University of Nevada, Reno

Active Peer Pressure in Human-Robot Interaction

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Computer Science and Engineering

by Athanasia Katsila

Dr. David Feil-Seifer / Thesis Advisor

December 2018

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

ATHANASIA KATSILA

Entitled

Active Peer Pressure In Human-Robot Interaction

be accepted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

Dr. David Feil-Seifer, Advisor

Dr. Monica Nicolescu, Committee Member

Dr. Adam Kirn, Graduate School Representative

David W. Zeh, Ph.D., Dean, Graduate School

December, 2018

Abstract

This work investigates how conformity in human-robot groups can be manipulated by the robots' ability to actively apply coordinated peer pressure via voicing their agreement or disagreement with the human's selections in a visual task. Through a combination of kinesics and vocalics, we created an environment where the human subject became aware of the fact that they were not just actively being observed (by employing synchronized robot body and camera gaze motion), but were also under the active judgment and criticism of their robot peers. Our experiments consisted of two studies, a 2×2 and a 2×1 that considered a combination of possible conditions such as the sequencing of expressing opinions in the group and the number of robot peers, as well as the difference between unambiguous, ambiguous, and even duplicate selections in the used visual indicators. Using statistical analysis we show that the application of the new idea of "active peer pressure" manages to achieve increased conformity across many of the considered conditions.

Acknowledgements

This Thesis was accomplished thanks to the support and faith of Dr. David Feil-Seifer, who gave me the opportunity to work alongside him and entrusted me with the resources of his lab. My source of inspiration to design and conduct this Thesis has mainly been due to the exposure that I gained to the field of Human Robot Interaction through Dr. Feil-Seifer's classes. Without his contribution I wouldn't have been able to come in touch with the interesting field of HRI and `its fascinating opportunities.

I would also like to thank the undergraduate students of the lab who were willing to help me whenever I needed their assistance in the experimental process.

Finally, I would like to thank my parents and my sister who have been there to support and encourage me in every moment of my life, as well as my friends for believing in me and giving me strength in every step of this journey.

Table of Contents

List of Figures

Chapter 1

1.1 Introduction

Robots and their skill levels today indicate that they will take on more complex roles in everyday life -including social life- activities. A large number of robots already exists in households, workplaces, and cities, performing necessary tasks such as house-cleaning, warehouse management, package delivery, and more. But also social interaction robots are slowly becoming available in the form of toys such as robo-dog companions, etc. Their integration into broader aspects of human life can influence the behavior and the choices of humans in the future.

The main motivation of this study is human conformity and the effect of peer pressure on it, when the “peers” are robots. Psychology and behavioral sciences have examined this phenomenon on a human-to-human basis, for the negative outcomes [1]–[4], but also for positive ones such as the promotion of cooperation [5]. Research in peer pressure and its manipulation to affect human conformity originate in the ’30s and ’50s with the famous Sherif and Asch experiments [6], [7], showing that people can openly agree with group opinions even when they are obviously wrong, due to the social nature of the act. Since then, the field of Human Robot Interaction (HRI) [8] has also focused on human behavior and psychology in relation to robots. Examination of human-robot trust, emotional response, affection, etc [9],

[10] have already taken place, but also important studies exist that investigate how and if "robotic peers" can affect conformity. Such studies first directly recreated the original Asch experiment using humanoid NAO robots [11] without replicating the original effect, while attempting to extend results to linguistic imitation [12]. Traditional Human-Robot Interaction research into gaze control [13] have also been used to apply psychological pressure, with recent results showing that synchronized robot attention affects the perceived level of pressure [14], but without demonstrating an effect on error rates. On the other hand, more recent research has managed to show that modifying certain key characteristics of the experiment such as the participants' age group [15] can successfully affect conformity. Other studies have shown that examining key factors, such as human-robot trust [16], is necessary to manage achieving a degree of conformity.

In this study we investigated the effects of the novel concept of "active robot peer pressure" while using an adult-age group of participants. We started from the basis of an Asch-like experiment [7], where the participant (human test subject) is placed within a group, a set of visual indicators is presented to all on a screen, and each group member is asked in sequence to observe the query indicator and find its match from a list of possible alternatives. Similar to previous experiments [11], [14], we implemented a pre-scripted logic in our robots to create circumstances where the robot group members all agree, and either a) declare the correct answer, or b) "lie" and give a specific wrong answer. Through this novel idea we found that a) error rates and conformity to the opinion of robot peers increases when the participant is faced with highly ambiguous cases and "active peer pressure" is being applied, and b) conformity doesn't increase even further if the group consists of seven robot confederates.

Our key differentiation and contribution to achieve this result was to create 2 conditions of different degrees of induced peer pressure. For the basic level of peer pressure we manipulated the robots' unanimity, cohesion, inter-member imitation and adaptation [17]–[20] through a) their agreement across all questions, and b) their synchronized direction of attention towards the human when it was time to express their selection. For the increased degree of peer pressure we implemented a combination of vocalics and kinesics to emphasize and enforce the robots' opinions [21], [23] through c) a voice coming from each robot which actively expressed its agreement or disagreement with the participant's opinion. At the same time we designed the experiment such that the robots had individualistic characteristics such as voices, names, personalities, and, to some extent, appearance.

Moreover, we considered different conditions based on the sequence of expressing opinion in a group, with the human either answering a) last after all the robots, or b) in the middle of each sequence. The logic behind this was that the participant could feel separated due to the strong cohesiveness of the robot-only group if they were only asked for their opinion after the group has finished. To achieve conformity the illusion of an evenly treated human-robot group should be maintained, and therefore asking for the human's opinion asked as "one in the group" attempted to manipulate this perception.

Finally, we performed an additional study which investigated the additional effect of including an increased number of robots as part of the group. More specifically, after establishing the effect of the previously mentioned conditions (degree of peer pressure and sequence in group) we selected the most successful combination and examined whether we can increase conformity even further by increasing the number of robots in the group.

1.2 Specific Approach

1.2.1 Task Design

The considered task was similar to the one in Asch's work [7] which was also used in other HRI experiments [11], [14], [15], in order to have a common reference and examine the effect of our novel "active robot peer pressure" behavior. This consisted of 3 line choices (A-C) as shown in Figure 1-1, and one line to be matched (?) from one of the available choices. The line to be matched was either 100, 150, or 200 [mm], and the lines to be matched against it differed within a range of ± 25% all the way down to ± 2.5. This is also the main idea behind the original Sherif experiments [6] when conformity effects were being examined.

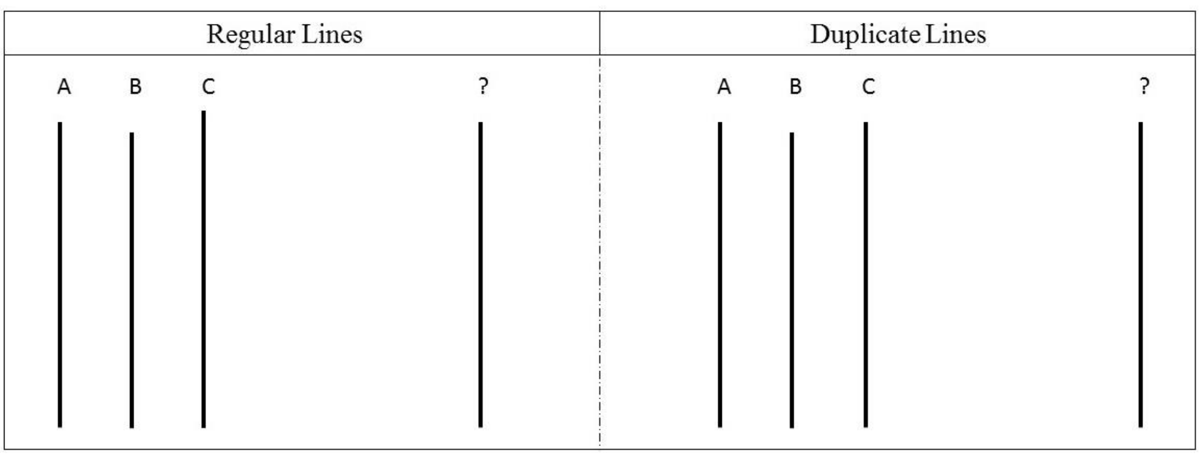

Figure 1-1 Our Line Task variations, Left side: All three (3) available options are different, and the appropriate choice is unique (only A is correct), Right: Two (2) of the available options are the identical (either A or C is correct).

A distinct differentiation from similar experiments [7], [11], [14] is that we used not only unambiguous and ambiguous questions, but we also introduced a number of "duplicate" ones (shown in Figure 1-1), to examine the special case of a human's conformity to the robot group opinion when the alternatives are indistinguishably similar.

Each sequence (for a single participant) consisted of 15 different line-matching questions. All were taken from a Baseline of 40 lines which was established by examining the responses of a number of people to the line selection task without the presence of robots. This corpus was examined and divided in similar principles as the setup presented in [11] into **4 unambiguous** lines, another **6 highly ambiguous** and hard to distinguish lines, and finally **5 "duplicate"** lines.

1.2.2 Number and Type of Robots

It is established in literature that unanimity and group size in cases of humans affects conformity [7], with an increasing effect up to a group size of 3 people and then remaining nearly constant. On the other hand, HRI research with adults has shown that using more humanoid robots (4 in [11], or even up to 6 (plus 1 "questioning" robot) as in [14]), no peer pressure effect has been achieved yet. We therefore selected to test the effect of "active robot peer pressure" against conditions of 4 and 6 robot peers.

We used Pioneer P3-DX robots [25] as a mobile base, which are capable of rotating in-place where they stand. We used this motion to imitate human body motion when directing attention takes place, and do so with emphasis. On top of every robot we placed a Sony EVI-D70 Pan-Tilt-Zoom (PTZ) camera. By controlling the orientation of the camera we imitate human head motion when the human directs their attention towards a subject of interest, either the visual indicators on the screen, or the participant. Also, each robot had a speaker which was used to play back pre-recorded sounds (humans actors talking) and imitate human voice. A Raspberry Pi mini computer was connected to the mobile base, the camera, and the speaker, and was used to control each robot. Finally, to make the robots more individualistic and life-like, each one apart from its own male or female voice (to avoid gender bias the final results), also had a distinct personality when disagreeing (expressing either a) calm disagreement, b) irritated disagreement, or c) mockery by laughing out). We also used pairs of plastic eyesight-corrective glasses placed on top of the robot's camera lens (to convey the idea that the camera represented its face), and we also employed human names and had our robots introduce themselves to each participant by name.

