Convergence in morphological behaviour and decision tasks with (human and) non-human peers

[Extended Abstract] *

Jürgen Brandstetter, Peter Racz, Jakub Zlotowski, Eduardo Sandoval, Jennifer Hay, Christoph Bartneck
Human Interface Technology Lab
P.O. Box 4800
Christchurch, 8140, New Zealand
juergen.brandstetter@pg.canterbury.ac.nz

ABSTRACT

This paper provides a sample of a LaTeX document which conforms to the formatting guidelines for ACM SIG Proceedings. It complements the document *Author's Guide to Preparing ACM SIG Proceedings Using LaTeX2$_\epsilon$ and BibTeX*. This source file has been written with the intention of being compiled under LaTeX2$_\epsilon$ and BibTeX.

The developers have tried to include every imaginable sort of "bells and whistles", such as a subtitle, footnotes on title, subtitle and authors, as well as in the text, and every optional component (e.g. Acknowledgments, Additional Authors, Appendices), not to mention examples of equations, theorems, tables and figures.

To make best use of this sample document, run it through LaTeX and BibTeX, and compare this source code with the printed output produced by the dvi file.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

General Terms

Theory

Keywords

ACM proceedings, LaTeX, text tagging

1. INTRODUCTION

Looking at this development, one can predict a future where robots help in the office, teach kids in the classroom, becoming companions, work in advertising, or help in our households. We can already see a rapid development in the shipping of industrial robots. In the year 1994 only 50.000 units where ships, worldwide. 2011 165.000 units, and there is a estimation of 207.500 units for 2015, which corresponds to a growth rate of 415% [12]. In all of this close futures, humans will interact physically and verbally with those robots. Especially the speech communication will become one of the most import ones because it is the post natural way to communicate. We can already see a movement where we use very primitive robots aka smartphone to communicate per speech with them. Modern smartphone operating system like Apples iOS or Googles Android provide system (Siri, Google Now) which analyse all our data and behaviour and use natural speech for communication. When we look at high developed countries a cellphone rate of 1.6 is not special anymore[9]. Samsung for example, implemented speech communication in all of their new smart TVs [10]. That means in theory there will be soon more virtual communication partners than human once.

Interestingly this scenario can create a new level of human language development. Till now the development of human language and spreading of new words and word forms was basically reduced to face to face communication and mass media. With robots connected to the internet and in theory connected to each other, the evolution of speech can move way faster than till now. We can built a scenario where one robot learns a new word for an object and all other robots synchronize immediately. Robots could also use a "wrong" from of a word or grammar, but because all robots use this form, humans tend to change their way of using this word or grammar. One possible negative scenario is that only a view providers create dictionaries for the robots. If one of those providers introduces a major change in their robot dictionary, all robots could change the way how they speak in a very short time period. Providers can now use this power to influence what words we use for some topics and can directly change the outcome of a political debate or other important topics. One example could be, that all robots start to change talking about climate change and switch to global warming. Frank Luntz suggested to the republic party in USA in 2001 after a focus group study that

*A full version of this paper is available as *Author's Guide to Preparing ACM SIG Proceedings Using LaTeX2$_\epsilon$ and BibTeX* at `www.acm.org/eaddress.htm`

George W. Bush should replace the word global warming with climate change, because it is less frightening for the citicens[4].

When robots get more similar to humans, we except them easier as friends, companions and living creatures. This social bond makes it easy to influence/persuade a human. It is the same effect as a friend tells us about a great product. Because it is a friend, we will trust him/her more than an advertisement. But robots, which have their dictionary from one source, can massively be used as highly influencing opinion/advertisement makers. Companies like Apple already use "happier" words instead less positive words in their retail stores to be more persuasive[8].

We believe it is important to understand what effect a wast among of robots can have on the human behavior, especially because robots can easily outnumber humans in the near future. To understand the power of robots, we developed three research questions based on two famous psychology experiments.

