profile screen (Experiment 2)

You’re logged in as user: 2 and Worker ID: 2! This is your profile page.

Thank you for choosing to participate in this experiment. Remember that you may withdraw from this experiment at any time. If you withdraw, your data will be omitted from any published results and you will forgo any rewards based on the completion of this experiment.

This experiment consists of 3 parts that must be completed in sequence: a design activity, questionnaire, and survey. After you have finished all 3 parts, this page will display a confirm code to enter into the Mechanical Turk page. Please follow the instructions below to complete the experiment.

Part 1: Design Activity

Part 1: NOT COMPLETED. Please click here to view the Design Activity instructions!

Part 2: Questionnaire

LOCKED! Please complete Part 1 to unlock!

Part 3: Survey

LOCKED! Please complete Part 2 to unlock!

If you wish to withdraw from the experiment, click here and follow the instructions.

(Initial profile screen)
You're logged in as user: 2 and Worker ID: 2! This is your profile page.

Thank you for choosing to participate in this experiment.
Remember that you may withdraw from this experiment at any time.
If you withdraw, your data will be omitted from any published results and you will forgo any rewards based on the completion of this experiment.

This experiment consists of 3 parts that must be completed in sequence: a design activity, questionnaire, and survey.
After you have finished all 3 parts, this page will display a confirm code to enter into the Mechanical Turk page.
Please follow the instructions below to complete the experiment.

Part 1: Design Activity

Part 1: COMPLETED!

Part 2: Questionnaire

Part 2: COMPLETED!

Part 3: Survey

Part 3: COMPLETED! Congratulations, you have completed the experiment.

You have just completed the Need for Closure Scale!
Need for Closure Scores range from 42 to 252. Need for Closure is an individual characteristic.
For an individual with a high Need for Closure Score (e.g. greater than 147), less information may be processed before forming a judgment (e.g. a shower concept).
The opposite is true for an individual with a low Need for Closure Score (e.g. less than 147).
Your Need for Closure Score is: 152

Return to the Mechanical Turk page and submit the Confirm Code: 379_2_3

Then click here to submit all data and exit the experiment.

(Profile screen after completing all 3 parts)
Appendix 3: Drawing canvas (Experiment 2: Group 3, high NFC<sub>e</sub> condition)

**Concept Drawings and Submissions**

**IMPORTANT:** Do not press the back, forward, or refresh buttons on your internet browser!

Please draw your concepts on the canvas below and then press the 'Submit Your Concept!' button to submit. Ensure that you describe how your concept works and note the different features using the textbox function.

You MUST submit at least THREE CONCEPTS (Draw only one concept per canvas - press Submit after you finish the current concept).

Do not consult any outside sources.

You have 5 minutes to draw and submit each concept. The amount of time remaining is indicated by the floating timer.

Once time has expired, the page will reload and all current canvas progress will be lost and this will be noted.
Appendix 4: Sample timestamp output

Homepage Worker: A________D (edited for anonymity) Condition: 3 Status: , Wed Mar 23 2016 @ 18:13:44

Looked at Design Problem, Wed Mar 23 2016 @ 18:14:02

entered drawing page (started drawing) (test), Wed Mar 23 2016 @ 18:16:00

submitted test design number 1, Wed Mar 23 2016 @ 18:17:40

entered drawing page (started drawing) (concepts), Wed Mar 23 2016 @ 18:17:51

submitted concept design number 1, Wed Mar 23 2016 @ 18:21:24

entered drawing page (started drawing) (concepts), Wed Mar 23 2016 @ 18:21:34

submitted concept design number 2, Wed Mar 23 2016 @ 18:23:38

entered drawing page (started drawing) (concepts), Wed Mar 23 2016 @ 18:23:40

submitted concept design number 3, Wed Mar 23 2016 @ 18:26:33

Homepage Worker: A________D Condition: 3 Status: , Wed Mar 23 2016 @ 18:26:35

On Personality Questionnaire, Wed Mar 23 2016 @ 18:26:39

Your Need for Closure Score is: 154 <br />

Order Scale: 47 <br />

Predictability Scale: 29 <br />

Decisiveness Scale: 25 <br />

Ambiguity Scale: 30 <br />

Closed Mindedness Scale: 23 <br />, Wed Mar 23 2016 @ 18:31:48

Completed Personality Questionnaire, Wed Mar 23 2016 @ 18:31:48
Q1: 1

Q2: Traditional ones

Q3: Female

Q4: 38

Q5: 136

Q6: , Wed Mar 23 2016 @ 18:34:06
Appendix 5: Inquisition of NFC estimate (Experiment 2)

Workers were first given a short description of NFC following completion of the 42-item NFC scale.

You have just completed the Need for Closure Scale!
Need for Closure Scores range from 42 to 252. Need for Closure is an individual characteristic. For an individual with a high Need for Closure Score (e.g. greater than 147), less information may be processed before forming a judgment (e.g. a shower concept). The opposite is true for an individual with a low Need for Closure Score (e.g. less than 147). You will be given your Need for Closure Score at the end of the experiment.

Questionnaire completed! Please return to profile page!

Return to your profile page.

Workers were asked to estimate their NFC score in question 5 of the exit survey.