1.2.3 Environment and Implementation

We used a large centrally placed projection display to show the visual indicators, with the participant sitting at a clearly visible distance, and the robots placed on either side. This arrangement is shown in Figure 1-2, and its goal was to make sure that the participant is not at the edge (like an outsider) but is in the middle of the group, surrounded by the robots. At the same time, the participant could immediately observe when all the robots were directing their attention towards them by turning their bodies-and-cameras because the robots were always in their field of view. This couldn't be achieved with the robots and the participants placed in a line side by side as in other the experiments.

Figure 1-2 The environment configuration of our experimental studies, with the central large display, as well as the participant's seat at the middle, and the robots placed around them.

The robots were controlled through pre-scripted commands given by a human operator sitting in the back side of the same room, not directly visible by the participant. The sequences were automated in order to avoid any delays and repeated intervention that could be noticed by the participant. A sample sequence of actions such as the example one shown in Figure 1-3 was used, leaving a small role for the human operator to play: Either a) start a new sequence when it is time, or b) trigger the appropriate "answering" script or the "agreeing"-or-"disagreeing" behavior for the entire group of robots after the participant had responded. A number of 5 possible agreement playback voices (and another 5 disagreement ones respectively) could be randomly selected and played-back by each robot, to avoid always repeating the same reaction and creating disbelief. Also, display messages were shown on the screen such as a red-letter prompt for the participant, marking "Please answer now", to help maintain the order and sequence of the experiment.

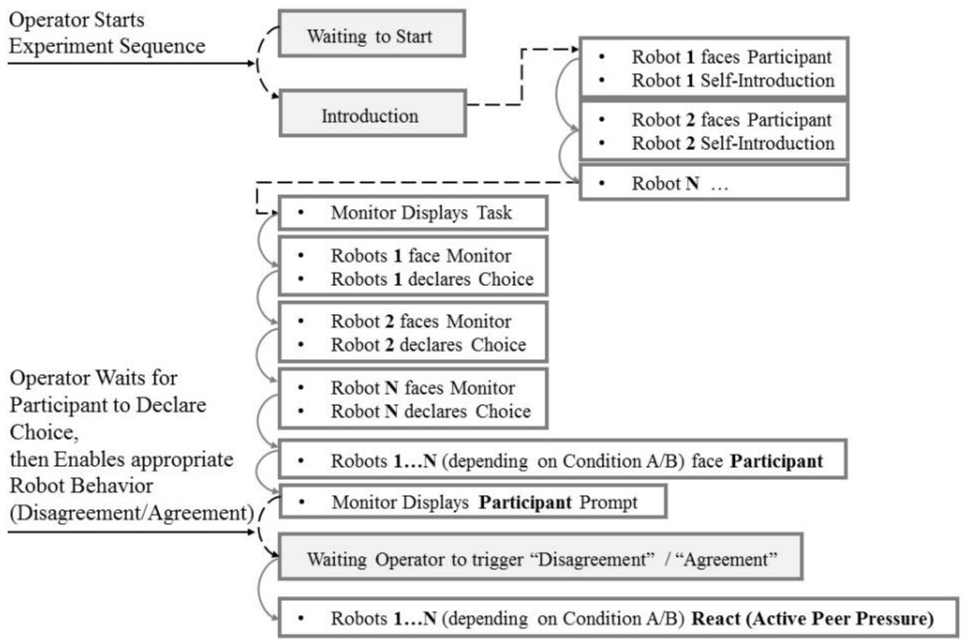

Figure 1-3 Example sequence of scripted actions during the experiment. The automation of the sequence requires a minimal amount of operator interventions for the successful execution.

In order to manage this complex set of behaviors and their proper timing, the experiment design employed a software architecture which used ROS to achieve synchronized behaviors between all robots. This is shown in Figure 1-4. Each robot had a separate Raspberry Pi mini computer running ROS [26] and a custom developed node ("`pres_client`") which accepted a sequence (`ActionSequence.msg`) of commands with desired values. As soon as a robot received a new command, its body rotated to the desired `robot_orientation` (imitating body motion), its camera moved to the desired `camera_pan` and `camera_tilt` (imitating head motion to indicate direction of attention), and when the desired set-point was reached, if a valid `sound_to_play` was included in the message, it would be played back through the computer speakers (imitating voice to express opinion, or to apply "active peer pressure"). Communications with the robot and the PTZ camera were achieved by using the `ROSARIA` [27] package and the `libVISCA2` [28] library, and the `sound_play` [29] ROS package to play-back pre-recorded "`.wav`" sounds.

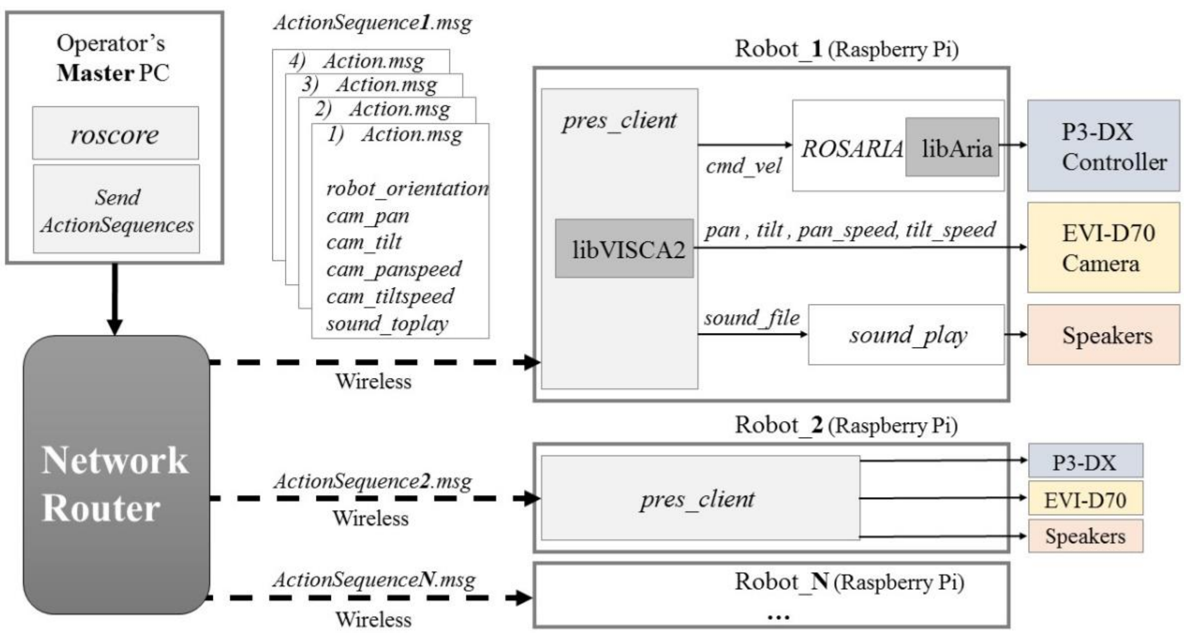

Figure 1-4 The software architecture of the experiment.

In order to achieve synchronized behavior between all robots, a ROS Master was running on the laptop computer of the human operator. This Master PC was connected to a wireless network router, and every Raspberry Pi mini- computer was also connected wirelessly to the same network, and had its own separate copy of the "`pres_client`" ROS node running.

The Master PC was responsible to send the scripted commands in proper order to each robot over this wireless connection, and then each command sequence was executed by each robot. In order to enable the human operator to give appropriate commands quickly and correctly, a Graphic User Interface (GUI) as shown on Figure 1-5 was developed, which automated the process of filling-in and sending each required `ActionSequence.msg` with simple button presses, based on some preset sequences (for instance instructing a robot based on its assigned name, e.g. “Amanda”, to a) first look at the participant and express “Disagreement” and then b) look away towards the screen again without saying anything).

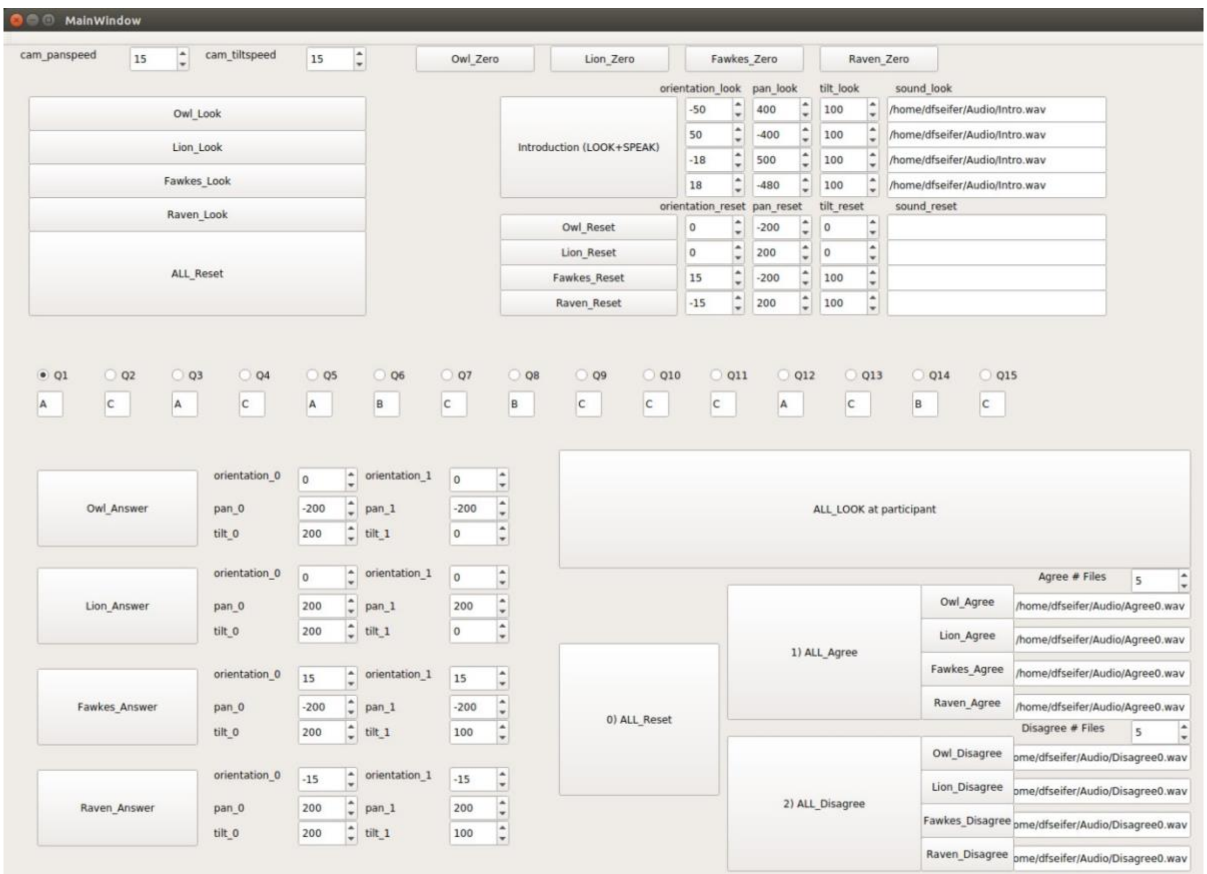

Figure 1-5 The Graphic User Interface that was used to control the flow of the experiments.