The first experiment was conducted by Muzafer Sherif who studied the conformity/peer pressure effect in a ambiguous situation. The experiment builds on the autokinetic effect which describes a situation in which people perceive a sudden movement of light point when [illegible]. Sherif create this effect by sitting participants into a dark room with one small light point in the room. The participants had to look at this point and afterwards they had to tell Sherif how much the light point moved. Actually the light never moved but the autokinetic effect created a sudden movement. This movement is different between most of the people. After he did the experiment with one person, he set two or three people in the same room and asked them to say out loud how much the point moved. Astonishingly after three rounds all the participants said the same number, even though everyone perceived a different movement. This effect is called informational conformity or social proof and describes the effect in what people in uncertain situation look at their neighbours to see what it true and confirm with them. [5] [11][7].

The second experiment is knows as the Asch experiment. The original title is "Effects of group pressure upon the modification and distortion of judgements" by Solomon Asch. Asch built up on the knowledge created by Sherif and wanted to find out if humans confirm also in non-ambiguous situations. The experiment simulates a simple visual line test. The participant saw three lines with different heights, labelled 1, 2, 3 on the right side of a board, and one reference line on the left side labelled with ? (see Figure 1). The task was to say what line matches the reference line. When the participants were alone in a room to perform the test they where almost every time right. In a second round Asch placed the participants in a room with other participants which where all actors and they all said the wrong answer. Even though, the real participant knew exactly the correct answer, in 32% (group size bigger 4) of all tasks the participants went along with the group. In this case, Asch had the evidence that peer pressure and conformity does not only work in ambiguous situation but also in non-ambiguous situation [6][2][1][3].

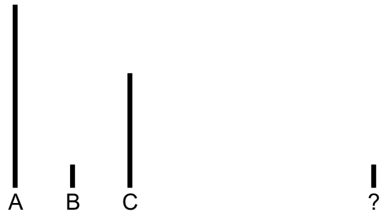

Figure 1: An example of an visual task. Based on Solomon Asch line experiment.[1]

Add somewhere her a definition of morphological productivity

Building on the above experiments and our interest in Human Robot Interaction and verbal communication we developed an experiment with includes all factors. The experiment is an extension of the original Asch experiment. It is dived into four part. One ambiguous and one non-ambiguous line experiment and one ambiguous and one non-ambiguous verbal task experiment. In the first question we want to see if humans confirm more to a robot group than to a human group. In this study we will only compare the data with the original Asch data. That means we look only at results given by the lines. In a later study we will run the same experiments with humans actors to compare all our conditions. In the second analyse we want to see if participants conform more to the robot group in a verbal task than in a visual task. And last question we want to see if participants conform more to the robot group in an ambiguous task than in a clear task?

2. METHODS

Rewrite first paragraph. Maybe nothing at all as intro is also fine. who knows

The setting consists of three parts: A physical setting, a virtual setting and a pilot study/test. To recreate the experiment it is important that the participants use the same physical setting. This is necessary because of visual perception, which can vary easily by sitting on a different position. The virtual setting (data + program) can be found on a public repository [github link]., In this case we try to give you as much reusable data as possible. And finally, the pilot study to create ground truth (reference) data.

2.1 Participants

For this experiments, a total of 23 (N=23) people participated. All of them were university related, which means they where students, had a university degree or were staff members. Half of the participants had English as their mother tongue. The other half had a different mother tongue and came from a non English background. Their age range is between 20 and 62 years. We had a high diversity rate in the participants scientific background. Participants came from an engineering, artistic or medical background.

2.1.1 Pilote

To compare the lines condition with the word condition, and to compare Sherifs studies with ours, we needed a baseline/pilot experiment. We created 107 line settings and asked twenty (20) people to participate in the line experiment. All settings were randomized for each participant. The apparatus was exactly the same as in the main experiment, the only difference was, that the participants did not say what line matches, they used the 1,2,3 button on the keyboard. We assume there is no difference between saying the answer out loud or pressing a button when no one is in the room except the participant self. We picked out of this 107 settings, exact 30 for our main experiment. The corpus of 30 lines is divided into two part. The first part (15 settings) contains all the settings where people made most of the mistakes. This lines are ambiguous and not easy to distinguish. The second part of the lines contains a set of lines where the two closest lines had no difference bigger than 30% and people made only one or zero mistakes. We choose this setting because Asch used a similar setting to prove his hypothesis[].

