# ASSESSING EFFORT DISUTILITY IN DECISION DELEGATION TASKS

An Undergraduate Research Scholars Thesis

by

CAMERON M. MCCANN

Submitted to the Undergraduate Research Scholars program at
Texas A&M University
in partial fulfillment of the requirements for the designation as an

UNDERGRADUATE RESEARCH SCHOLAR

Approved by Research Advisor: Dr. Heather C. Lench

May 2018

Major: Psychology

# TABLE OF CONTENTS

# ABSTRACT

Assessing Effort Disutility in Decision Delegation Tasks

Cameron M. McCann
Department of Psychological and Brain Sciences
Texas A&M University

Research Advisor: Dr. Heather C. Lench
Department of Psychological and Brain Sciences
Texas A&M University

There is an ongoing debate in the field of engineering over different approaches to the process of delegation, two of which are requirements allocation (RA) and value-driven design (VDD). In the RA approach, requirements are used to communicate the desired outcome. In the VDD approach, certain values of the project are used to communicate the desired outcome. Drawing on economic principal-agent theory, we expect individuals in the requirements allocation approach to have greater effort disutility than those in the value-driven approach. Participants played a computer game with a goal to receive points. Participants were randomly assigned to one of two conditions. In the requirements allocation condition, the task was described by the need to meet a specified threshold; in the value-driven design condition, the task was described by the need to optimize. The game required participants to make multiple decisions at different difficulty levels with measurements over their effort being recorded throughout the experiment. Our results reveal evidence supporting the value-driven design approach as well as patterns found in the data.

# ACKNOWLEDGEMENTS

Data collection was supported by the National Science Foundation Award #1563379 to Drs. Richard Malak, Heather C. Lench, and Rachel Smallman. I would like to thank my adviser, Dr. Heather C. Lench, for her guidance throughout the course of this research. I would also like to thank Yidou Wan for his help, especially during the data analysis, and the large amount of time he has sacrificed for me. Thanks also go to Masden Stribling, from the D.A.T.A. (Dissertation, Article, and Thesis Assistance) program, for her assistance in the writing process. Additionally, I would like to thank Andrea Mendes and Zari K. Haggenmiller for their wonderful feedback they have provided me with.

I would like to extend my gratitude to the Emotion Science lab for data collection and their support. Also, thanks to my parents and friends for their support throughout this process. Finally, I would like to thank my girlfriend, Aira Martin, for her patience, constant encouragement, unending support, and continuous feedback and help throughout this entire project.

# CHAPTER I

# INTRODUCTION

During the designing stages of an engineering project, like large projects in other fields, the project is broken into parts and delegated to specialists. There are different types of approaches in this stage, two of which are requirements allocation (RA) and value-driven design (VDD). In the requirements allocation approach, requirements are used to communicate the desired outcome. In the value-driven design approach, certain values of the project are used to communicate the desired outcome. For example, if a committee is allocating the project of building a spacecraft that will reach Mars to a team of engineers, the instructions to a team based on the requirements allocation approach would be: "Build a spacecraft to reach Mars that goes X velocity." While the instructions based on the value-driven design approach would say, "Build a spacecraft to reach Mars, while maximizing its velocity."

There is ongoing debate, in the field of engineering, about the efficacy of these two approaches. While some researchers believe approaches such as requirements allocation guarantee the coordination of multiple tasks, they further note that these approaches fail in other areas, like facilitating knowledge sharing, during the designing stages (Bertoni, Bertoni, & Isaksson, 2018). Many researchers see value-driven design as an alternative to requirements allocation (Collopy & Hollingsworth, 2011). However, more evidence is needed to support this approach. Our study examined the impact that the type of design delegation approach (requirements allocation or value-driven design) has on disutility of effort during the design process. Disutility of effort occurs at the point in time when an individual's effort becomes less beneficial, counterproductive, or discontinued.

## Principal-Agent Theory

The principal-agent theory provides a frame for understanding processes when one individual/group (the agent) makes decisions on behalf of another individual/group (the principal). Previous studies have shown that there are negative effects on the agent's effort and performance when a principal decides to implement a minimum requirement on tasks (Kajackaite & Werner, 2015; Falk & Kosfeld, 2006). Masella, Meier, and Zahn (2014) showed that negative effects occur due to the implementation of a minimum requirement by the principal in both within- and between-group relationships. They argue that, in the case of the within-group relationship, agents do not expect to be controlled and therefore respond negatively. Likewise, Riener and Widerhold (2016) find negative effects in response to the implementation of control by a principal who had interacted with the agent prior to the task. Kajackaite and Werner (2015) speculate that control implemented by the principal undermines the intrinsic motivation of the agent. Research in psychology has supported this by showing that control lowers an individual's intrinsic motivation (Fisher, 1978). Additionally, in experiments where agents performed poorly when a minimum requirement was implemented most perceived the control as a sign of distrust or limiting autonomy of choice for the agent during the task (Falk & Kosfeld, 2006). Although Schnedler and Vadovic (2011) suggest that control costs are lessened when the agents perceive control as legitimate or necessary.

These observed responses to control are consistent with Self-Determination Theory (Ryan & Deci, 2000), a psychological theory of motivation. The theory states that we have three innate psychological needs that should be met for self-motivation to improve: autonomy, competence, and relatedness. Self-Determination Theory draws a distinction between autonomous and controlled motivation. Autonomous motivation "involves acting with a sense of volition and having the experience of choice" (Gagne & Deci, 2005). Controlled motivation, on the other hand, "involves acting with a sense of pressure, a sense of having to engage in the actions" (Gagne & Deci, 2005).

Studies show that lowering an individual's autonomy, via surveillance (Lepper & Greene, 1975) or deadlines (Amabile, Dejong, & Lepper, 1976), also lowers their interest in the task. However, as Amabile and colleagues (1976) mention, the fact that there is an external constraint, not necessarily the specific form of constraint, is what affects the participants' interest in the activity. Furthermore, when examining individuals' effort, goals are an important aspect. Studies have shown that during tasks, individuals' performance tends to decline as they reach their goal (Goerg & Kube, 2012) and in some cases quit working (Kajackaite & Werner, 2015). Goerg and Kube (2012) argue this is due to the disappearance of intrinsic and extrinsic motivation.

**The Present Investigation**

The framing of statements and questions can affect individuals' perception (Levin & Gaeth, 1988) and decisions (Tversky & Kahneman, 1981). Additionally, Goldsmith & Dhar (2013) showed that the framing of incentives (positive or negative) can cause participants' effort to be significantly different between the two conditions. This led us to question whether the framing of instructions could influence individuals' effort in a task. Based on past work demonstrating lowered motivation under conditions of control and when goals are met, we hypothesized that the requirements allocation approach will have a higher rate of effort disutility than the value-driven design approach. In other words, participants will experience less beneficial, if not counterproductive, effects on their effort when they are given instructions that include requirements to meet.

# CHAPTER II

# METHODS

## Participants

Participants were 109 Texas A&M University undergraduate college students (41 men, 67 women, and 1 missing) who completed the study for course credit in an introductory psychology course. The requirements allocation condition had 55 participants, and the value-driven design condition had 54 participants. Participants' age varied from 18 years-old to 25 years-old, with the mean age being 19.01 years (*SD* = 1.45). However, 35 participants (32.1%) failed to record their age. Furthermore, participants described themselves as African-American (6.4%), Asian (7.3%), White (74.3%), American Indian or Alaska Native (1.8%), Native Hawaiian or Other Pacific Islander, White (0.9%), American Indian or Alaska Native, Asian (0.9%), American Indian or Alaska Native, White (1.8%), and Other (5.5%). In addition to one of the previous categories, 24.1% of participants described themselves as Hispanic. The majority of participants were not engineering students, although the most common major was engineering (25.7%) with the next most common being: biomedical science (11.0%), business (9.2%), and psychology (7.3%). The majority of the participants were freshmen (59.6%) followed by sophomores (21.1%), juniors (10.1%), and seniors (7.3%).

## Materials

Our study consisted of a between-subjects research design with the independent variable being the framing of the instructions for the task. Participants played a variation of a computer game called, "Manned Mission: Mars," which has been used in previous research to study decision making (Vermillion, Malak, Smallman, & Fields, 2014). The lab—where participants played the computer game—had three single-person cubicles laterally aligned, side-by-side, with one computer in each, allowing for one to three participants to participate in each lab session. The computers used in the experiment randomly assigned the participants to one of two conditions, requirements allocation condition or value-driven design condition, with the instructions differing between the two. In the requirements allocation condition, the instructions stated, “Your goal is to get a robustness score of 70%.” Participants in this condition received target robustness scores of 60%, 70%, 80%, 90%, or 100%. In the value-driven design condition the instructions stated, “Your goal is to maximize the robustness percentage.” Prior to the start of the game, participants were told to read the instructions on the computer screen carefully. The on-screen instructions explained how to play the game before the first level began. Then the participants played all four levels of the game, while the computer recorded their effort by measuring the amount of time spent on each level, the number of components moved, and the number of times they analyzed their system before submission.

After the game, participants completed a short survey consisting of the revised Need for Cognition Scale (Cacioppo, Petty, & Kao, 1984), the Achievement Goal Questionnaire (Elliot & McGregor, 2001), and a few demographic questions. The revised Need for Cognition Scale included a 5-point Likert scale (from 1 = “Extremely uncharacteristic” to 5 = “Extremely characteristic of me”) for 18-items. This scale assesses an “individuals… tendency to engage in and enjoy thinking” (Cacioppo & Petty, 1982). So those with high scores on this assessment tend to think more and enjoy the process more than those with low scores.

The Achievement Goal Questionnaire included a 7-point Likert scale (from 1 = “Strongly disagree” to 7 = “Strongly agree”) for 12-items. This scale measures the participants’ achievement goal (e.g., mastery or performance) and its direction (e.g., avoidance or approach).