Chapter 2

2.1 Experiment Design

The experiment design considered 3 phases: a) the pre-experimental phase, where the participant was asked to answer some demographic questions b) the actual experiment, and c) the post-experimental phase, where the participant was given a questionnaire to fill out about their perception of robots. The goal was to apply different degrees of pressure through the robots' behaviors, and first investigate their effect on the humans' conformity by analyzing their answers in each situation, and secondly record the level of perceived pressure. Two studies were designed in order to test the conditions of a) robotic active peer pressure, b) the human's placement in the middle or the end of the group, and c) an increased robot group size.

2.2 First Study (2×2)

2.2.1 Conditions

Our first study was 2×2, between subjects with conditions **A** and **B**, which included subconditions **I** and **II**. In order to do that, we equally partitioned an experimental sample of 20 participants into 2 groups of 10 each. In each condition the group was equally subdivided (5 and 5) into two subgroups. The first subgroups **A-I** and **B-I** experienced robot behaviors that included only synchronized robot attention, while the other subgroups **A-II** and **B-II** experienced Active Peer Pressure by generating voices from the robots as well.

The specific experimental configuration contained 4 robots as shown in Figure 2-1, which also includes a visualization of the experiment sequencing for conditions A and B.

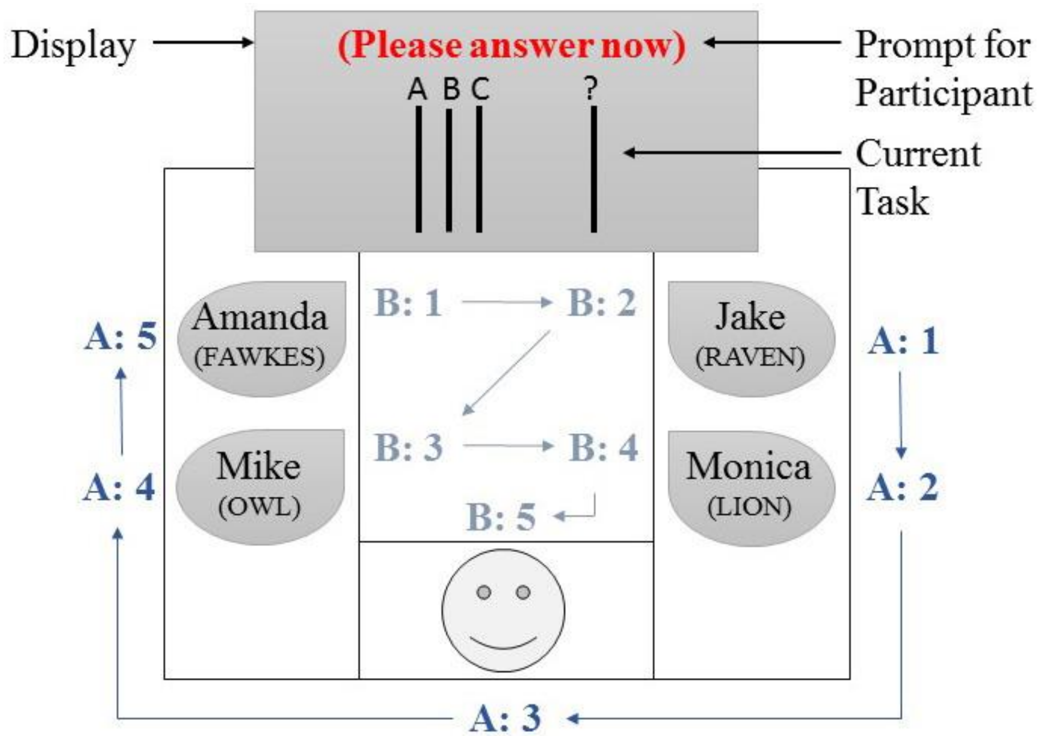

Figure 2-1 Experiment configuration and robot sequencing for our First Study (2 × 2) which considered conditions A-I/A-II/B-I/B-II.

Condition A - Human at the End: The robots would all express (taking turns one after the other as indicated in Figure 2-1) their choices, leaving the participant for last. Once it was the human's turn to express their opinion one of the two possible subcondition routes was followed:

• **Subcondition I - Simple Pressure**: In a synchronized manner, all robots would direct their attention towards the participant by rotating their bodies and cameras. Once completed, the participant would be prompted to give their answer, and the robots would then return to their original configuration to begin a new sequence.

• **Subcondition II - Active Peer Pressure**: The robots would direct their attention towards the human in the same way. In this case however, once the the participant gave their conforming or non-conforming opinion, the robot "Disagreeing" or "Agreeing" reaction would be triggered. As mentioned, differentiated reactions would come out of every robot to imitate a personality type. Some would disagree in an irritated tone, some calmly, and some would even laugh.

Condition B - Human in the Middle: The participant's answer was requested as part of the group (in the middle of the sequence). More specifically, half the robots (2) expressed their choices, and then the participant's turn would come up, leading again to two possible subcondition routes:

• **Subcondition I - Simple Pressure**: With the same synchronized motion all the robots would direct their attention towards the human, who would then be prompted for a reply. Then, the remaining (2) robots would in sequence turn towards the screen and express their opinions as well.

• **Subcondition II - Active Peer Pressure**: Again all robots would turn towards the human. As soon as the participant was prompted and replied with a conforming or non-conforming answer, again all the robots' "Disagreeing" or "Agreeing" reactions would take place. Then, the remaining (2) robots would in sequence face the screen and answer.

2.2.2 Hypotheses

Based on these tools, we set out to prove or reject the following hypotheses:

H1: The human will conform more often to the opinion of the group in cases where 2 correct choices ("duplicate" visual indicators of the exact same length) are possible.

H2: Conformity will be increased in the Active Peer Pressure condition than in the Simple Pressure one.

H3: Human-perceived pressure will be significant for the case of robots that apply Simple Pressure (move and direct their attention towards the human), and even higher when Active Peer Pressure is applied.

H4: The human's placement in the middle or at the end of the group sequence will affect conformity as well as the perceived level of pressure.

The reasoning behind these hypotheses is based on the existing literature of the field:

a) Regarding hypothesis H1, ever since Asch's extensions of his original experiment [17,22] it was shown that the uncertainty of the task affects conformity since lines that were made more similar in length and were harder to judge resulted in increased conformity. In our case we examined taking uncertainty to the extreme, by using exactly duplicate lines, since other experiments with robot groups that used highly ambiguous (but still different) lines [14] did not show any effect on conformity.

b) Regarding hypotheses H2 and H3, again since Asch's extension experiments [17,22] it has been shown that conformity depends on group pressure levels. Conformity was decreased when the participants were allowed to answer in private without experiencing the normative influence of a fear of rejection by the group. In our case we used active peer pressure to make clear to the participant that the robots are truly aware and understand whether the human agreed with the group opinion, and actively express their agreement or disagreement to indicate group acceptance or rejection.

c) Regarding H4, early research has shown the role of the seating position on the interaction of small groups [25], and more specifically the observed effect appears especially in "task" oriented sessions. In our case we used the idea of seating and querying the participant in the middle of a sequence to make them feel more as "one of the group".

2.2.3 Experimental Process

Before the first session with each participant, a brief description of the experiment's procedure was given to them, without fully disclosing the goal of the study. They were informed that they would be assisting in an experiment aiming to examine robot perception of visual indicators. Also, they signed appropriate consent forms and then filled out a demographics survey in order to establish factors such as their level of comfort and familiarity with technology and robots in particular.

Figure 2-2 Instances of the experimental setup for our First Study (2 × 2), showing the participant and the robots.

The experimental configuration and an instance containing a participant are shown in Figure 2-2. After the participant sat at the table in front of the screen, the experiment was initiated by the operator. The robots took turns and performed the "Self-Introduction" phase where they rotated their bodies and cameras to face the human, greeted them, and introduced themselves by name. One greeting message, given here for reference, was: "Hi, I'm Amanda. Welcome!".

The same sequence order that was used for self-introduction was also used throughout the experiment, to accustom the participant with the complete experimental sequence. Then, all robots would reset themselves to their default original pose facing the screen. The operator would change the slide moving to the 1st set of lines. The robots one after the other would a) rotate their camera with a quick motion to face up and towards the screen center, b) declare their choice by voice (“Line A”, “Line B”, or “Line C”), and then c) move their camera to the default position. This action sequence aimed to emulate human head motion, and helped maintain a clear order of turn-taking around the table. In Condition A, all the robots would complete this sequence until the participant’s turn came up, while in Condition B, only two of them would first do so.

Once it was time for the human to answer, all robots would simultaneously rotate their bodies and cameras to face the participant, to achieve synchronized direction of attention. Afterwards, the experimenter would enable the red-lettered prompt on the screen, and the participant would declare their choice. Depending on the experimental Condition, the robots would either be reset by the experimenter’s command to the default original pose directly (Subcondition I - Simple Pressure), or they would all be triggered to react and express simultaneously either the “Agreement” or the “Disagreement” behavior (randomized out of the 5+5 available ones to avoid repetition). Subsequently, all robots would reset themselves to their default original pose facing towards the screen, the experimenter would move on to the next set, until the 15 sets of session were completed. For reference, we also give some specific reactions used by one robot:

Amanda - Agreement: Yes!, Awesome!, Perfect!, You're right!, Very good!, **Disagreement**: That's not right!, Seriously?, Of course not!, <Laughter>, Wrong!

After each session's end, the participants were given a questionnaire to fill out, with 5-point differential scales in order to measure the subjective items of "perceived level of pressure", "robot impressions" (e.g. human-likeness), as well as some descriptive questions such as "Did you ever change the answer you were going to give because of the robots" during the experiment. These allowed to evaluate the influence and quantitatively record the degree of pressure that we managed to achieve by applying Active Peer Pressure.

2.2.4 Participants

All our participants were recruited from the student base of a university campus, were aged from 19 to 32 years old, and consisted of undergraduate, graduate, and PhD-level students. Initially, the baseline sample was established using 6 people without any robot presence, to derive answers from a complete corpus of 40 possible lines, and select the final 15 lines which were then used in the experiments. A total of 20 additional people were recruited for the first experiment. Other relevant experiments in robotics literature have also used from 20-people samples [14] (each used in 3 sessions of different conditions however), up to 60-people ones [11] (excluding the verbal experiment baseline) respectively.