Add average impact when people where alone

l is to create a situation in which one partic- peer pressure by a group of robots. The in- s design goal came from the Asch and Sherif hich are introduced in the introduction sec- e experiment on a new level, we introduced task, a verbal task. This part of the experi- he verbal communication. This is especially use robots will work mainly in tasks where ate and work with humans, but not compete

sists of ten parts but only four parts are im- ntification. First the introduction for the first n (lines or words). The introduction is pre- ritten on the projection. We wanted to make participant gets exactly the same introduc- the noise in our data. Second, an example k (lines or words). Third, three questions s will say the correct form. Fourth, fifteen stions where all robots will say the wrong fifteen non ambiguous questions where all the wrong answer. Sixth, an introduction rt (lines or words). Seventh, an example of k (lines or words). Eighth, three questions s will say the correct form. Ninth, fifteen questions where all robots will say the wrong , fifteen non ambiguous questions where all the wrong answer.

tion consists of three horizontal parallel lines C which are on the left side of the projection belled with ? which is on the left side. An ex- een in figure []. The line on the left, marked hes with exact one line on the right. The rticipant is it to say which line, A, B, or C, he line labelled ?. This setting uses the Asch ing as a model. All in all, thirtythree differ- s are used. This setting can now be divided s. The first part consists of three settings cial focus. Those three settings are only used

at the beginning of the experiment. The robots will in all three trials say the correct answer. The second part consists of fifteen ambiguous settings. Ambiguous in this case means, that the participants in the pilot study could not easily say the correct matching line. The third part consists of fifteen non ambiguous settings. Non ambiguous in this case means, that the participants in the pilot study could easily answer the correct matching line, plus the difference between the two closest lines is not more than 30 percent. The goal of the robot is it to convince the participant that an other number, which is obviously wrong, is the right one. This gives us the chance to measure peer pressure.

The word projection consists of five vertical words. The task is it to say the word on the projection and its past tens form. For example Like - Liked. Each participant says the word corresponding to its position. That means participant one says the first word and its past tense. Participant two says the second word and its past tense. And so on. Similar to the line setting, this setting consists of thirtythree trials. In the first three trials all robots will form the correct past tense form. The next fifteen trials consists of ambiguous words. That means words which can have a regular or irregular past tense form. For example, drive - drove/drived. The robots will all form the regular past tense. The last fifteen trials consist of non ambiguous words. That means words which have only a irregular past tense form. For example, run - ran. Nevertheless, the robots will always form the regular past tense. In our example, run - runed. The task is it again, that the robots convince the person to use their form instead of the "correct" form.

2.3 Apparatus

The setting contains of a projector, a high quality wireless microphone, a table with five chairs, four robots and laptop to control the recordings and experiment. The projecting area has a dimension of 243x177cm. Nevertheless, the lines will not fill the whole screen, therefore, the maximum line length is

add dimention

X cm. The table is exact parallel to the projecting area and is on a distance of 2m. The dimensions of the table are 80x250cm.

To remotely control the robots and the presentation at the same time, a program/script is used which runs step by step through the experiment. All steps except when the participant speaks or the description runs is automated. To make the exact setting reusable, a quasi random generator script generates setting files for every participant. Those setting files are used by the controller script. All setting files are available in our online repository (see Appendix).

The room/lab of the experiment contains only black/dark surfaces except the presentation surface, which should reduce distractions. No real daylight was used, only standard sealing office lights.

Missing figure Add a picture of the setting

2.4 Procedure

Before the experiment started we set all four robots on chairs on their final position. The participants could also see a welcome page on the project.

The first part the participants had to do was to sign the consent form and to answer some questionnaires. The next step was to sit next to the robots and put on their microphone. Some participants asked what are the robots for. Normally we referred to the presentation and said the experimenter is not the researcher because she (we said the researcher is female) is at a conference, but she recorded the description for the experiment and the participant will hear her talking about why the robots are in the room. Some other times, we started to chat a bit with them and told them the robots have a artificial intelligence and lived the last 12 month with a New Zealand family to learn how to see, read, talk, ... Similar to a human baby. The robots are doing exactly the same experiment as the humans but for scalability reasons we put all robots and people on one room. In this case we can test five participants at the same time. The same explanation was used during the researchers description of the experiment at the beginning of the presentation. The experiment was over after about twenty minutes and we debriefed them and gave them a 10$ (NZD) voucher.