Mastery goals focus on developing the complete competency of a task; whereas performance goals focus on demonstrating one's ability on a task. Each type of goal is divided into two dimensions in this scale for a total of four dimensions being measured: mastery-approach, mastery-avoidance, performance-approach, and performance-avoidance. Those high in mastery-approach strive to improve their competency and the complete mastery of the task. In contrast, individuals high in mastery-avoidance aim to avoid the state of being incompetent. Furthermore, those high in performance-approach strive to demonstrate their ability or performance relative to an external standard; whereas individuals high in performance-avoidance focus on avoiding the demonstration of their inability.

*Game Description*

The game asks participants to build structures for the rocket associated with a manned mission to mars and were provided with the robustness score for their structure as feedback on how well they were doing. In the game there were four levels, each required the participants to make multiple decisions on tasks, with the difficulty increasing with each level. For every level, participants moved multiple components around to build the structure and, depending on the placement of these components, the robustness of the system would change. The game starts off with three different component colors: red, blue, and yellow. The robustness of the system would increase if certain colored components were placed near specific vicinities of the system: red toward the top left, blue near the center, and yellow toward the bottom right. Participants could move the components around and analyze the robustness of the system as many times as they wished before submitting the system and moving on to the next level. However, there was one requirement: the system had to be feasible—meaning the components all had to fit within the system's square perimeter.

Instructions appeared on the screen before each level. These instructions were framed differently depending on which condition the student was in. In the requirements allocation condition, the task was described by the need to meet a specified robustness threshold (e.g., “Your goal is to get a robustness score of 70%”). Participants in the requirements allocation condition received target robustness scores of 60%, 70%, 80%, 90%, or 100%. In the value-driven design condition, the task was described by the need to optimize the robustness percentage of the system (e.g., “Your goal is to maximize the robustness percentage”).

The first level contains nine square homogeneous components that sit in a three-by-three fashion (see Appendix A, Image 1). Each component was one of the following colors: red, blue, or yellow. The second level replaces a few square shaped components with rectangular shaped ones—making the components no longer homogenous (see Appendix A, Image 2). These rectangular components were the size of two square components. The third level introduces two new colors—green and orange (see Appendix A, Image 3). These new colors are mixtures of two existing colors (e.g., blue and yellow forming green; red and yellow forming orange), with the composing colors signaled in the corners of the component. In relation to the robustness of the system, these new components are considered the two colors they are made up of; this introduces the challenge of finding the optimal position for these new components. In the fourth level (see Appendix A, Image 4), the difficulty increases drastically in multiple ways compared to the third level. First, the increase in difficulty is partly due to the addition of another color: purple—which is composed of red and blue. So, to optimize the purple components’ effect on the robustness of the system, they had to be treated as both a red and blue component. Thus there were six different colors and their associated requirements in the fourth level: red, blue, yellow, green, orange, and purple. Second, the change in difficulty was also due to the large increase in the number of components. The third level contained seven components, while this number increased to 19 for the fourth level.

# CHAPTER III

# RESULTS

Our preliminary results showed that, on average, participants in the VDD condition attained higher robustness scores than participants in the RA condition in every level of the game with significantly higher robustness scores in the third level, $t(107) = 2.19$, $p = .031$, $d = .42$ (see Table 1). However, compared to those in the VDD condition, participants in the RA condition tended to move more components in every level of the game (see Table 2).

**Table 1.** *Results for t-test and descriptive statistics for VDD and RA on Robustness Scores*

| | Condition | | | | | | 95% Confidence Interval of the Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | VDD | | | RA | | | | | |
| | *M* | *SD* | *N* | *M* | *SD* | *N* | | *t* | *df* |
| Level 1 | 93.31 | 20.15 | 54 | 93.04 | 12.89 | 55 | -6.13, 6.69 | 0.09 | 107 |
| Level 2 | 94.70 | 13.66 | 54 | 91.18 | 12.21 | 55 | -1.40, 8.44 | 1.42 | 107 |
| Level 3 | 94.04 | 13.55 | 54 | 87.60 | 16.97 | 55 | .60, 12.28 | 2.19* | 107 |
| Level 4 | 51.43 | 33.26 | 54 | 50.33 | 32.80 | 55 | -11.45, 13.64 | 0.17 | 107 |

* $p < .05$. ** $p < .01$.

Furthermore, participants in the RA condition tended to spend more time in the first, second, and fourth levels of the game than those in the VDD condition (see Table 3). In the first three levels of the game, participants in the RA condition, on average, analyzed their systems more than participants in the VDD condition. In the first level though, those in the RA condition ($M = 7.29$, $SD = 8.67$) analyzed their system significantly more than participants in the VDD condition ($M = 4.08$, $SD = 3.52$), $t(68) = -2.46$, $p = .016$, $d = .60$ (see Table 4). The degrees of freedom were adjusted from 99 to 68, because the Levene's test indicated unequal variances (*F* = 19.64, $p < .001$) between these two conditions.

**Table 2.** *Results for t-test and descriptive statistics for VDD and RA on components moved*

| | Condition | | | | | | 95% Confidence Interval of the Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | VDD | | | RA | | | | | |
| | *M* | *SD* | *N* | *M* | *SD* | *N* | | *t* | *df* |
| Level 1 | 18.57 | 15.11 | 54 | 24.25 | 22.49 | 53 | -13.04, 1.70 | -1.53 | 91 |
| Level 2 | 15.13 | 12.87 | 53 | 18.96 | 17.76 | 53 | -9.81, 2.15 | -1.27 | 95 |
| Level 3 | 17.26 | 16.62 | 54 | 19.63 | 20.23 | 54 | -9.43, 4.69 | -0.67 | 106 |
| Level 4 | 166.06 | 105.04 | 54 | 198.44 | 130.72 | 55 | -77.46, 12.70 | -1.42 | 107 |

* $p < .05$. ** $p < .01$.

**Table 3.** *Results for t-test and descriptive statistics for VDD and RA on time spent*

| | Condition | | | | | | 95% Confidence Interval of the Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | VDD | | | RA | | | | | |
| | *M* | *SD* | *N* | *M* | *SD* | *N* | | *t* | *df* |
| Level 1 | 150749 | 43654 | 53 | 160956 | 66732 | 52 | -32122, 11708 | -0.93 | 88 |
| Level 2 | 63163 | 39376 | 53 | 70869 | 53748 | 53 | -25875, 10462 | -0.84 | 95 |
| Level 3 | 89518 | 56500 | 54 | 88269 | 57528 | 54 | -20506, 23003 | 0.11 | 106 |
| Level 4 | 733531 | 454721 | 54 | 814232 | 552575 | 55 | -273045, 111644 | -0.83 | 107 |

* $p < .05$. ** $p < .01$.

The threshold participants were given in the RA condition varied, and therefore we also assessed robustness by the specific threshold provided. However, this represented a comparison among six groups [i.e., VDD, RA(60), RA(70), RA(80), RA(90), and RA(100)], resulting in small sample sizes for each level of the RA condition in comparison to the VDD condition, making the statistical comparisons less reliable. Therefore, we looked for patterns in the data. The most common pattern in the data was that the RA conditions used more effort (i.e., time spent, components moved, times analyzed) than the VDD condition but finished with a lower robustness score. This pattern was present in all four levels of the game; although, it was more prevalent in the lower levels. In the first level RA(60), RA(80), and RA(90) spent more time (see Appendix B, Figure 1A), moved more components (see Appendix B, Figure 1B), and analyzed their systems more frequently than the VDD condition (see Appendix B, Figure 1C) but finished with lower robustness scores (see Appendix B, Figure 1D). This pattern occurred in level two for both RA(70) and RA(80) (see Appendix B, Figures 2A-D). However, RA(70) had a significantly lower final robustness score ($M$ = 77.33, $SD$ = 11.53) than VDD ($M$ = 94.70, $SD$ = 13.66), $t(61) = 3.60$, $p = .001$, $d = .92$ (see Table 5). Although for the third and fourth level, the pattern was only noticeable in the RA(80) condition (see Appendix B, Figures 3A-D & 4A-D). But, in the fourth level, RA(80) moved significantly more components ($M$ = 242.30, $SD$ = 121.64) compared to VDD ($M$ = 166.06, $SD$ = 105.04), $t(62) = -2.06$, $p = .044$, $d = .52$ (see Table 6).

**Table 4.** *Results for t-test and descriptive statistics for VDD and RA on times analyzed*

| | Condition | | | | | | 95% Confidence Interval of the Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | VDD | | | RA | | | | | |
| | *M* | *SD* | *N* | *M* | *SD* | *N* | | *t* | *df* |
| Level 1 | 4.08 | 3.52 | 49 | 7.29 | 8.67 | 52 | -5.81, -.61 | -2.46* | 68 |
| Level 2 | 5.40 | 5.26 | 52 | 7.42 | 8.09 | 52 | -4.68, .64 | -1.51 | 88 |
| Level 3 | 6.34 | 6.60 | 53 | 7.15 | 7.58 | 52 | -3.56, 1.93 | -0.59 | 103 |
| Level 4 | 32.85 | 20.49 | 52 | 32.59 | 27.01 | 51 | -9.10, 9.62 | 0.06 | 101 |

* $p < .05$. ** $p < .01$.

**Table 5.** *Results for t-test and descriptive statistics for VDD and RA(70) on second level*

| | Condition | | | | | | 95% Confidence Interval of the Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | VDD | | | RA(70) | | | | | |
| | *M* | *SD* | *N* | *M* | *SD* | *N* | | *t* | *df* |
| Time Spent | 63162.60 | 39376.08 | 53 | 63736.22 | 46102.73 | 9 | -29663.63, 28516.39 | -0.04 | 60 |
| Times Analyzed | 5.40 | 5.26 | 52 | 7.22 | 6.67 | 9 | -5.77, 2.13 | -0.92 | 59 |
| Components Moved | 15.13 | 12.87 | 53 | 18.89 | 21.52 | 9 | -20.47, 12.96 | -0.51 | 9 |
| Robustness Score | 94.70 | 13.66 | 54 | 77.33 | 11.53 | 9 | 7.72, 27.02 | 3.60** | 61 |

* $p < .05$. ** $p < .01$.