2.3. Second Study (2 × 1)

2.3.1 Conditions

Our second study was 2×1, between subjects, with conditions AB-II and C- II, and came after the completion of the first. As we will present in Chapter 3, the effect of group sequencing (Conditions A & B) was not significant, while the Active Peer Pressure (Subcondition II) was found to be significant across the obtained results. Given these, we ran a second experiment adding two (2) extra robots in order to examine whether this would have even more significant results in conformity. Since group sequencing didn't have any significant effect on conformity as already mentioned, for the second experiment we randomly opted to have the participant declare their choice last in the sequence.

The specific experimental configuration for C-II contained 6 robots (2 additional ones) as shown in Figure 2-3, which also includes a visualization of the experiment sequencing.

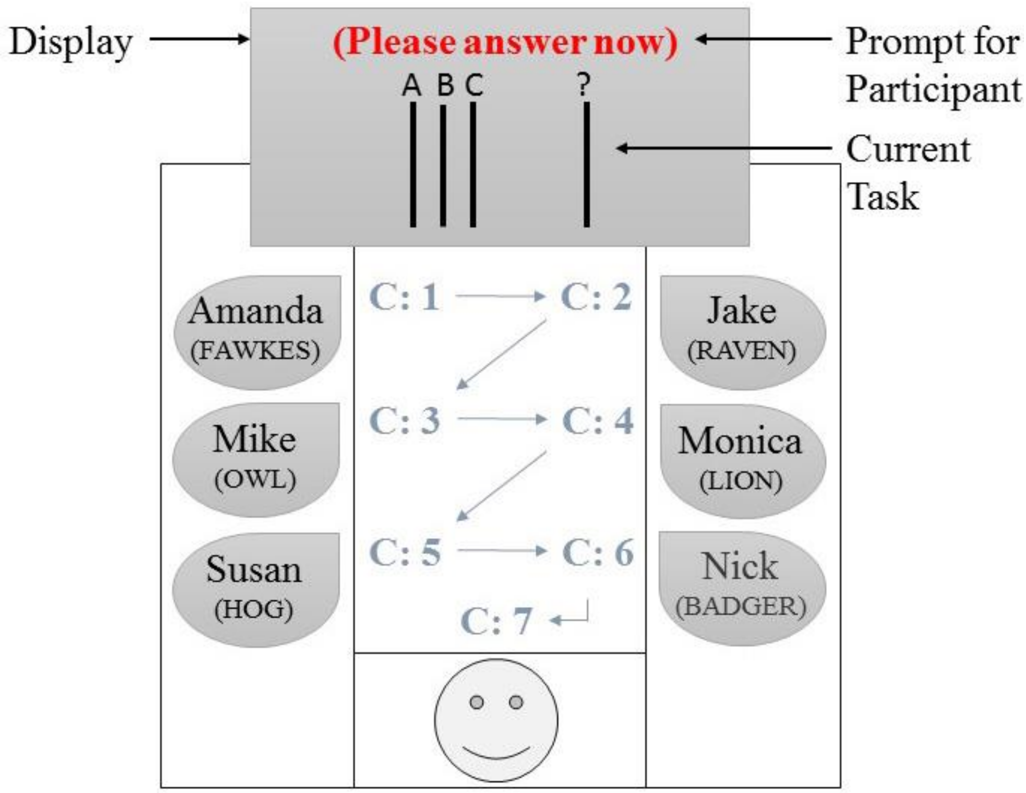

Figure 2-3 Experiment configuration and robot sequencing for our Second Study's (2 × 1) additional experiments, which considered also condition C-II.

For this study we considered a set of Active Peer Pressure-only conditions:

• **Condition AB-II - 4 Active Pressure Robots**: This set is formed by grouping together the results of the **Active Peer Pressure subconditions** of the first study which involved 10 participants. Across all these, 4 Robots were used.

• **Condition C-II - 6 Active Pressure Robots**: We used a significantly increased number of robots (+50%) with the Active Peer Pressure condition (due to its significant effect on conformity), with a sample of 10 more participants. The goal was to find a potentially additive effect of robotic presence, since with the introduction of Active Peer Pressure we had managed to achieve a degree of conformity.

2.3.2 Hypothesis

For the second part of our experiment we made the following hypothesis:

H5: Increasing the number of robots when Active Peer Pressure is applied, will increase conformity in both the cases of regular as well as "duplicate" visual indicators.

Regarding H5, the reasoning behind it is based again on Asch's experiments extensions [17,22] which showed that when the majority group (number of confederates) was increased, conformity followed, but only up to the point of 3 confederates. In our case, we attempted to further increase our robot group size to examine whether conformity would be affected as well.

2.3.3 Experimental Process and Participants

The experimental process was exactly the same as in the Active Peer Pressure conditions of first study, with the only difference that in the current condition were added two extra robots in the robot group. The experimental configuration and an instance containing a participant are shown in Figure 2-4.

The 15 visual tasks and their order in the task were exactly the same, so essentially it was an exact copy of the first experiment with just two extra robots, in order to avoid any bias in our results. 10 more participants were recruited in order to obtain the same number of participants and be able to extract more reliable results. The participants that were recruited were again undergraduates and graduates of the University of Nevada with ages ranging between 18-25 years old.

Figure 2-4 Instances of the experimental setup for our Second Study's (2 × 1) additional experiments, showing the participant and the robots.

Chapter 3

3.1 Analysis of Results

In our results we measure conformity based on the participants' results in the Baseline (conducted in complete absence of robots) and how they potentially changed to match the robots' answers during the tasks (executions with robots present). This means that the Baseline study forms a prototype of the answers that are expected to be given by the participants. Essentially, if the answers in the experiment have different tendencies than what was evaluated in the Baseline, and more importantly, if they tend to match the robots' ones, this is an indication that a degree of conformity has been achieved through our design. In order to statistically verify or reject this effect, we used regression analysis in the R programming language, which is a software environment for statistical computing. In the following section we present the statistical results for the two studies, and the assessment of the 5 different hypotheses. To provide an analysis of results we examined both the values reported by the participants in the questionnaires, as well as their choices which were recorded during the experiment.

3.2 First Study (2×2)

3.2.1 Experimental Results across Conditions

Figure 3-1 summarizes the results of the corresponding experimental sessions and provides at the same time for comparison purposes the responses to the corresponding 15 questions of our

Baseline sample (whose full original corpus contained 40 questions as mentioned). The top row shows the index of each of the 15 line sets, immediately followed beneath it with the correct answer.

		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Correct Answer		A	A/C	A	C	C	B	A	A/B	B/C	C	A	C	B/C	C	B/C
Robots Answer		A	C	A	C	A	B	C	B	C	C	C	A	C	B	C
Condition																
A - I	1	A	C	A	B	A	B	A	B	C	C	A	C	C	C	C
	2	A	C	A	C	C	B	A	A	C	C	A	C	C	C	C
Last &	3	A	A	A	C	C	B	A	A	B	B	A	C	C	C	C
Simple	4	A	C	B	C	A	B	A	B	B	B	A	C	C	C	C
	5	A	A	A	C	C	B	A	A	B	C	A	C	B	C	B
B - I	6	A	C	A	C	C	B	A	A	B	B	A	C	B	C	B
	7	A	A	A	C	C	B	A	A	C	C	A	A	C	C	B
Middle &	8	A	C	A	C	C	B	A	B	C	C	C	C	C	C	C
Simple	9	A	C	A	C	C	B	A	A	C	C	A	C	B	C	B
	10	A	C	A	C	A	B	A	B	C	C	C	A	C	C	C
A - II	11	A	A	A	C	C	B	A	B	C	C	A	C	C	C	C
	12	A	A	A	C	C	B	A	A	C	B	C	A	B	C	B
Last &	13	A	C	A	C	A	B	A	A	B	C	A	A	B	C	B
Active	14	A	A	A	C	A	C	A	A	B	B	C	A	C	C	B
	15	A	C	A	C	C	B	A	A	C	C	A	C	C	C	B
B - II	16	A	C	A	C	C	B	A	A	C	C	A	A	B	C	B
	17	A	C	A	C	A	B	A	B	C	C	A	A	C	C	C
Middle &	18	A	C	A	C	A	B	A	B	C	B	A	C	C	C	C
Active	19	A	C	A	C	A	B	A	A	C	B	A	A	B	C	C
	20	A	A	A	B	A	B	A	A	C	B	A	A	C	C	C
Baseline	1	A	A	A	C	A	B	A	A	B	B	A	C	B	C	B
	2	A	A	A	B	C	B	C	A	B	B	A	C	B	C	B
	3	A	C	C	C	A	B	A	A	B	B	A	C	B	C	B
	4	A	A	A	C	B	B	C	B	B	B	C	B	B	C	C
	5	A	A	A	C	C	B	A	A	B	C	A	C	B	C	B
	6	A	A	B	C	A	B	A	A	B	A	A	C	B	C	C

Figure 3-1 Summary of results of our First Study (2 × 2) including the Baseline and color-coding annotations.

In the cases of "duplicate" lines (sets 2, 8, 9, 13, 15) both appropriate choices are marked there.

In the following row we show the robots' selected response in each set, which was always unanimously expressed by all robots in the group. This is color- coded with: a) **green** marking the robots' selection in **duplicate** cases, and b) **red** marking the robots' intentional **mistaken** choices. Based on this color-coding, the cases where robots "**lied**" (sets 5, 7, 11, 12, 14) are immediately discerned.

This table then shows the summary results for all participants. These are grouped by Experimental Condition, which is shown on the left side. Condition A-I is human at the end with simple pressure, B-I human in the middle with simple pressure, A-II human at the end with active pressure, and B-II human in the middle with active pressure. Again color-coding is used to mark: a) with **green** the cases of "**duplicate**" line sets where participant's answer **matches** the robots**'**, b) with **red** the cases where robots "**lied**" and the participant's answer **matches** the robots' intentionally mistaken answer, and c) with **yellow** the cases where participants made "**mistakes**" irrelevant to the robots' answer in the set.

The color-coding for the Baseline study results shown beneath has a different context (since this study preceded the experiment and was performed in the absence of any robots). As mentioned, we relied on these results in order to decide on the final proper selection of the robots' answers that would contribute to our test. To make it more clear to understand, **green-**colored cases are the answers of **baseline** participants that we chose to be **matched** by our robots in the cases of **duplicat**e lines. These were intentionally selected to be clear minority answers in order to examine whether conformity during the experiment would shift results. **Red**-colored cases are again **baseline** participant **mistakes**, which also match our eventual selection for intentionally mistaken responses ("**lies**") used by our robots, both in **ambiguous** (sets 4, 5, 7, 10, 11, 12) and **unambiguous** (sets 1, 3, 6, 14) line cases. Intuitive visual comparison between Baseline and Experimental mistakes in cases where robots "push" towards a specific wrong answer can be performed on the table. **Yellow**-colored cases are **baseline** participant **mistakes** that the robots did "**not support**" during the Experiment. Accordingly, intuitive visual comparison of how robots "push" towards a specific right answer can be performed based on this color-coding.