3. RESULTS

Describe basic strategies used to analyse. Write how Asch did his experiment with an example

3.1 Calculation with original ash data as example

Since it was hard to understand how the original calculation by Asch worked we use here the original Asch data[link to table] to show you how the data is calculated. We believe it is always easier to understand a subject if we can put it into a context [Effects of Group Pressure upon the Modification and Distrotion of Judgments].

N... Number of Participants
n... number of Rounds per participant
k... total number of critical questions for N participants
m... total number of errors for N participants
x... conformity in percent

$$k = n * N \tag{1}$$

Table 1: Original Asch data

Number of Critical Erros	Mistakes per Critical Group (N)
0	13
1	4
2	5
3	6
4	3
5	4
6	1
7	2
8	5
9	3
10	3
11	1
12	0

$$m = \sum_{i=0}^{n} i * "MistakesperCriticalGroup"[i] \tag{2}$$

$$x = m * 100/k \tag{3}$$

When we use now the data from tabel [x] (N=50, n=12) we get k=600, m=192, x=32%. To review if outcome x=32% is correct, we can take a look at the "Size of Majority" table in the paper "Opinions and Social Pressure" and we will see that the conformity lies between 30% and 35%.

3.2 End Results

At this point we describe our data in general and rewrite the three hypothesises and compare them with our data.

3.2.1 Do participants conform more to the robot group than the human group?

The first research question was: "Do participants conform more to the robot group than the human group?". Since we did not run a human condition till this point, we can only compare our data with original Asch data. First of all we can compare the total impact between Asch data and our data. Second, we want to see if we had a significant difference between the robots and Asch's human condition. Third, we can give a qualitative answer between Sherif and our data.

We performed a one-sample t-test to compare our robot/non-ambiguous condition (four robots) to the human condition described by Asch. Across all participants and trials, 32% of errors occurred. In our experiment we only observed 14% of errors (see Figure **??**). The test shows an impact given by the robots but not as high as the one given by humans. The t-test revealed that this difference is significant (t(68)=2.238, p=0.028)).

At this point, we must point out that the average impact is not a very stable calculation. One participant who is in all fifteen trials consistent with the rest of the group can have a major impact on the total impact level. Similar to Asch, only about one fourth of the participants went along with the group. Most of the participants where "stock" and could not free themselves of changing their mind back to a

Table 2: no signficant differenct (f(1,17)=0.984, p=0.335)

Type	Diffe	noIdea	noIdea
La	false	7.38	4.926
	true	8.83	4.448
	Total	8.25	4.575
Lc	false	3.13	5.939
	true	1.42	4.316
	Total	2.10	4.951
Wa	false	10.25	3.694
	true	6.58	3.147
	Total	8.05	3.762
Wc	false	3.38	5.290
	true	.83	.937
	Total	1.85	3.528

non-pressured answer. Out of 23 participant x went along with the group at least one time.

When we look at the impact level of the ambiguous condition, which is a comparison between Sherifs experiment, we can see an impact of 53% vs [mistaces when alone]%.

Sherif stuff here

rticipants conform more to the robot group erbal task than in a visual task?

search question was: "Do participants con-
he robot group in a verbal task than in a
To understand more about this question we
epeated measure ANOVA in which the task
he modality (lines vs. words) were the in-
ables and the number of errors made by the
s the dependent variable.

the ambiguous lines versus the ambiguous
hat we had a confirmation rate of 52% for
1% for the words. Furthermore all partici-
d at least once with the robots. On the other
nly a conformation rate of 12.7% for the lines
the words in the non-ambiguous condition.
easure ANOVA unveils that conformity level
s is not significant different (f(1,17)=0.984,
en lines/words and ambiguous/non-ambiguous
.

rticipants conform more to the robot group ambiguous task than in a clear task?

arch question was: "Do participants conform
ot group in an ambiguous task than in a clear
on two unveiled there is no significant effect
1,17)=0.984, p=0.335), but there was a sig-
or task difficulty (f(1,17)=20.857, p<0.001).
the ambiguous condition made an average of
rors while participants in the non-ambiguous
e on average only 1.975 (7.8%) errors. On
, there was no significant interaction effect
fficulty and modality.