**Table 6.** *Results for t-test and descriptive statistics for VDD and RA(80) on fourth level*

| | Condition | | | | | | 95% Confidence Interval of the Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | VDD | | | RA(80) | | | | | |
| | *M* | *SD* | *N* | *M* | *SD* | *N* | | *t* | *df* |
| Time Spent | 733531.31 | 454720.89 | 54 | 1002646.80 | 596265.95 | 10 | -597978, 59747 | -1.636 | 62 |
| Times Analyzed | 32.85 | 20.49 | 52 | 43.40 | 28.13 | 10 | -25.62, 4.51 | -1.401 | 60 |
| Components Moved | 166.06 | 105.04 | 54 | 242.30 | 121.64 | 10 | -150.30, -2.19 | -2.058* | 62 |
| Robustness Score | 51.43 | 33.26 | 54 | 37.70 | 25.03 | 10 | -8.43, 35.88 | 1.238 | 62 |

* $p < .05$. ** $p < .01$.

The next three most common patterns occurred three times each. One of these three shows the RA conditions spending less time but moving more components and analyzing their systems more compared to the VDD condition, while still finishing with a higher robustness score. This pattern shows that the RA condition, in these cases, is exerting more effort in a shorter period of time. This pattern was present in level one for the RA(100) condition (see Appendix B, Figures 1A-D) and in level three for RA(90) and RA(100) (see Appendix B, Figures 3A-D). In the case of the RA(100) condition in level three, they had a significantly higher final robustness score (*M* = 100.00, *SD* = .00) than VDD did (*M* = 94.04, *SD* = 13.55), *t*(53) = -3.23, *p* = .002, *d* = .89 (see Table 7). The degrees of freedom were adjusted from 59 to 53, because the Levene's test indicated unequal variances (*F* = 6.18, *p* = .016) between these two conditions.

**Table 7.** *Results for t-test and descriptive statistics for VDD and RA(100) on third level*

| | Condition | | | | | | 95% Confidence Interval of the Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | VDD | | | RA(100) | | | | | |
| | *M* | *SD* | *N* | *M* | *SD* | *N* | | *t* | *df* |
| Time Spent | 89517.59 | 56499.75 | 54 | 73364.71 | 38303.24 | 7 | -27997.82, 60303.58 | 0.73 | 59 |
| Times Analyzed | 6.34 | 6.60 | 53 | 7.14 | 9.74 | 7 | -6.43, 4.83 | -0.29 | 58 |
| Components Moved | 17.26 | 16.62 | 54 | 18.86 | 17.89 | 7 | -15.07, 11.87 | -0.24 | 59 |
| Robustness Score | 94.04 | 13.55 | 54 | 100.00 | .00 | 7 | -9.66, -2.26 | -3.23** | 53 |

* *p* < .05. ** *p* < .01.

The next pattern was that the RA conditions spent more effort (i.e., time spent, components moved, times analyzed) than VDD and finished with a higher robustness score. This pattern is present in the second and fourth level. In the second level, this pattern was experienced by RA(90) and RA(100) (see Appendix B, Figures 2A-D). In the fourth level the RA(100) condition experienced this pattern (see Appendix B, Figures 4A-D); they spent significantly more time than the VDD condition, *t*(62) = -2.00, *p* = .049, *d* = .51, with mean times of 1,060,259.10 ms (*SD* = 572,097.87 ms) and 733,531.31 ms (*SD* = 454,720.89 ms), respectively. They also moved significantly more components (*M* = 237.60, *SD* = 88.14) compared to VDD (*M* = 166.06, *SD* = 105.04), *t*(62) = -2.02, *p* = .047, *d* = .51, and analyzed their system more, while finishing with a higher final robustness score (see Table 8).

**Table 8.** *Results for t-test and descriptive statistics for VDD and RA(100) on fourth level*

| | Condition | | | | | | 95% Confidence Interval of the Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | VDD | | | RA(100) | | | | | |
| | *M* | *SD* | *N* | *M* | *SD* | *N* | | *t* | *df* |
| Time Spent | 733531.31 | 454720.89 | 54 | 1060259.10 | 572097.87 | 10 | -652626.05, -829.52 | -2.00* | 62 |
| Times Analyzed | 32.85 | 20.49 | 52 | 45.78 | 38.11 | 9 | -42.47, 16.61 | -.99 | 9 |
| Components Moved | 166.06 | 105.04 | 54 | 237.60 | 88.14 | 10 | -142.26, -.83 | -2.02* | 62 |
| Robustness Score | 51.43 | 33.26 | 54 | 59.20 | 41.38 | 10 | -31.56, 16.01 | -.65 | 62 |

* $p < .05$. ** $p < .01$.

The next pattern, opposite to the previous one, showed the RA conditions exerting less effort (i.e., time spent, components moved, times analyzed) and finishing with lower robustness scores than the VDD condition. Unlike the previous two patterns, this one is noticeable in three different levels: level two, with the RA(60) condition (see Appendix B, Figures 2A-D); level three, with the RA(70) condition (see Appendix B, Figures 3A-D); and level four, with the RA(90) condition (see Appendix B, Figures 4A-D). In the second level RA(60) finished with a significantly lower final robustness score ($M$ = 77.75, $SD$ = 20.60) than VDD ($M$ = 94.70, $SD$ = 13.66), $t(56) = 2.32$, $p = .024$, $d = .62$ (see Table 9). Likewise, in the third level the RA(70) condition finished with significantly lower final robustness score ($M$ = 79.69, $SD$ = 13.31) compared to VDD ($M$ = 94.04, $SD$ = 13.55), $t(65) = 3.44$, $p = .001$, $d = .85$ (see Table 10).

**Table 9.** *Results for t-test and descriptive statistics for VDD and RA(60) on second level*

| | Condition | | | | | | 95% Confidence Interval of the Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | VDD | | | RA(60) | | | | | |
| | *M* | *SD* | *N* | *M* | *SD* | *N* | | *t* | *df* |
| Time Spent | 63162.60 | 39376.08 | 53 | 43147.67 | 20182.10 | 3 | -26191.66, 66221.54 | 0.87 | 54 |
| Times Analyzed | 5.40 | 5.26 | 52 | 4.67 | 3.79 | 3 | -5.47, 6.94 | 0.24 | 53 |
| Components Moved | 15.13 | 12.87 | 53 | 10.67 | 7.23 | 3 | -10.65, 19.58 | 0.59 | 54 |
| Robustness Score | 94.70 | 13.66 | 54 | 77.75 | 20.60 | 4 | 2.30, 31.61 | 2.32* | 56 |

* $p < .05$. ** $p < .01$

Note: consistent with the pattern mentioned, the RA condition spent less time, analyzed their system less, moved less components, and finished with a significantly lower robustness score compared to the VDD condition.

**Table 10.** *Results for t-test and descriptive statistics for VDD and RA(70) on third level*

| | Condition | | | | | | 95% Confidence Interval of the Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | VDD | | | RA(70) | | | | | |
| | *M* | *SD* | *N* | *M* | *SD* | *N* | | *t* | *df* |
| Time Spent | 89517.59 | 56499.75 | 54 | 77019.62 | 60226.40 | 13 | -22797.47, 47793.42 | 0.71 | 65 |
| Times Analyzed | 6.34 | 6.60 | 53 | 5.46 | 5.72 | 13 | -3.11, 4.86 | 0.44 | 64 |
| Components Moved | 17.26 | 16.62 | 54 | 14.31 | 18.53 | 13 | -7.53, 13.43 | 0.56 | 65 |
| Robustness Score | 94.04 | 13.55 | 54 | 79.69 | 13.31 | 13 | 6.01, 22.68 | 3.44** | 65 |

* $p < .05$. ** $p < .01$.

To see the participants' performance relative to the threshold provided to them, we subtracted their achieved robustness score from what score was presented in their instructions for that level. We treated the VDD condition like the RA(100) condition and subtracted by 100. We then checked if any relationships existed between this new variable, the difference, and any effort measurements (i.e., time spent on level, components moved, or times system was analyzed) by running Pearson's correlation test. The trend we noticed was that, for the first three levels, the participants' effort (i.e., time spent, components moved, times analyzed) was inversely related with their difference variable—meaning that when the participants exerted more effort, they tended to score lower on robustness—but in the fourth level their effort was directly related to this variable.

When broken out by specific threshold in the RA condition, there were significant correlations between effort and performance for multiple thresholds in the first, second, and third level and two thresholds in the fourth game level: RA(60) and VDD. In the RA(60) condition, the participants' difference score was moderately correlated with the number of components moved, $r$ (14) = .63, $p$ = .008, and time spent on the level, $r$ (14) = .65, $p$ = .007. In the VDD condition, performance was significantly correlated with the number of components moved, $r$ (52) = .33, $p$ = .014, time spent on the level, $r$ (52) = .33, $p$ = .016, and number of times the system was analyzed, $r$ (50) = .50, $p < .001$.

**Individual Differences**

We found individual differences influenced participants' performance and persistence. For example, in the third level, those that scored higher on the need for cognition scale moved less components, $r$ (100) = -.28, $p$ = .005. Although, in the fourth level, those that scored higher on the need for cognition scale were more persistent on the task (i.e., longer time spent, $r$ (101) = .28, $p$ = .004; more components moved, $r$ (101) = .32, $p$ = .001; and more analyses, $r$ (95) = .41, $p < .001$), which led them to finish with higher robustness scores as well, $r$ (101) = .48, $p < .001$. Additionally, those who scored high on this scale also spent longer on the whole game, $r$ (101) = .21, $p$ = .038.