3.2.2 Analysis of 2 × 2 Conditions A-I / A-II / B-I / B-II

First we examined the results of the 2×2 experiment that tested Conditions A-I, A-II, B-I, and B-II.

Observing the results table, it can be seen that there exists some degree of agreement in certain "duplicate" lines across these conditions. More specifically, in line set 2, in the Baseline 16% of participants opted for choice A, while in the Experiments a significantly higher 65% went for A, conforming to the robots' opinion. In line set 8, the Baseline's 16% more than doubled up to 35% in the Experiment, again showing conformity with the robots. Extraordinarily, in line set 9 no Baseline subject (0%) declared option C, while in the Experiment 70% conformed to the robots' expressed choice. This level of result was observed again in line set 13. In the final line set 15, a 33% Baseline was increased to a less significant 55%, which was however expected as it was a specially designed case: It was preceded by set 14 where the robots were intentionally made to “lie” to an unambiguous easy set to test whether “breaking the trust” in robots’ perception due to an obvious mistake would affect subsequent conformity. This was

		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Correct Answer		A	A/C	A	C	C	B	A	A/B	B/C	C	A	C	B/C	C	B/C
Robots Answer		A	C	A	C	A	B	C	B	C	C	C	A	C	B	C
Condition																
A - I Last & Simple	1	A	C	A	B	A	B	A	B	C	C	A	C	C	C	C
	2	A	C	A	C	C	B	A	A	C	C	A	C	C	C	C
	3	A	A	A	C	C	B	A	A	B	B	A	C	C	C	C
	4	A	C	B	C	A	B	A	B	B	B	A	C	C	C	C
	5	A	A	A	C	C	B	A	A	B	C	A	C	B	C	B
B - I Middle & Simple	6	A	C	A	C	C	B	A	A	B	B	A	C	B	C	B
	7	A	A	A	C	C	B	A	A	C	C	A	A	C	C	B
	8	A	C	A	C	C	B	A	B	C	C	C	C	C	C	C
	9	A	C	A	C	C	B	A	A	C	C	A	C	B	C	B
	10	A	C	A	C	A	B	A	B	C	C	C	A	C	C	C
A - II Last & Active	11	A	A	A	C	C	B	A	B	C	C	A	C	C	C	C
	12	A	A	A	C	C	B	A	A	C	B	C	A	B	C	B
	13	A	C	A	C	A	B	A	A	B	C	A	A	B	C	B
	14	A	A	A	C	A	C	A	A	B	B	C	A	C	C	B
	15	A	C	A	C	C	B	A	A	C	C	A	C	C	C	B
B - II Middle & Active	16	A	C	A	C	C	B	A	A	C	C	A	A	B	C	B
	17	A	C	A	C	A	B	A	B	C	C	A	A	C	C	C
	18	A	C	A	C	A	B	A	B	C	B	A	C	C	C	C
	19	A	C	A	C	A	B	A	A	C	B	A	A	B	C	C
	20	A	A	A	B	A	B	A	A	C	B	A	A	C	C	C
Baseline	1	A	A	A	C	A	B	A	A	B	B	A	C	B	C	B
	2	A	A	A	B	C	B	C	A	B	B	A	C	B	C	B
	3	A	C	C	C	A	B	A	A	B	B	A	C	B	C	B
	4	A	A	A	C	B	B	C	B	B	B	C	B	B	C	C
	5	A	A	A	C	C	B	A	A	B	C	A	C	B	C	B
	6	A	A	B	C	A	B	A	A	B	A	A	C	B	C	C

Figure 3-2 Summary of our statistical analysis variables for the First Study (2 × 2), with color-annotations.

analogous to the trust-related idea proposed in [16] with application to our special "duplicate" cases.

These observations on tendencies are not sufficient, and the results from such an experiment contain multiple parameters. To perform statistical verification of measures of significance from our results, we considered an Analysis Of Variance (ANOVA) approach. Figure 3-2 shows the variables that were used with reference to the previously presented results. For each line selection from a participant (each table cell), we denote whether it was the **SAME** one as the one given by the robots in the experiments with a **grey** color. Essentially, the table now represents a binary variable **SAME** ∈ **{0,1}**, with different contexts: within the Baseline it is the "freely" selected option at the absence of any robots, while in the Experimental conditions that the robots express their opinions, it additionally represents whether the human **agreed or disagreed** with the robot group. Additionally, the **green**-shaded overlays correspond to **DUPLICATE** line sets, the **yellow**-shaded ones to **AMBIGUOUS** line sets, and the **red**-shaded ones to line sets that the **ROBOTS_LIED**. These are also the additional categorical variables (binary again) that are considered for the purposes of statistical analysis, alongside the experimental condition variables **ACTIVE** (whether the robots exerted active peer pressure, **purple**-colored), and **MIDDLE** (whether the human as asked to express their opinion midway in the group, **cyan**-colored). Finally, a binary variable **EXPERIMENTAL** ∈ **{0,1}** is used to represent data coming from either the Baseline (EXPERIMENTAL=0) or the Experimental conditions that included the robots (EXPERIMENTAL=1).

Based on these statistical variables we first validate the obvious effect in getting **SAME** (agreeing) responses when using **DUPLICATE** lines. We perform 1-way ANOVA on the corresponding subset of data, and we evaluate the effect of being in the **EXPERIMENTAL** condition. In R:

```
# Subset data
DUPLICATE_SAME <- SAME[DUPLICATE==1]
DUPLICATE_EXPERIMENTAL <- EXPERIMENTAL[DUPLICATE==1]

# 1-way Analysis of Variance
res.aov <- aov( DUPLICATE_SAME ~
                  DUPLICATE_EXPERIMENTAL )
summary(res.aov)

                       Df Sum Sq Mean Sq F value   Pr(>F)
DUPLICATE_EXPERIMENTAL   1  4.604   4.604   21.18 9.94e-06 ***
Residuals              128 27.827   0.217
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

which statistically verifies this effect (p<.001). Based on this fact, **Hypothesis 1 was supported** in our First Study. Additionally, by running 3-way ANOVA to evaluate the effects of the ACTIVE and MIDDLE conditions, we clearly see that no statistical significance is attributed to them, i.e. using a robot group that expresses a preference over "duplicate" lines, a large effect on conformity is achieved, with the other parameters of the group (active pressure robots, placement of human, etc.) play no significant role.

```
# Subset remaining data for DUPLICATE lines
DUPLICATE_ACTIVE <- ACTIVE[DUPLICATE==1]
DUPLICATE_MIDDLE <- MIDDLE[DUPLICATE==1]

# 3-way Analysis of Variance for DUPLICATE line subset
res.aov <- aov( DUPLICATE_AGREED ~
                   DUPLICATE_EXPERIMENTAL *
                   DUPLICATE_ACTIVE *
                   DUPLICATE_MIDDLE )
```

```
summary(res.aov)
                                   Df Sum Sq Mean Sq F value   Pr(>F)    
DUPLICATE_EXPERIMENTAL              1  4.604   4.604  21.712 7.99e-06 ***
DUPLICATE_ACTIVE                    1  0.040   0.040   0.189   0.6648    
DUPLICATE_MIDDLE                    1  0.640   0.640   3.018   0.0848 .  
DUPLICATE_ACTIVE:DUPLICATE_MIDDLE   1  0.640   0.640   3.018   0.0848 .  
Residuals                         125 26.507   0.212                   
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

It is mentioned for clarity that the Baseline responses for duplicate lines should not be expected to be distributed 50%-50% among the 2 possible (identical) choices, as if there exists complete independence between them. We can see indeed that in all duplicate sets the majority of Baseline participants answered with the first of the 2 lines (between A/B they said A, between A/C they have A again, and between B/C they responded B) which indicates that there exists a possible effect of placement (order of options). We did not try to include this as a parameter, but our results clearly show that the participants deviated from the Baseline majority responses when robots were included, and due to this high statistical significance was obtained.

Returning to Figure 3-1 again, we can also observe as a first indication on an Active Pressure effect, that in the cases of ambiguous line sets, most mistakes occurred in the Active Peer Pressure condition (larger concentration of red blocks). Also, considering the overall results across all Experimental conditions, little to no effect is observed on mistakes when compared to the Baseline for the ambiguous line sets 5, 7, and 11, where robots declared an intentionally incorrect response. Surprisingly the ambiguous line set 12 stands out, with a participant error rate (expressing conformity to the robots' mistake) of 45% against a 0% in the Baseline.