4. CONCLUSIONS/DISCUSSION

Our main goal was it to find out if robots can create peer pressure and if so, how much difference is between humans peers and robot peers. The study uses the idea from Sherif, who studied conformity/peer pressure in ambiguous situations and Asch who studied conformity/peer pressure in non-ambiguous situations. To see if there is a difference between a less social task and a highly social task, we created a line conformity situation and a word conformity situation.

In our first question (hypotheses) we wanted to know if robots create the same peers pressure as humans. Our experiment and the analyses showed that robots are not able to create the same amount of peer pressure than humans. We only had a impact level of 14% for the robots, versus. 32% for humans. We understand there could be many different reasons for this outcome, but we belief the main reason is that humans do not see themselves as part of a group. They do not consider robots as their peers. Therefore they do not care if they go with they robots or not. We know from other studies about conformity [ref] that the similarity effect[refence] is very important. For example, is the conformity level higher when the participant beliefs that all group members study the same.

Go to social psychology video and find out how the effect is called when only one person is different in the group. Does this effect work also in the other way around. that means does one person feel unconfortable?

In our second question (hypothesis) we wanted to know if robots have a different impact in a visual task versus verbal task. Since unanimity and consistency is crucial for human communication our hypothesis was that the impact in the verbal test (word test) is higher than in the visual test (line test). We considered the verbal task as more social one. Especially by looking at the average impact of just 20% for non-ambiguous lines but 55% impact for non-ambiguous words. Nevertheless, the test results proofed us wrong. We could not find a significant difference between the impact in the verbal versus visual experiment.

Our explanation for this result is similar to the one in question one. We believe the social power of robots is still far less than human power. The reason for the difference between non-ambiguous lines vs. words is, that the priming effect[] took place. That means, when humans hear a word, does not matter if it is "wrong" or right, and they have to reaped it, there is a high probability that the humans will repeat the word how it was given, but not what was asked. As an example. The robots said at least two times drive-drived, before the human had to form the past tense. In this case, the human brain what is "primed" with the word 'drived' what made them use this form instead of the correct 'drove' form.

In our third question (hypothesis) we wanted to know if there is a high conformity in a non-ambiguous situation versus an ambiguous situation. Looking at the literature, we see that a human group creates a higher impact in a ambiguous situation than in a non-ambiguous situation. Our robot conditions showed the same outcome, with significant different result between those two test. At the end 100%;95% (Lines;Words) of all participant confirmed at least once with the robot group in a ambiguous situation and only 20%;55% confirmed in the non-ambiguous situation. To add a second level of security in our data, we conducted a pretest for the line condition where we tested participants being alone in a room to answer the line conditions. In the ambiguous line situation the average mistake level was 29% and 2% for the non-ambiguous. Compared with the group situation 53% made a mistake in the ambiguous situation, respectively 13% in the non-ambiguous situation. This is in direct comparison 29% vs. 53% (ambiguous) and 2% vs. 13% (non-ambiguous). Our conclusion for this effect is that robots have a certain power, which is enough to cause more humans committing errors in the ambiguous condition.

Tweaking the experiment. To find out if the group behaviour or the social behaviour of the robots is the reason for a lower conformity level we can change two values in a future experiment. First of all, we can simulate a group situation. For example, we give all participants and robots a coloured t-shirt and tell them that there is an other group which what we compare them. This should create a group situation. The second tweak could be a more social introduction. We could let the participants perform some task with the robot group. For example build together a Lego house, while doing some chatting. This situation could create a more trustful situation between humans and robots.

Future Work. In a future experiment we want to change several parts. First, we want to create a full human condition. That means a recreate of the Asch experiment for lines and words. Second, we want to use a more homogeneous group. For example, only female children which are as similar as possible. This could reduce our noise ratio, but limit our conclusion. A third tweak would be a mixed human and robot situation. Bla at al. [beeing alone in group] shows that being an outside/insider has a big influence on the individual behaviour. Finally, we want to a new test group. For example the participants have to form a sentence together or give a short speech. This could have the benefits that the words respectively lines used in the experiment are set into a context which create a more natural situation.