Furthermore, participants' achievement orientations seemed to influence their performance as well. For example, in the first level, those who scored high in performance approach tended to analyze their system less, $r$ (99) = -.23, $p$ = .020, but obtained a higher robustness score, $r$ (106) = .26, $p$ = .006. In the third level, if participants scored high in performance approach, mastery approach, or performance avoidance, they tended to achieve higher robustness scores (performance approach: $r$ (106) = .22, $p$ = .020; mastery approach: $r$ (105) = .19, $p$ = .047; performance avoidance: $r$ (106) = .20, $p$ = .036). On the most difficult level, the fourth level, the participants that scored high in mastery approach tended to be more persistent in the game (i.e., spent longer, $r$ (105) = .20, $p$ = .041; moved more components, $r$ (105) = .27, $p$ = .004; and analyzed more, $r$ (99) = .31, $p$ = .002); this persistence seemed to have paid off, since these individuals also finished with higher robustness scores, $r$ (105) = .29, $p$ = .002. Additionally, participants high in mastery approach tended to, during the whole game, analyze their system more, $r$ (105) = .20, $p$ = .037, and move more components, $r$ (105) = .19, $p$ = .046.

These individual differences may be factors that influenced the results we observed. Take, for instance, the RA(70) condition's significantly lower robustness score, compared to the VDD condition seen in the second level. Participants in this condition that had a higher need for cognition tended to achieve a lower robustness score, $r$ (6) = -.86, $p$ = .006. Likewise, participants in this same condition who scored high in mastery approach also tended to achieve lower robustness scores, $r$ (7) = -.67, $p$ = .049. Furthermore, in the fourth level, the significantly higher amount of components moved by RA(100) compared to VDD may be due in part to the participants' achievement orientation in the RA(100) condition; those that scored high in performance avoidance also moved more components in that condition, $r$ (8) = .68, $p$ = .032.

Interestingly, we found that Engineering students had significantly higher need for cognition scores (*M* = 3.59, *SD* = .74) than non-engineering students (*M* = 3.21, *SD* = .50), t(102) = 3.03, p = .003, d = .60. Additionally, compared to non-engineering students (*M* = 4.19, *SD* = 1.37), the Engineering students had significantly higher mastery approach scores (*M* = 4.83, *SD* = 1.30), *t*(105) = 2.12, *p* =.037, *d* = .41 (see Table 11). Engineering students also spent a significantly longer time on the fourth level (*M* = 941,160.39 ms, *SD* = 567,636.99 ms) than non-engineering students (*M* = 711,550.38 ms, *SD* = 473,512.64 ms), *t*(106) = 2.10, *p* = .039, *d* = .41, which resulted in a significantly higher final robustness score as well (*M* = 69.36, *SD* = 32.08), compared to non-engineering students (*M* = 43.95, *SD* = 30.66), *t*(106) = 3.73, *p* < .001, *d* = .72 (see Table 12).

**Table 11.** *Need for Cognition and Goal Orientation results for Engineering and Non-Engineering Students*

| | Condition | | | | | | 95% Confidence Interval of the Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Engineering | | | Non-engineering | | | | | |
| | *M* | *SD* | *N* | *M* | *SD* | *N* | | *t* | *df* |
| Need for Cognition | 3.59 | .74 | 27 | 3.21 | .50 | 77 | -.13, .64 | 3.03** | 102 |
| Mastery Approach | 4.83 | 1.30 | 27 | 4.19 | 1.37 | 80 | .04, 1.23 | 2.12* | 105 |

* *p* < .05. ** *p* < .01.

**Table 12.** *T-test results and descriptive statistics for Engineering and Non-Engineering Students on fourth level*

| | Condition | | | | | | 95% Confidence Interval of the Difference | | |
|---|---|---|---|---|---|---|---|---|---|
| | Engineering | | | Non-engineering | | | | | |
| | *M* | *SD* | *N* | *M* | *SD* | *N* | | *t* | *df* |
| Time Spent | 941160.39 | 567636.99 | 28 | 711550.38 | 473512.64 | 80 | 12302.20, 446917.84 | 2.10* | 106 |
| Times Analyzed | 61.61 | 56.33 | 28 | 32.48 | 27.79 | 80 | 6.54, 51.73 | 2.63* | 32 |
| Components Moved | 224.46 | 153.36 | 28 | 168.36 | 102.67 | 80 | -7.13, 119.33 | 1.80 | 36 |
| Robustness Score | 69.36 | 32.08 | 28 | 43.95 | 30.66 | 80 | 11.90, 38.92 | 3.73** | 106 |

* $p < .05$. ** $p < .01$.

Note: By spending more time, analyzing more, moving more components, and obtaining higher robustness scores, engineering students seem to be more persistent in the fourth level than non-engineering students.

# CHAPTER IV

# DISCUSSION

These findings suggest there may be signs that people under the RA approach experience greater rates of effort disutility than people under the VDD approach, as we hypothesized, but further research must be conducted to confirm this. This is supported by the fact that the most common pattern in the data, that the RA condition exerts more effort but finishes with a lower robustness score, indicates that the RA condition may be experiencing higher rates of effort disutility throughout the game compared to the VDD condition. In addition, this pattern occurs seven times compared to the next most common pattern, which occurs three times throughout the game. One of the next most common patterns, that the RA condition spent less time but moved more components and analyzed their system more while finishing with a higher robustness score, suggests that the RA condition is exerting more effort in a shorter amount of time compared to the VDD condition. This pattern only occurred in the highest two levels of the RA condition, RA(90) and RA(100), so there is potential that some sort of boundary conditions may exist with this pattern. Future research plans to examine these two patterns further and will see which patterns hold in more difficult tasks.

As seen in previous research, like Kajackaite and Werner (2015), our results show participants performed worse when a minimum requirement exists than when not. This is noticeable when looking at performance in the VDD condition compared to each level of the RA condition. Out of the 20 opportunities, participants in the VDD condition outperformed the RA condition 13 times. In addition, this is noticeable in Table 1—before the RA condition was divided—where the VDD condition outperformed the RA condition in every level of the game.

Another pattern, which appears three times, was that the RA condition exerted less effort and finished with a lower robustness score than the VDD condition. This pattern suggests the RA approach might cause individuals to exert less effort compared to the VDD approach. However, the exact opposite pattern occurs three times as well therefore making this speculation questionable. Nonetheless, in future research, it is worth seeing whether these two patterns occur, and, if so, in what situations do they appear.

In the fourth level, when participants' difference variable, the difference between their achieved robustness score and the threshold mentioned in their instructions, correlated with their effort measurements (i.e., time spent on level, components moved, or times system was analyzed), this suggested some sort of relationship exists between the participants' effort and performance. Although this only occurred for the RA(60) and VDD conditions, there is a possibility that due to the smaller sample sizes for each level of the RA condition that some trends were not able to be detected.

Interestingly the engineering students in our sample scored higher on the need for cognition and the mastery approach on the achievement orientation scale, spent longer on the fourth level, and scored higher on the fourth level than the non-engineering students. All of this seems to suggest that engineering students may be characteristically different than their non-engineering counterparts.

Because the small sample sizes we had at each threshold level in the RA conditions are not ideal for comparison with the VDD condition, these findings should be taken with some caution. Additionally, this study was meant to be exploratory, meaning it was designed to test the waters and guide future research by checking how best to record certain measurements and what aspects to be aware of when setting up future study designs. Furthermore, because we ran multiple participants in the same room there was potential for distractions to occur. Although we told participants to remain seated once they finished the game, during the first week of the experiment we had participants attempt to leave the room. In some cases, this caused a distraction for other participants, which could have some effect on their measurements, like time spent on the level. Furthermore, we had a participant mention that they intentionally lowered their robustness score to get closer to the target goal mentioned in their instructions (and we changed the instructions as a result). This, in combination with the small sample sizes, could have had an impact on the data.

In conclusion, this study has explored different patterns seen in the data of a computer game designed to examine participants effort during a task meant to emulate the decision-making process of the designing stages of an engineering project. There are signs that the requirements allocation (RA) approach may be experiencing higher rates of effort disutility than the value-driven design (VDD) approach. However, further, more extensive research is needed.

## REFERENCES


Amabile, T. M., Dejong, W., & Lepper, M. R. (1976). Effects of externally imposed deadlines on subsequent intrinsic motivation. *Journal of Personality and Social Psychology*, 34(1), 92-98. http://dx.doi.org/10.1037/0022-3514.34.1.92

Bertoni, M., Bertoni, A., & Isaksson, O. (2018). Evoke: A value-driven concept selection method for early system design. *Journal of Systems Science and Systems Engineering*, 27(1), 46-77. https://doi.org/10.1007/s11518-016-5324-2

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116-131. http://dx.doi.org.ezproxy.library.tamu.edu/10.1037/0022-3514.42.1.116

Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48(3), 306-307. http://dx.doi.org.ezproxy.library.tamu.edu/10.1207/s15327752jpa4803_13

Collopy, P. D., & Hollingsworth, P. M. (2011). Value-driven design. *Journal of Aircraft*, 48(3), 749-754. https://doi.org/10.2514/1.C000311

Elliot, A. J., & McGregor, H. A. (2001). A 2X2 achievement goal framework. *Journal of Personality and Social Psychology*, 80(3), 501-519. http://dx.doi.org/10.1037/0022-3514.80.3.501

Falk, A. & Kosfeld, M. (2006). The hidden costs of control. *The American Economic Review*, 96(5), 1611-1630. http://www.jstor.org.ezproxy.library.tamu.edu/stable/30034987

Fisher, C. D. (1978). The effects of personal control, competence, and extrinsic reward systems on intrinsic motivation. *Organizational Behavior and Human Performance*, 21(3), 273-288. http://dx.doi.org.lib-ezproxy.tamu.edu:2048/10.1016/0030-5073(78)90054-5