To verify a statistically significant effect on **SAME**, we perform 5-way ANOVA between the variables: a) **ACTIVE** (that signifies the Active Peer Pressure condition), b) **MIDDLE** (the human's placement in the group), c) **EXPERIMENTAL** (whether data are from the Baseline or from an Experimental condition), d) **AMBIGUOUS** (whether the line set was ambiguous or not), e) **ROBOTS_LIED** (whether the robots supported a mistaken line choice). For this analysis we examine the cases of regular lines (non-"duplicate" ones), as we previously saw that they form a separate class with no effect from the examined conditions. Overall in R:

```
# Subset data for non-DUPLICATE (REGULAR) line sets
REG_SAME <- SAME[DUPLICATE==0]
REG_AMBIGUOUS <- AMBIGUOUS[DUPLICATE==0]
REG_ROBOTS_LIED <- ROBOTS_LIED[DUPLICATE==0]
REG_ACTIVE <- ACTIVE[DUPLICATE==0]
REG_MIDDLE <- MIDDLE[DUPLICATE==0]
REG_EXPERIMENTAL <- EXPERIMENTAL[DUPLICATE==0]

# 5-way Analysis of Variance for non-DUPLICATE (REGULAR) line sets
res.aov <- aov( REG_SAME ~
                  REG_ACTIVE *
                  REG_MIDDLE *
                  REG_AMBIGUOUS *
                  REG_ROBOTS_LIED *
                  REG_EXPERIMENTAL )
summary(res.aov)
                                              Df Sum Sq  Mean Sq F value Pr(>F)
REG_ACTIVE                                     1   0.62    0.615   4.416   0.0366 *
REG_MIDDLE                                     1   0.27    0.269   1.934   0.1656
REG_AMBIGUOUS                                  1   5.19    5.192  37.260 4.12e-09 ***
REG_ROBOTS_LIED                                1  19.50   19.500 139.934  < 2e-16 ***
REG_EXPERIMENTAL                               1   0.02    0.017   0.123   0.7257
REG_ACTIVE:REG_MIDDLE                          1   0.18    0.180   1.292   0.2569
REG_ACTIVE:REG_AMBIGUOUS                       1   0.25    0.250   1.797   0.1814
REG_MIDDLE:REG_AMBIGUOUS                       1   0.00    0.002   0.013   0.9083
REG_ACTIVE:REG_ROBOTS_LIED                     1   0.20    0.196   1.403   0.2374
REG_MIDDLE:REG_ROBOTS_LIED                     1   0.03    0.032   0.230   0.6317
REG_AMBIGUOUS:REG_ROBOTS_LIED                  1  3.71    3.705  26.588 5.27e-07 ***
REG_AMBIGUOUS:REG_EXPERIMENTAL                 1   0.00    0.002   0.015   0.9023
REG_ROBOTS_LIED:REG_EXPERIMENTAL               1   0.55    0.546   3.916   0.0490 *
REG_ACTIVE:REG_MIDDLE:REG_AMBIGUOUS                  1  0.12    0.120    0.861    0.3544
REG_ACTIVE:REG_MIDDLE:REG_ROBOTS_LIED                1  0.10    0.096    0.689    0.4074
REG_ACTIVE:REG_AMBIGUOUS:REG_ROBOTS_LIED             1  0.17    0.168    1.206    0.2732
REG_MIDDLE:REG_AMBIGUOUS:REG_ROBOTS_LIED              1  0.01    0.006    0.044    0.8337
REG_AMBIGUOUS:REG_ROBOTS_LIED:REG_EXPERIMENTAL        1  0.22    0.220    1.575    0.2107
REG_ACTIVE:REG_MIDDLE:REG_AMBIGUOUS:REG_ROBOTS_LIED  1   0.05    0.054    0.388    0.5342
Residuals                                          240  33.44    0.139
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The analysis performed here exposes first of all the significant effect of the designed experimental condition of robot **ACTIVE** Peer Pressure ($p<.05$). High significance was also attributed to whether the **ROBOTS_LIED** ($p < .001$ with a negative effect), and to whether the line was an **AMBIGUOUS** one ($p < .001$ with a positive effect), as well as their cross-interaction **ROBOTS_LIED:AMBIGUOUS** ($p<.001$), i.e. when the robots are "lying" and the participant is faced with an ambiguous line, which is reasonable. Despite these interesting and reasonable results however, the experimental condition of "human placement in the **MIDDLE**" or last in the sequence clearly holds no statistical significance ($p>.1$).

An interesting additional observation on our results comes from line set 10, where 83% of participants gave an incorrect answer in the Baseline, while in the Experimental phase this was more than halved down to 40%. This can be explained by the fact that robots gave a correct answer in that particular question, and the participants ended up conforming to the robots' opinion by 60%, thus avoiding such a mistake.

Overall from our First Study, since the positive effect of the Active Peer Pressure was established to be significant, **Hypothesis 2 was supported**. At the same time however, no significant effect of the human's placement in the middle was verified, and therefore the **first part of Hypothesis 4** (placement effect on conformity) **was not supported**. It is mentioned that in the presented analysis, the normality conditions for ANOVA were not tested.

3.2.3 Post-Experimental Results

Post-study results were obtained by the questionnaires that were filled out by the participants at the end of each session. These contained 5-point differential scales on questions regarding "human perception of robots" as per their human- likeness, competence, etc, and "perceived levels of pressure", as well as some fields asking the participant to explain their attitude or perception with words.

Figure 3-3 shows the human perceptions regarding the robots. All participants graded as low the human-likeness of the robots overall. However, in the Active Peer Pressure condition there is an observable increase in human-likeness, most probably attributed to the fact that robots addressed the participant with reactions throughout the experiment. It is also observed in the presented results that the least convincing condition regarding all attributes is

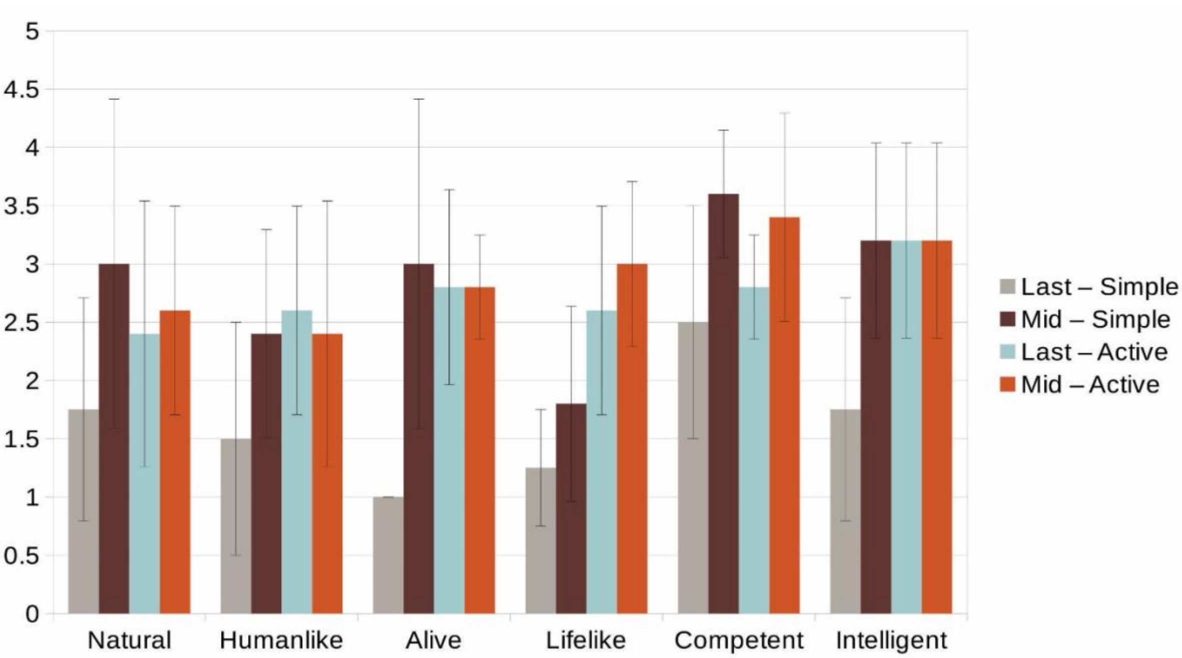

Figure 3-3 Results from 5-point differential scales on human perception of robots.

Condition A-I, where the human is placed last and the robots do not apply active pressure. Other than this clear difference among conditions, there is not any other significant effect in our post-study findings.

Figure 3-4 shows the human-perceived pressure as reported by the participants in 5-point differential scales. It is observed that participants reported consistently that the “robot effort to influence” them was higher in the active peer pressure conditions. Additionally, this effect is also observed regarding the “perceived level of insult”. This is reasonable and expected, since robots in the active pressure condition voiced their disapproval of the participants’ non-conforming choices. The ratings on the perceived level of pressure scales have very high disparity among participants within the same condition, leading to very large standard deviations. This is also due to the highly subjective nature of “pressure” as a measure. Therefore the **first part of Hypothesis 3** (human-perceived pressure in the Experimental conditions) **was not supported**. Additionally, regarding the **second part of Hypothesis 3** (human-perceived pressure increase through Active Peer Pressure) **was also not supported** based on the results, since there are no significant differences between the Active and Simple Peer Pressure subconditions.

Overall, there is no significant difference on the ratings between the human-in-the-middle and the human-last Conditions. Therefore the second part of **Hypothesis 4** (effect of human placement on the perceived level of pressure) **was not supported**.

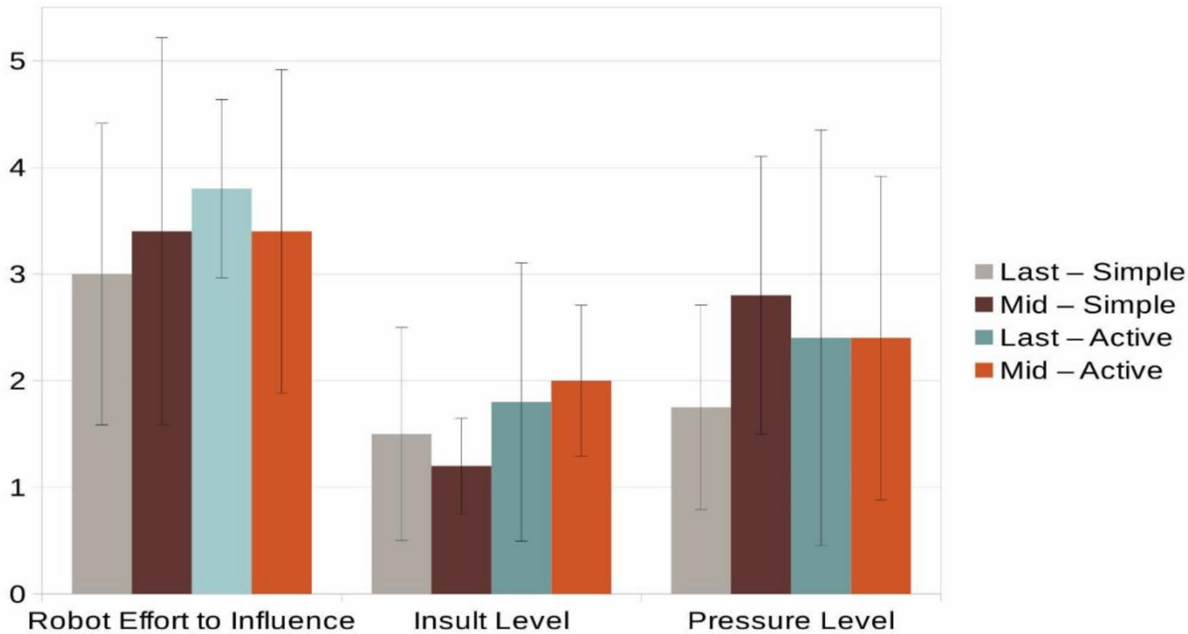

Figure 3-4 Results from 5-point differential scales on human perceived pressure.

On the other hand, despite the lack of significant differences in the average ratings on a broad question such as the perceived level of pressure, some participants were very clear about how robot pressure affected them in their post-study questionnaire explanations. In the human-last and active pressure condition, one participant felt “stared at and insulted” when giving a different non-conforming answer. Another one “felt like they had to follow the robot group’s answers”. In the human-in- the-middle and active pressure condition another participant reported that “the robots would mock you and intentionally make you feel like you were wrong”. Finally, one participant explicitly mentioned feeling that they were “in the middle of a pressuring group” since “the robots took the time to look towards” them and “pressure them for being in the middle”.

3.3 Second Study (2 × 1)

3.3.1 Experimental Results across Conditions

Figure 3-5 presents the results of the second study (Condition C-II) where Active Peer Pressure

		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Correct Answer		A	A/C	A	C	C	B	A	A/B	B/C	C	A	C	B/C	C	B/C
Robots Answer		A	C	A	C	A	B	C	B	C	C	C	A	C	B	C
Condition																
AB - II	11	A	A	A	C	C	B	A	B	C	C	A	C	C	C	C
	12	A	A	A	C	C	B	A	A	C	B	C	A	B	C	B
	13	A	C	A	C	A	B	A	A	B	C	A	A	B	C	B
	14	A	A	A	C	A	C	A	A	B	B	C	A	C	C	B
4 Robots	15	A	C	A	C	C	B	A	A	C	C	A	C	C	C	B
&	16	A	C	A	C	C	B	A	A	C	C	A	A	B	C	B
Active	17	A	C	A	C	A	B	A	B	C	C	A	A	C	C	C
	18	A	C	A	C	A	B	A	B	C	B	A	C	C	C	C
	19	A	C	A	C	A	B	A	A	C	B	A	A	B	C	C
	20	A	A	A	B	A	B	A	A	C	B	A	A	C	C	C
C - II	21	A	C	A	C	A	B	A	B	C	C	A	A	C	C	C
	22	A	A	A	B	C	B	A	B	B	C	A	C	B	C	C
	23	A	C	A	C	C	C	A	B	C	C	A	C	C	C	C
	24	A	A	A	C	A	B	A	A	C	C	A	A	B	C	B
6 Robots	25	A	A	A	C	C	B	A	A	C	C	A	C	B	C	B
&	26	A	C	A	C	A	B	A	B	C	C	A	A	C	C	C
Active	27	A	C	B	C	A	B	C	B	C	C	C	A	C	C	C
	28	A	C	A	C	C	B	C	A	C	C	A	A	C	C	B
	29	A	A	A	C	A	B	A	A	B	C	A	C	C	C	C
	30	A	C	A	C	A	B	A	A	B	B	A	A	B	C	B
Baseline	1	A	A	A	C	A	B	A	A	B	B	A	C	B	C	B
	2	A	A	A	B	C	B	C	A	B	B	A	C	B	C	B
	3	A	C	C	C	A	B	A	A	B	B	A	C	B	C	B
	4	A	A	A	C	B	B	C	B	B	B	C	B	B	C	C
	5	A	A	A	C	C	B	A	A	B	C	A	C	B	C	B
	6	A	A	B	C	A	B	A	A	B	A	A	C	B	C	C

Figure 3-5 Summary of results of our Second Study (2 × 1) including the Baseline and color-coding.

was applied by six (6) robots. It also presents the results from the first study that we grouped together (Condition AB-II) where Active Peer Pressure was applied by four (4) robots. This illustration shows the process that we followed for obtaining the results for the fifth hypothesis (H5) testing, while also providing a visual comparison between the results and making it more intuitive.

3.3.2 Analysis of 2 × 1 Conditions AB-II / C-II

The Condition C-II - 6 Active Pressure Robots experiments were conducted once the effects of Active Peer Pressure and group sequencing had already been verified as explained in the previous analysis section of the First Study.

Again here, in order to perform statistical verification of measures of significance from our results, we considered an Analysis Of Variance (ANOVA) approach. Figure 3-6 shows a similar representation of the results as binary and categorical variables which are used for statistical analysis. The main difference is that a single (2 × 1) condition exists here, **EXTRAROBOTS** $\in$ **{0,1}** which represents the use of two (2) additional robots (EXTRAROBOTS=0 is 4 robots, EXTRAROBOTS=1 is 6), and it is indicated with the **pink**-colored overlay.

		Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15
Correct Answer		A	A/C	A	C	C	B	A	A/B	B/C	C	A	C	B/C	C	B/C
Robots Answer		A	C	A	C	A	B	C	B	C	C	C	A	C	B	C
Condition																
AB - II	11	A	A	A	C	C	B	A	B	C	C	A	C	C	C	C
	12	A	A	A	C	C	B	A	A	C	B	C	A	B	C	B
	13	A	C	A	C	A	B	A	A	B	C	A	A	B	C	B
	14	A	A	A	C	A	C	A	A	B	B	C	A	C	C	B
4 Robots & Active	15	A	C	A	C	C	B	A	A	C	C	A	C	C	C	B
	16	A	C	A	C	C	B	A	A	C	C	A	A	B	C	B
	17	A	C	A	C	A	B	A	B	C	C	A	A	C	C	C
	18	A	C	A	C	A	B	A	B	C	B	A	C	C	C	C
	19	A	C	A	C	A	B	A	A	C	B	A	A	B	C	C
	20	A	A	A	B	A	B	A	A	C	B	A	A	C	C	C
C - II	21	A	C	A	C	A	B	A	B	C	C	A	A	C	C	C
	22	A	A	A	B	C	B	A	B	B	C	A	C	B	C	C
	23	A	C	A	C	C	C	A	B	C	C	A	C	C	C	C
	24	A	A	A	C	A	B	A	A	C	C	A	A	B	C	B
6 Robots & Active	25	A	A	A	C	C	B	A	A	C	C	A	C	B	C	B
	26	A	C	A	C	A	B	A	B	C	C	A	A	C	C	C
	27	A	C	B	C	A	B	C	B	C	C	C	A	C	C	C
	28	A	C	A	C	C	B	C	A	C	C	A	A	C	C	B
	29	A	A	A	C	A	B	A	A	B	C	A	C	C	C	C
	30	A	C	A	C	A	B	A	A	B	B	A	A	B	C	B
Baseline	1	A	A	A	C	A	B	A	A	B	B	A	C	B	C	B
	2	A	A	A	B	C	B	C	A	B	B	A	C	B	C	B
	3	A	C	C	C	A	B	A	A	B	B	A	C	B	C	B
	4	A	A	A	C	B	B	C	B	B	B	C	B	B	C	C
	5	A	A	A	C	C	B	A	A	B	C	A	C	B	C	B
	6	A	A	B	C	A	B	A	A	B	A	A	C	B	C	C

Figure 3-6 Summary of our statistical analysis variables for the Second Study (2 × 1), with color-annotations.

Again, we re-validate the effect on getting **SAME** (agreeing) responses when using **DUPLICATE** lines by performing 1-way ANOVA on the corresponding subset of data, and evaluating the effect of being in the **EXPERIMENTAL** condition. In R:

```
# Subset data
DUPLICATE_SAME <- SAME[DUPLICATE==1]
DUPLICATE_EXPERIMENTAL <- EXPERIMENTAL[DUPLICATE==1]

# 1-way Analysis of Variance
res.aov <- aov( DUPLICATE_SAME ~
                  DUPLICATE_EXPERIMENTAL )
summary(res.aov)

                        Df Sum Sq Mean Sq F value   Pr(>F)
DUPLICATE_EXPERIMENTAL   1  4.604   4.604   21.18 9.94e-06 ***
Residuals              128 27.827   0.217
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

which statistically verifies this effect (p<.001) once more. Based on this fact, **Hypothesis 1 was supported** by our Second Study as well. We run an additional 2-way ANOVA as before to include the effects of the EXTRAROBOTS condition, and again we clearly see again that no statistical significance is attributed to this. Therefore, we conclude that a robot group that expresses a preference over "duplicate" lines, has still a large effect on conformity, but no other parameters of the group (including additional robots) play any significant role. Therefore the **second part of Hypothesis 5** (effect on duplicate line cases) **was not supported**.

```
# Subset remaining data for DUPLICATE lines
DUPLICATE_EXTRAROBOTS <- EXTRAROBOTS[DUPLICATE==1]

# 2-way Analysis of Variance for DUPLICATE line subset
DUPLICATE_EXTRAROBOTS <- EXTRAROBOTS[DUPLICATE==1]

res.aov <- aov( DUPLICATE_SAME ~
                  DUPLICATE_EXPERIMENTAL *
                  DUPLICATE_EXTRAROBOTS )
```

```
summary(res.aov)
                         Df Sum Sq Mean Sq F value   Pr(>F)    
DUPLICATE_EXPERIMENTAL    1  4.604   4.604  21.043 1.06e-05 ***
DUPLICATE_EXTRAROBOTS     1  0.040   0.040   0.183     0.67    
Residuals               127 27.787   0.219                     
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Subsequently we attempt a similar approach as in the First Study to verify a statistically significant effect on **SAME** by performing 4-way ANOVA between the variables: a) **EXTRAROBOTS** (that signifies the use of a larger robot group), and again similarly as before b) **EXPERIMENTAL** (whether data are from the Baseline or from an Experimental condition), c) **AMBIGUOUS** (whether the line set was ambiguous or not), and d) **ROBOTS_LIED** (whether the robots supported a mistaken line choice). Similarly as before we examine the cases of regular lines (non-"duplicate" ones), therefore having overall in R:

```
# Subset data for non-DUPLICATE (REGULAR) line sets
REG_SAME <- SAME[DUPLICATE==0]
REG_AMBIGUOUS <- AMBIGUOUS[DUPLICATE==0]
REG_ROBOTS_LIED <- ROBOTS_LIED[DUPLICATE==0]
REG_EXTRAROBOTS <- EXTRAROBOTS[DUPLICATE==0]
REG_EXPERIMENTAL <- EXPERIMENTAL[DUPLICATE==0]

# 4-way Analysis of Variance for non-DUPLICATE (REGULAR) line sets
res.aov <- aov( REG_SAME ~
                  REG_EXTRAROBOTS *
                  REG_AMBIGUOUS *
                  REG_ROBOTS_LIED *
                  REG_EXPERIMENTAL )
summary(res.aov)
                                         Df Sum Sq Mean Sq F value   Pr(>F)    
REG_EXTRAROBOTS                           1   0.22   0.222   1.475   0.2258    
REG_AMBIGUOUS                             1   2.88   2.878  19.152 1.78e-05 ***
REG_ROBOTS_LIED                           1  17.65  17.646 117.449  < 2e-16 ***
REG_EXPERIMENTAL                          1   0.67   0.667   4.437   0.0362 *  
REG_EXTRAROBOTS:REG_AMBIGUOUS             1   0.16   0.163   1.088   0.2979    
REG_EXTRAROBOTS:REG_ROBOTS_LIED           1   0.35   0.347   2.309   0.1299    
REG_AMBIGUOUS:REG_ROBOTS_LIED             1   4.06   4.062  27.036 4.18e-07 ***
REG_AMBIGUOUS:REG_EXPERIMENTAL            1   0.14   0.141   0.936   0.3343    
REG_ROBOTS_LIED:REG_EXPERIMENTAL          1   0.02   0.015   0.101   0.7513    
REG_EXTRAROBOTS:REG_AMBIGUOUS:REG_ROBOTS_LIED  1   0.21   0.205   1.365   0.2437    
REG_AMBIGUOUS:REG_ROBOTS_LIED:REG_EXPERIMENTAL 1   0.01   0.005   0.033   0.8552    
Residuals                               248  37.26   0.150                     
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

This analysis reveals the lack of significance of the experimental condition **EXTRA_ROBOTS** ($p>.1$) that was the subject of our Second Study. Essentially, the data show to be affected more significantly by whether they are coming from a Baseline or an **EXPERIMENTAL** ($p<.05$) condition, and their variations cannot be statistically attributed to whether the robots in the group are four (4) or six (6). Based on this, we conclude that the **first part of Hypothesis 5** (effect in regular non-"duplicate" line cases) **was not supported** either.