Create table with all important number.

effect

uman condition - simulate a peer situation.
ne color and tell them they are a group now
sozial introduction. Make the robots more

ntroduction in starting with a summary of
s. - Start the Discussion by restating the
eriment and the hypotheses that wre under
Summerize the research findings - consider
between the findings and the hypothesis.
ould also be discussed, plus methodological
ibe future research.

checklist 1) Briefly restate the studys aim and summarize the results in relation to hypotheses 2) Remark on the consistency of the findings with previous literature 3) Clarify theoretical implications of the findings 4) Note methodological or applied implications, including limitations of the study 5) Make suggestions for future research 6) Draw conclusions related to the general issue/problem identified at the outset.

No difference word vs. line. Robots are not our peers, therefore no need to agree on vocabulary.

The robots do have a certain power, which is enough to cause more humans committing errors in the ambigous condition.

This power is probablt still far less than human power.

The robots are not able to apply peer pressure because humans do not consider robots to be their peers.

5. ACKNOWLEDGMENTS

Tempeltan Foundation .... This section is optional; it is a location for you to acknowledge grants, funding, editing assistance and what have you. In the present case, for example, the authors would like to thank Gerald Murray of ACM for his help in codifying this *Author's Guide* and the **.bls** and **.tex** files that it describes.

6. ADDITIONAL AUTHORS

Additional authors: John Smith (The Thørväld Group, email: `jsmith@affiliation.org`) and Julius P. Kumquat (The Kumquat Consortium, email: `jpkumquat@consortium.net`).

7. REFERENCES

[1] S. E. Asch. Effects of group pressure upon the modification and distortion of judgments. *Groups, Leadership, and Men. S*, pages 222–236, 1951.
[2] S. E. Asch. Opinions and social pressure. 1955.
[3] 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.
[4] O. Burkeman. Memo exposes Bush's new green strategy | Environment | The Guardian, 2003.
[5] R. B. Cialdini. *Influence : science and practice.* HarperCollins, 2009.
[6] S. McLeod. Asch Experiment.
[7] S. A. McLeod. What is Conformity? - Simply Psychology, 2007.
[8] D. Murphy. Leaked Apple store manual reveals madness behind the genius | News | TechRadar, 2012.
[9] NSA. The world factbook Austria, 2011.
[10] Samsung. Samsung Smart TV 2013.
[11] M. Sherif. A study of some social factors in perception. *Archives of Psychology (Columbia University)*, 1935.
[12] Worldrobotcs.org. World Robotics - Industrial Robots 2012 Executive Summary. Technical report, 2012.

APPENDIX

Add github repo and descripton how to install and so on!!!

A. HEADINGS IN APPENDICES

The rules about hierarchical headings discussed above for the body of the article are different in the appendices. In the **appendix** environment, the command **section** is used to indicate the start of each Appendix, with alphabetic order designation (i.e. the first is A, the second B, etc.) and a title (if you include one). So, if you need hierarchical structure *within* an Appendix, start with **subsection** as the highest level. Here is an outline of the body of this document in Appendix-appropriate form:

A.1 Introduction

A.2 The Body of the Paper

A.2.1 Type Changes and Special Characters

A.2.2 Math Equations

Inline (In-text) Equations

Display Equations

A.2.3 Citations

A.2.4 Tables

A.2.5 Figures

A.2.6 Theorem-like Constructs

A Caveat for the TEX Expert

A.3 Conclusions

A.4 Acknowledgments

A.5 Additional Authors

This section is inserted by LATEX; you do not insert it. You just add the names and information in the `\additionalauthors` command at the start of the document.

A.6 References

Generated by bibtex from your .bib file. Run latex, then bibtex, then latex twice (to resolve references) to create the .bbl file. Insert that .bbl file into the .tex source file and comment out the command `\thebibliography`.

B. MORE HELP FOR THE HARDY

The acm_proc_article-sp document class file itself is chock-full of succinct and helpful comments. If you consider yourself a moderately experienced to expert user of LATEX, you may find reading it useful but please remember not to change it.