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational behavior*, 26(4), 331-362. https://doi.org/10.1002/job.322

Goerg, S., & Kube, S. (2012). Goals (th)at Work: Goals, Monetary Incentives, and Workers' Performance. *MPI Collective Goods Preprint*, No. 2012/19. http://dx.doi.org/10.2139/ssrn.2159663

Goldsmith, K., & Dhar, R. (2013). Negativity bias and task motivation: Testing the effectiveness of positively versus negatively framed incentives. *Journal of Experimental Psychology: Applied*, 19(4), 358. http://dx.doi.org.ezproxy.library.tamu.edu/10.1037/a0034415

Kajackaite, A., & Werner, P. (2015). The incentive effects of performance requirements – A real effort experiment. *Journal of Economic Psychology*, 49, 84-94. http://dx.doi.org.lib-ezproxy.tamu.edu:2048/10.1016/j.joep.2015.03.007

Lepper, M. R., & Greene, D. (1975). Turning play into work: Effects of adult surveillance and extrinsic rewards on children's intrinsic motivation. *Journal of Personality and Social Psychology*, *31*(3), 479-486. http://dx.doi.org/10.1037/h0076484

Levin, I. P., & Gaeth, G. J. (1988). How consumers are affected by the framing of attribute information before and after consuming the product. *Journal of Consumer Research*, 15(3), 374-378. https://doi.org/10.1086/209174

Masella, P., Meier, S., & Zahn, P. (2014). Incentives and group identity. *Games and Economic Behavior*, *86*, 12-25. https://doi.org/10.1016/j.geb.2014.02.013

Riener, G., & Wiederhold, S. (2016). Team building and hidden costs of control. *Journal of Economic Behavior & Organization*, 123, 1-18. https://doi.org/10.1016/j.jebo.2015.12.008

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, *55*(1), 68. http://dx.doi.org/10.1037/0003-066X.55.1.68

Schnedler, W., & Vadovic, R. (2011). Legitimacy of control. *Journal of Economics & Management Strategy*, *20*(4), 985-1009. https://doi.org/10.1111/j.1530-9134.2011.00315.x

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453-458. http://www.jstor.org.ezproxy.library.tamu.edu/stable/1685855

Vermillion, S. D., Malak, R. J., Smallman, R., & Fields, S. (2014). Linking normative and descriptive research with serious gaming. *Procedia Computer Science*, 28, 204-212. https://doi.org/10.1016/j.procs.2014.03.026