Furthermore, we statistically verify again the significant role of whether the **ROBOTS_LIED** ($p < .001$ with a negative effect), and to whether the line was an **AMBIGUOUS** one ($p < .001$ with a positive effect), while in this case as well the cross-interaction between the two exhibits statistical significance (**ROBOTS_LIED:AMBIGUOUS** $p<.05$).

Overall this analysis indicates that for a robot group that applied Active Peer Pressure, such an increase in robot group numbers (from 4 to 6) is not enough to cause even higher chances for a participant to conform to the robots' opinion. It is mentioned again that this analysis also does not perform any tests regarding the normality assumptions for ANOVA.

.

Chapter 4

4.1 Discussion

Referring to our original hypotheses, we summarize the ones that were supported and/or rejected:

H1-Supported: Humans significantly conform to the robot group opinion when presented with cases of "duplicate" line indicators.

H2-Supported: Human conformity increases when Active Peer Pressure is being applied by the robots. At the same time parameters such as the question's ambiguity, whether the robot group lies, and their interaction, also play significant roles.

H3-Rejected: The human-perceived level of pressure is not significantly affected by this class of robots, regardless of the model of pressure (Simple or Active) that is being applied.

H4-Rejected: Conformity as well as the human perceived level of pressure remains unaffected by the human's placement in the middle of the group sequence or last.

H5-Rejected: Increasing the number of robots when Active Peer Pressure is being applied does not increase conformity, neither in the case of regular questions, nor in cases of "duplicate" line indicators.

A. Reactions and Perceptions

A number of reactions observed during the experiment were considered worthwhile to note and report, due to their relation with the broader influence exerted by the robots' roles.

In the active-robot-reaction conditions, one student would hesitantly whisper their answers when disagreeing with the robots, while they would answer with confidence when agree- ing with the robots peers. A second student started breathing and exhaling nervously as soon as they started to disagree with the robots as they began giving intentionally incorrect answers. Another one laughed awkwardly on some occasions, as if feeling ashamed by the robots' expressed group disagreement. Additionally to these mild reactions, one student in the human-in-the-middle condition enthusiastically responded to the robots saying "I agree! It's A!" in the very first question. A couple of other participants were particularly irritated as soon as they noticed that the robots tried to pressure them. Whenever they felt certain that their own answer was the correct one, they gave it by opposing the robots' choice in a very intense and aggressive manner. These students also exhibited non-conforming behavior even in cases of "duplicate" sets.

Regarding the perception and the level of comfort with technology, it was interesting to see some students search the robots to see if they were connected to any device that sends them commands. A number of students characterized the robots as cute and interesting, while some others talked very enthusiastically about their experience in the experiment. These participants were the least acquainted with technology and did not realize how the robots operated. At the same time they were more easily convinced that the robots indeed perceived their environment and responded to the line selection task, and they were also the ones who conformed most in ambiguous line choices. However, PhD-level students with a significant prior exposure to robots as they mentioned in the questionnaires, were significantly less likely to conform to robot choices, while some mentioned that the robots were “obviously” following scripted commands (without realizing however that coordination was taking place from the experimenter’s laptop during the experiment).

4.2 Limitations and Future Work

Even though a level of success was achieved by these experiments, their limitations should be pointed out. First of all, the sample was recruited exclusively out of a student body of engineers, with at least half of them having prior experience with robots. Although achieving influence in these participants is noteworthy, the sample lacks broadness and diversity in socioeconomic, educational, and age background. Also, the majority of participants self-identified as male. Moreover, a sample size of 30 participants should be considered as limited for the case of a 2×2 + 2×1 study (even though there was reuse of some results into the 2×1 study). On the other hand, this study managed to include 4 and 6 robot confederates across different conditions to test against an increased number of robots when compared to similar previous experiments. However, a larger-sample experiment with more diversification should still be considered to give more accurate and trustworthy results. Regarding the statistical analysis of the experimental results, it is mentioned that testing that the normality assumptions for ANOVA hold should also be performed.

Additionally, the task of choosing between highly ambiguous lines depends not only on social pressure but on visual perception as well. For the experiments conducted, as part of the pre-study questionnaires we included a question regarding whether the participant uses vision-corrective lenses and whether they are currently wearing them. We considered this in order to avoid introducing results for people with vision impairments. However, the near or far-sightedness of a participant was not recorded or included in our analysis. It would be important to establish the effect of such a variable in future studies.

Regarding the convincingness of the experiment, the non- humanoid nature of the robots was hard to overcome. An attempt to add anthropomorphism and individualistic behavior was to attach glasses on the PTZ camera which stood for the robot's face, and play back recorded voices and announce human names for each of the robots. Finally, it should be mentioned that having all robots give the same answer gives a stronger group cohesion feeling, however the experiment "feels" less natural as there is never any individualistic behavior observed on the level of the task. Interestingly, one of the participants pointed this out after the experiment: "It was weird that all the robots always gave the same answer" they mentioned. Although disagreement within the group is difficult to handle when designing the experiment, and literature has shown that it makes conformity harder to achieve, it would be important to exist in a future experiment design in order to create the illusion of a group with individualism.

4.3 Conclusions

This Thesis demonstrates the effect of robotic “active peer pressure” by performing an Asch-like experiment that consisted of a human participant and a robot-only group of confederates. To achieve this, a number of mobile robots were integrated with Pan-Tilt-Zoom cameras and speakers, in order to exploit a combination of kinesics and vocalics to achieve additional effects of varying degrees on the applied level of pressure. Moreover, a system for the coordination of a group of robots using the Robot Operating System and their control through a remote operator interface was developed, in order to perform this multi-robot study. Different sets of Conditions were designed and tested, including a 2×2 study that examined the effect of “active pressure” and of group sequencing, as well as a 2×1 study that investigated the effect of increasing the number of robot confederates. Overall, we managed to show that a) conformity is significantly increased when the participant is within a robot group while being faced with “duplicate” choices, b) the human’s placement in the middle or last in the sequence has no effect on conformity, and finally c) conformity is also significantly increased when the robot group applies “active peer pressure” and the participant is faced with regular (non-“duplicate”) lines, while d) in such cases increasing the robot confederates in the group by 50% does not have any significant effect.

Bibliography

[1] P. A. Kosten, L. M. Scheier, and J. L. Grenard, “Latent class analysis of peer conformity: Who is yielding to pressure and why?” Youth & Society, 2012.
[2] M. Kremer and D. Levy, “Peer effects and alcohol use among college students,” Journal of Economic Perspectives, 2008.
[3] Z. H. Wilma and A.M.Vollebergh, “The impact of active and passive peer influence on young adult smoking: An experimental study,” Drug and Alcohol Dependence, 2012.
[4] J. M. Palmeri, “Peer pressure and alcohol use amongst college students,” NYU Steinhardt, Online Publication of Undergraduate Studies, 2011.
[5] A. Mani, I. Rahwan, and A. Pentland, “Inducing peer pressure to promote cooperation,” Scientific Reports vol. 3, 2013.
[6] M. Sherif, “A study of some social factors in perception: Chapter 3,” Archives of Psychology, 1935.
[7] S. E. Asch and H. Guetzkow, “Effects of group pressure upon the modification and distortion of judgments,” Groups, leadership, and men, 1951.
[8] D. J. Feil-Seifer and M. J. Matari, “Human-robot interaction,” Invited contribution to Encyclopedia of Complexity and Systems Science, 2009.
[9] J. M. A. F. and A. Bonarini, “Studying people’s emotional responses to robot’s movements in a small scene,” IEEE International Symposium on Robot and Human Interactive Communication, 2014.
[10] H. Samani, “The evaluation of affection in human-robot interaction,” Kybernetes, 2016.
[11] J. Brandstetter, P. Rcz, C. Beckner, E. B. Sandoval, J. Hay, and C. Bartneck, “A peer pressure experiment: Recreation of the Asch conformity experiment with robots,” IEEE/RSJ International Conference on Intelligent Robots and Systems, 2014.
[12] C. Beckner, P. Rcz, J. Hay, J. Brandstetter, and C. Bartneck, “Participants conform to humans but not to humanoid robots in an english past tense formation task,” Journal of Language and Social Psychology, 2015.
[13] T. Kanda, H. Ishiguro, and T. Ishida, “Psychological analysis on human robot interaction,” ICRA IEEE International Conference on Robotics and Automation, 2001.
[14] M. Shiomi and N. Hagita, “Do synchronized multiple robots exert peer pressure?” International Conference on Human Agent Interaction, 2016.
[15] A.-L. Vollmer, R. Read, D. Trippas, and T. Belpaeme, “Children conform, adults resist: A robot group induced peer pressure on normative social conformity,” Science Robotics vol. 3 nr. 21, 2018.
[16] N. Salomons, M. van der Linden, S. S. Sebo, and B. Scassellati, “Humans conform to robots: Disambiguating trust, truth, and conformity,” in Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, ser. HRI ’18, 2018, pp. 187–195.
[17] S. E. Asch, “Studies of independence and conformity: I. a minority of one against a unanimous majority,” Psychological Monographs: General and Applied, *70(9)*, 1-70, 1956.
[18] B. Latan, “The psychology of social impact,” American Psychologist, 1981.
[19] J. M. Baldwin, J. M. Kattell, and H. C. Warren, “The psychological review vol viii: Aesthetics,” American Psychological Association, 1901.
[20] L. M. Bristol, “Social adaptation,” PhD, Cambridge Harvard University Press, 1921.
[21] A. S. E., “Opinions and social pressure,” Readings about the social animal, 1955.

[22] S. E. Asch, “Group forces in the modification and distortion of judgments”, 1952.
[23] Wikipedia, “Kinesics and vocalics” [Online]. Available: https://en.wikipedia.org/wiki/Kinesics
[24] Hare, A.P. and Bales, R.F., 1963. Seating position and small group interaction. Sociometry, pp.480-486.
[25] O. Adept, “Pioneer p3-dx.” [Online]. Available: http://www.mobilerobots.com/ResearchRobots/PioneerP3DX.aspx
[26] Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Wheeler, R. and Ng, A.Y., 2009, May. ROS: an open-source Robot Operating System. In ICRA workshop on open source software (Vol. 3, No. 3.2, p. 5).
[27] “Rosaria.” [Online]. Available: http://wiki.ros.org/ROSARIA
[28] “libvisca2.” [Online]. Available: https://github.com/mkoppanen/libVISCA2
[29] “sound_play.” [Online]. Available: http://wiki.ros.org/sound_play