# APPENDIX A

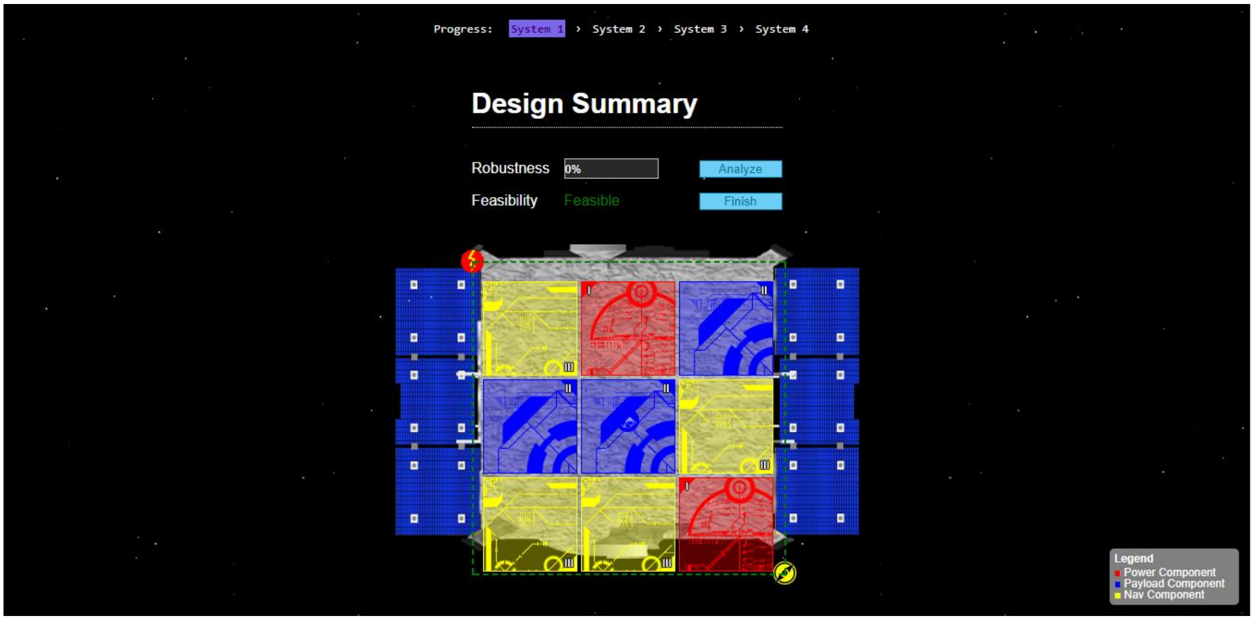


Image 1 First level of game

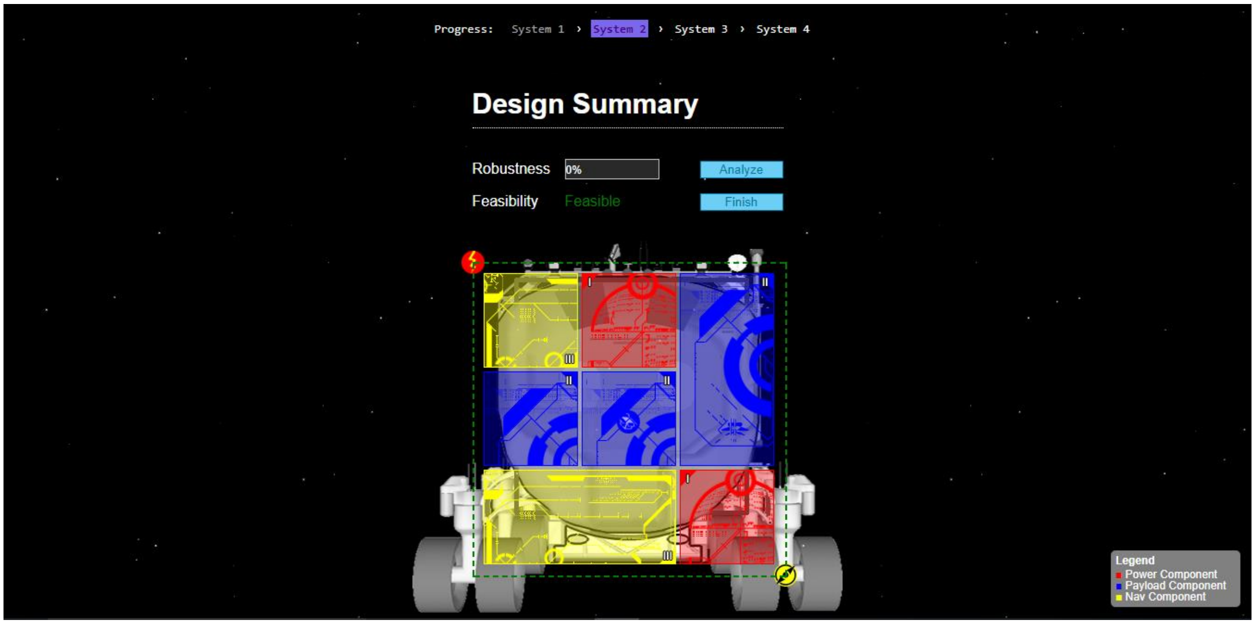


Image 2. Second level of game

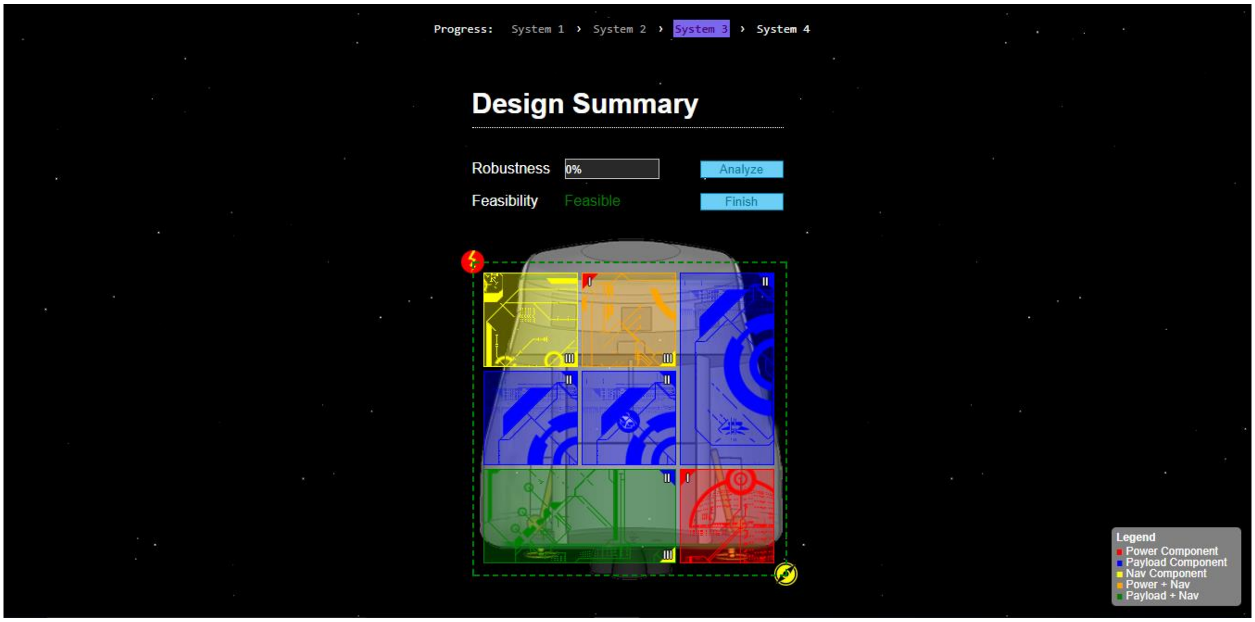


Image 3. Third level of game


Image 4. Fourth level of game

## APPENDIX B

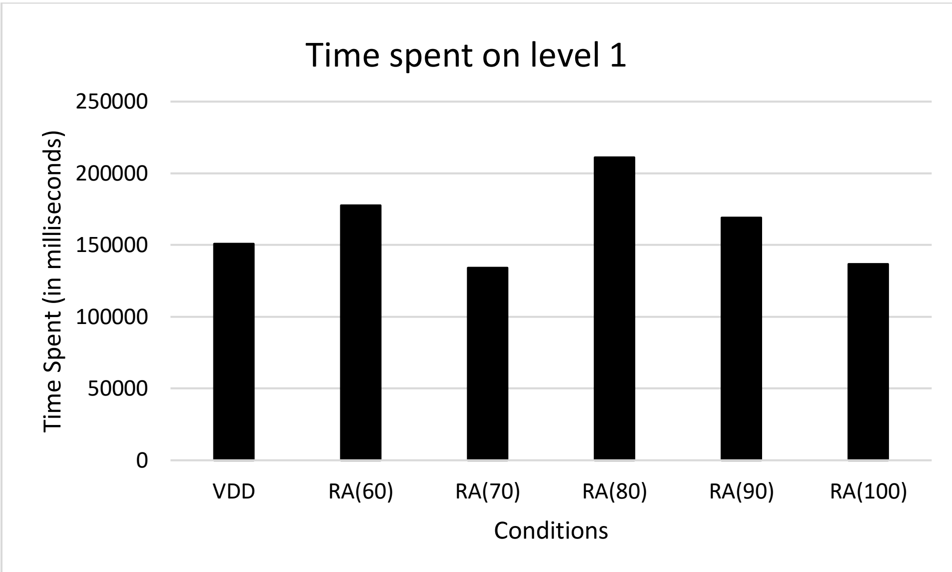


Figure 1A. Time spent on the first level

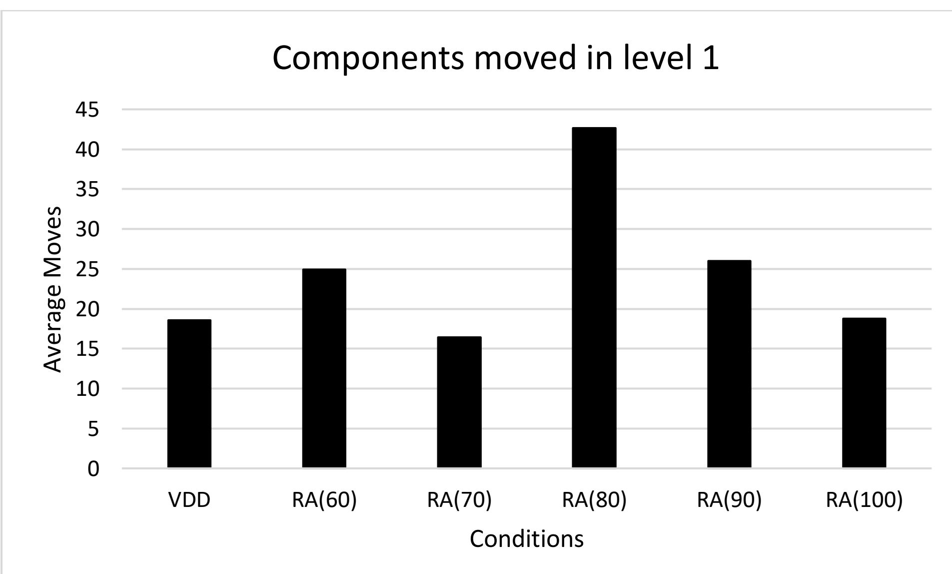


Figure 1B. Components moved in the first level

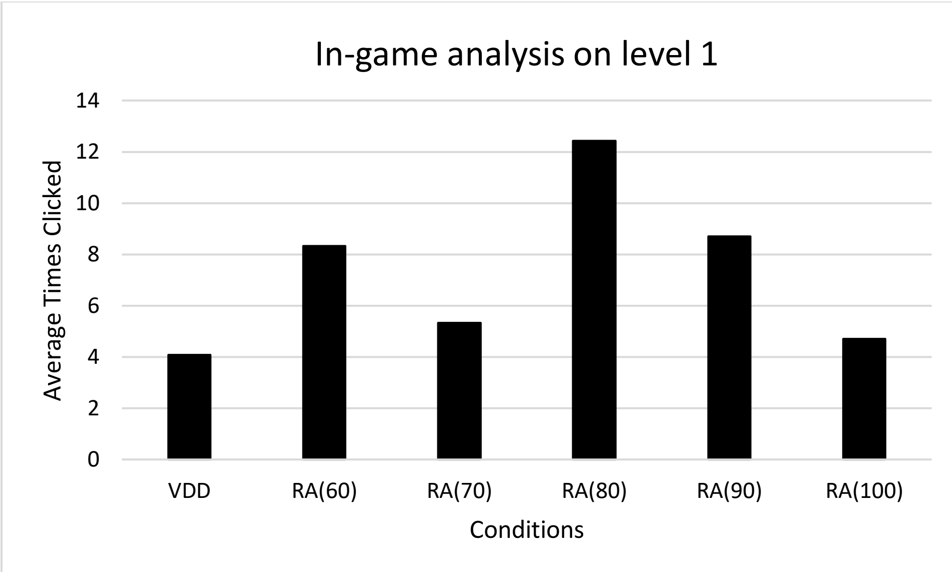


Figure 1C. Amount of times system was analyzed in the first level

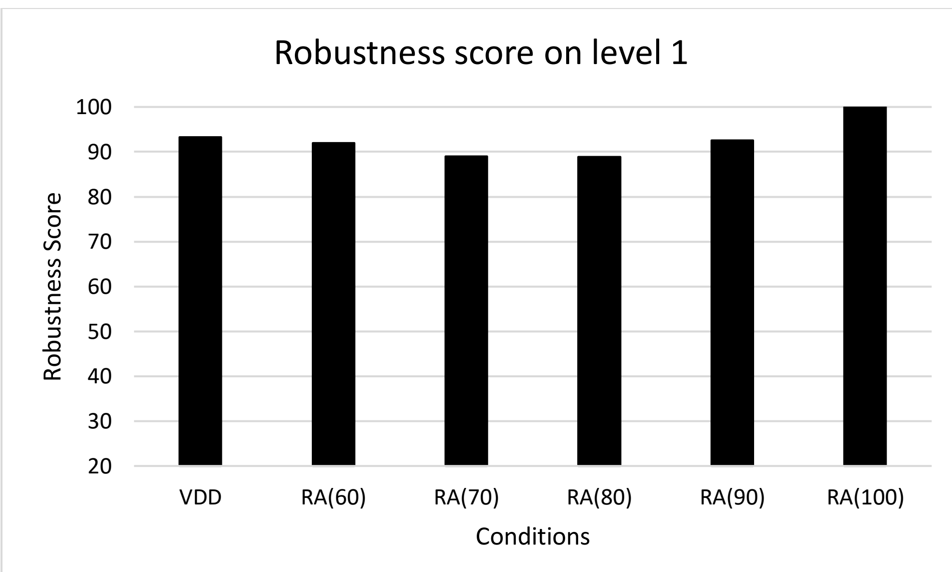


Figure 1D. Final robustness score for the first level

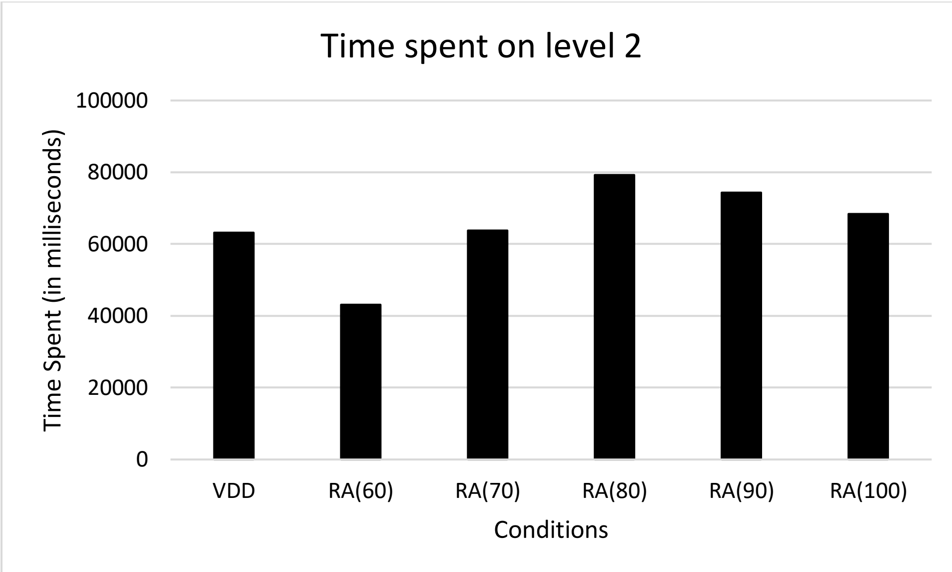


Figure 2A. Time spent on the second level

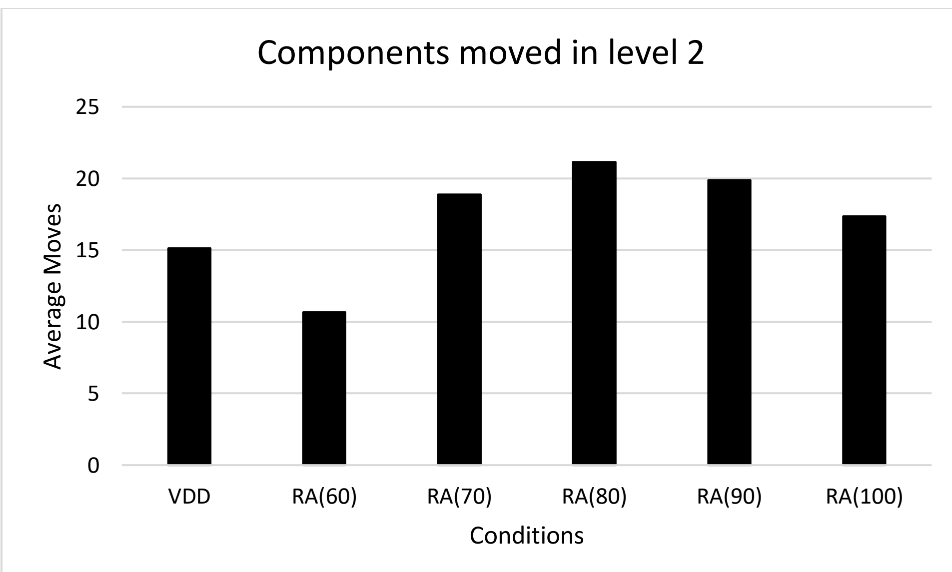


Figure 2B. Components moved in the second level

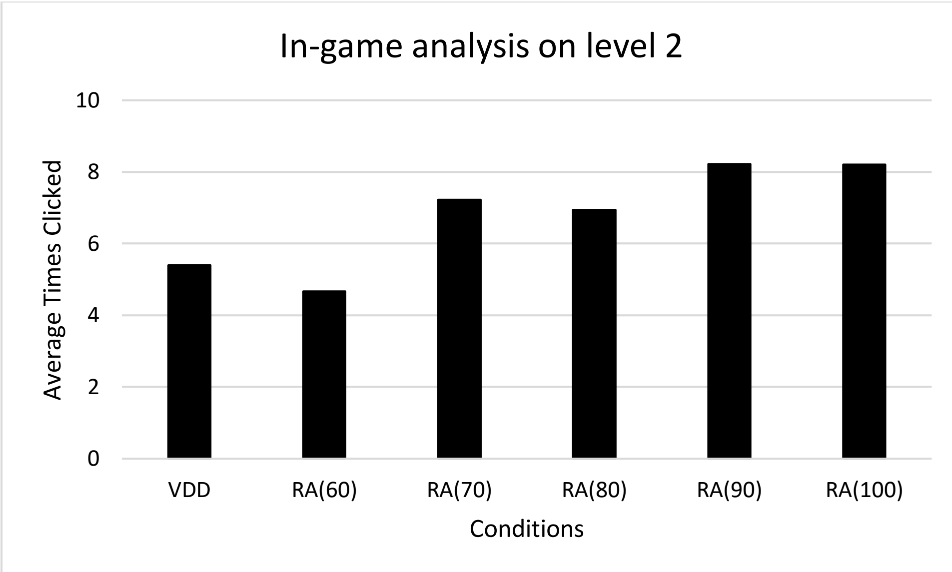


Figure 2C. Amount of times system was analyzed in the second level

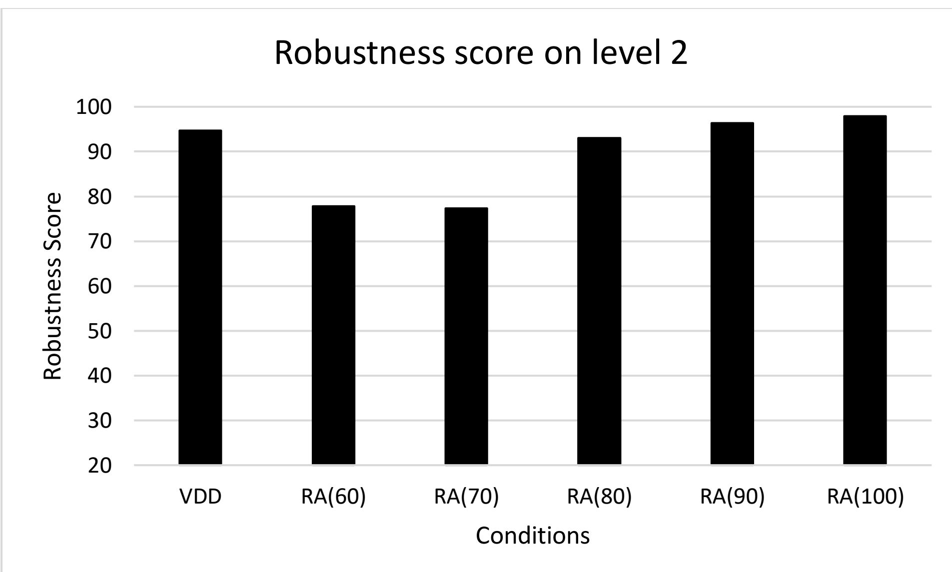


Figure 2D. Final robustness score for the second level

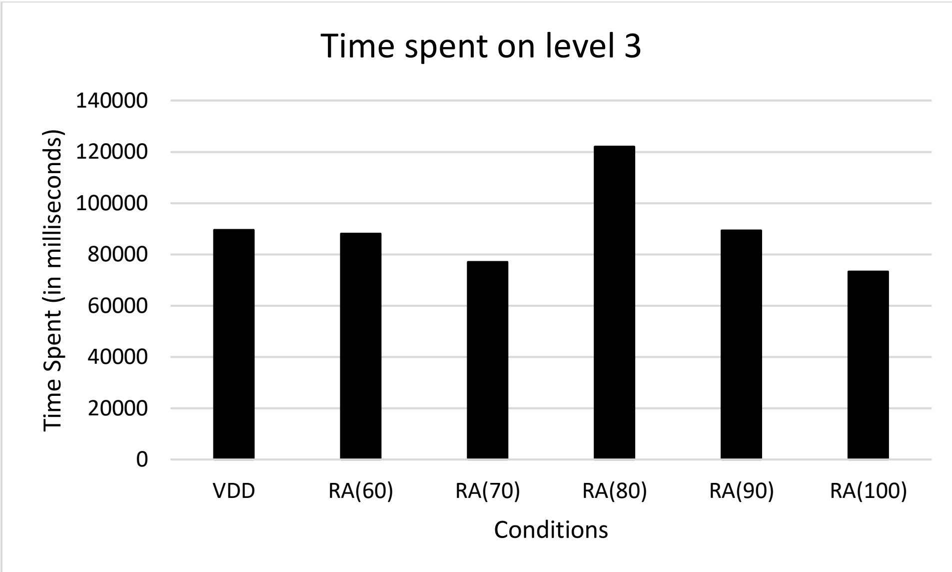


Figure 3A. Time spent on the third level

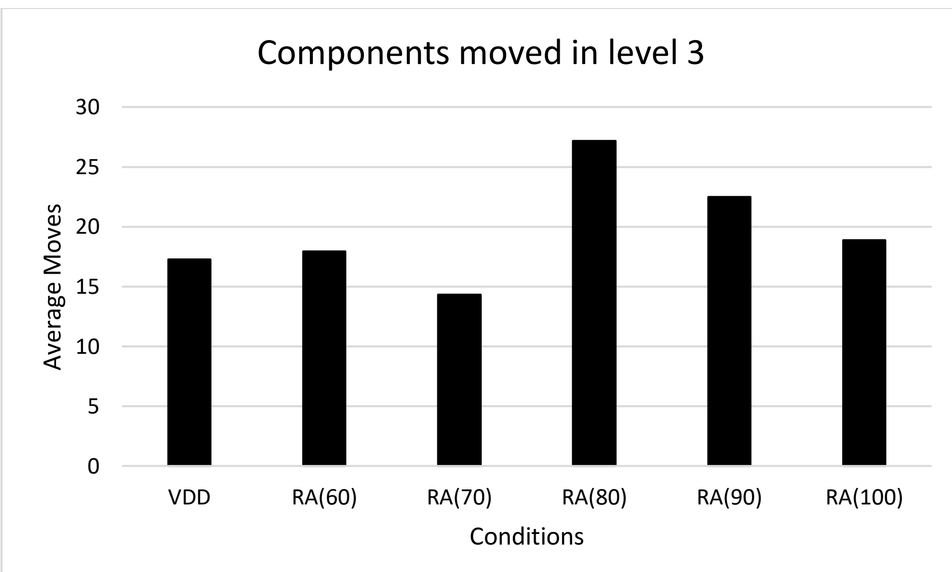


Figure 3B. Components moved in the third level

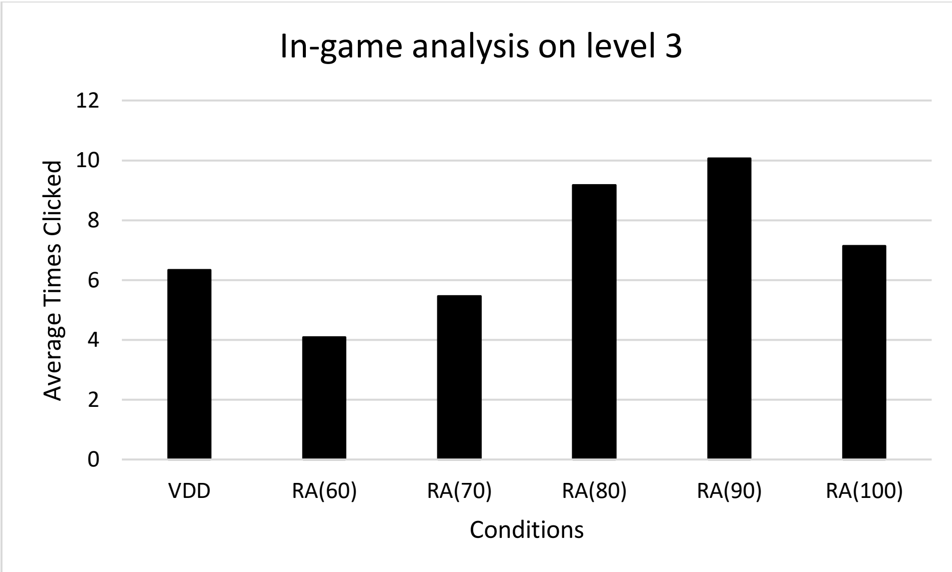


Figure 3C. Amount of times system was analyzed in the third level

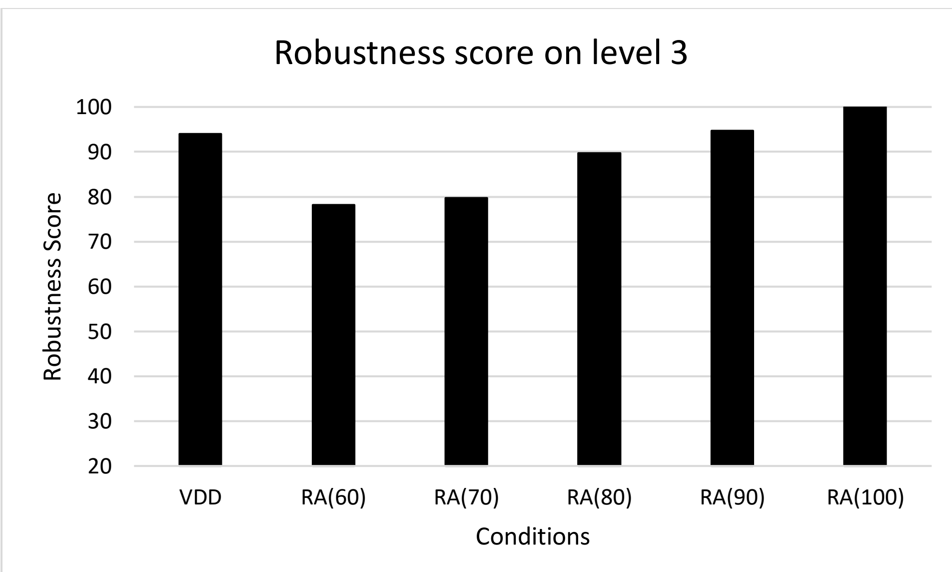


Figure 3D. Final robustness score for the third level

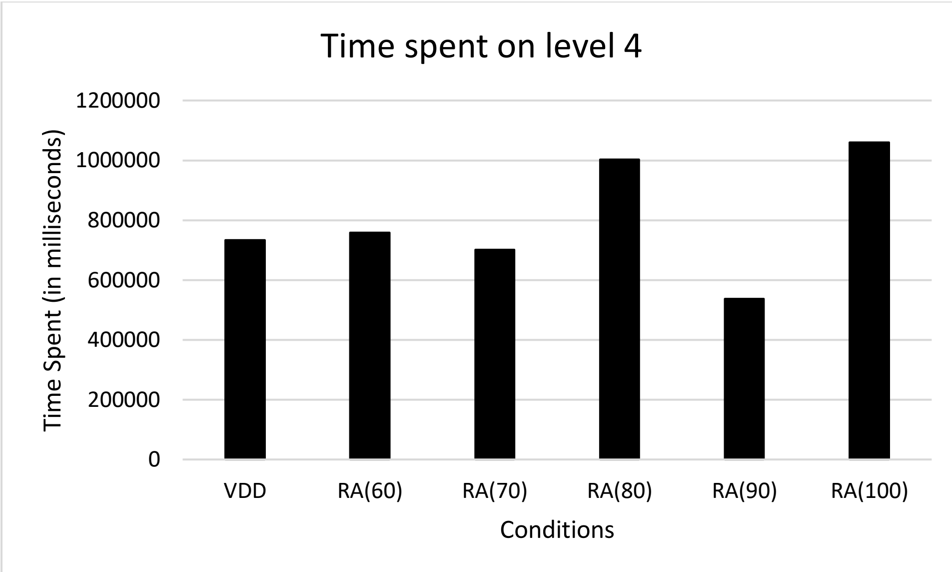


Figure 4A. Time spent on the fourth level

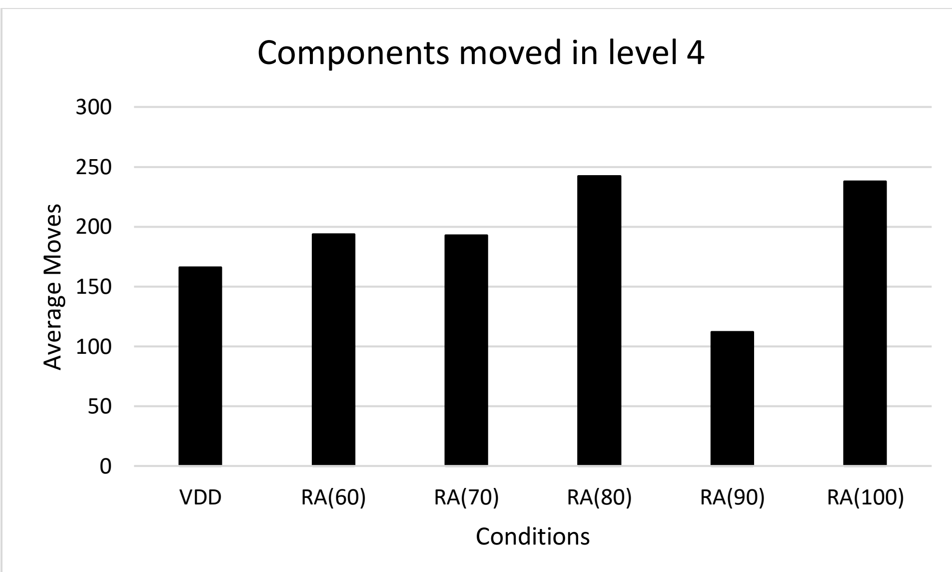


Figure 4B. Components moved in the fourth level

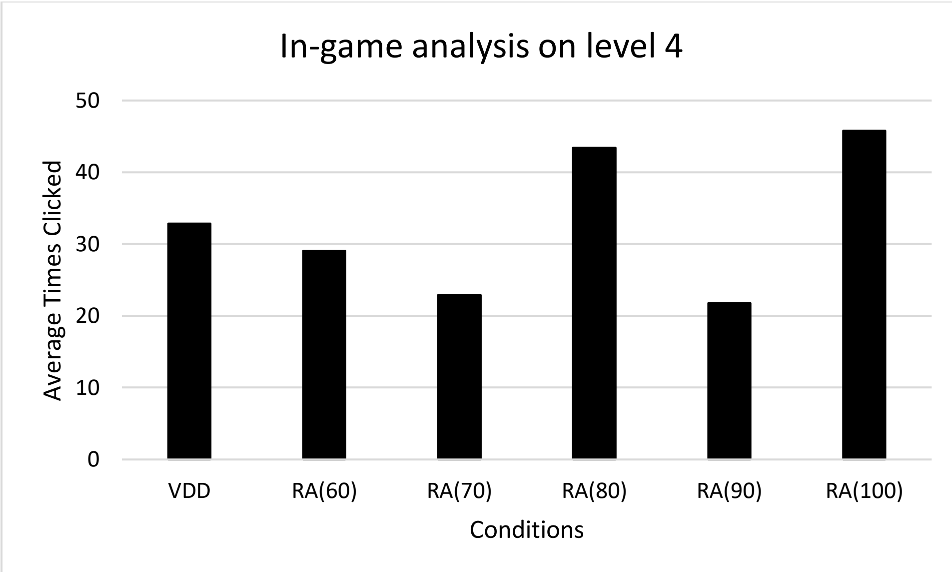


Figure 4C. Amount of times system was analyzed in the fourth level

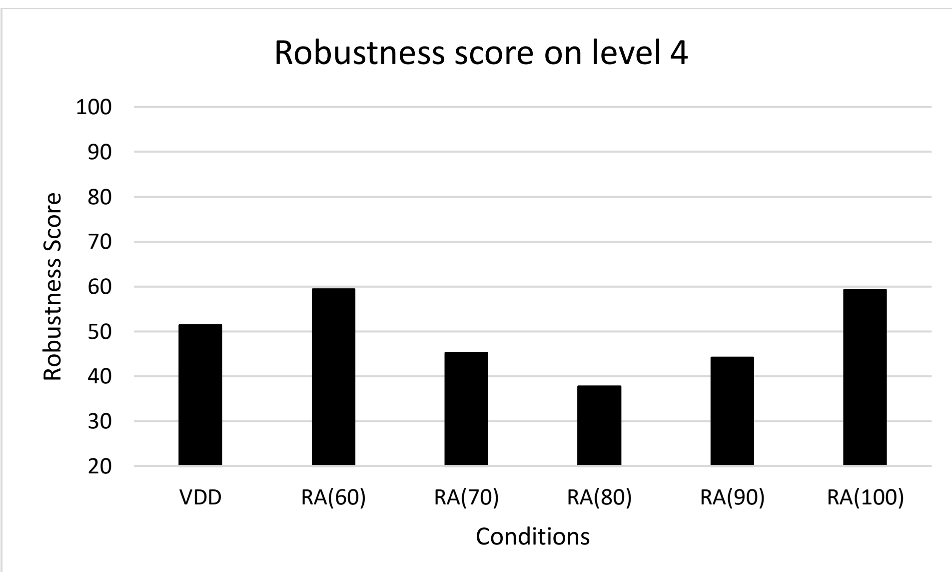


Figure 4D. Final robustness score for the fourth level