Linking actions to value categories
a first step in categorization for easier value elicitation

Moonen, D.D.M.; Tielman, M.L.

Publication date
2020

Document Version
Final published version

Published in
11th International Workshop Modelling and Reasoning in Context

Citation (APA)
Moonen, D. D. M., & Tielman, M. L. (2020). Linking actions to value categories: a first step in categorization for easier value elicitation. In *11th International Workshop Modelling and Reasoning in Context* (pp. 17-21)

Linking actions to value categories - a first step in categorization for easier value elicitation

Djoshua D. M. Moonen and **Myrthe L. Tielman** [1]

Abstract. Computer systems are increasingly involved in making decisions. Therefore, it is increasingly important that they understand our values. To make values usable, context is important, both of the individual and the actions they underlie. This work aims to study if it is possible to make it easier to elicit an individual's values by using the context of the action. Practically, we first held an expert survey (n = 7) to see if some values are more likely to underlie some actions than others. The results were positive on this score, so a second study (user, (n = 135)) was done showing that restricting the number of values made it easier to elicit values from users while not unnecessarily limiting their expression. This work shows that when linking actions to values, it is possible to make the elicitation easier by only showing the applicable options. This is an important step in being able to incorporate values in computerized decision making.

1 Introduction

Computer systems are increasingly helping us to make and stick to important decisions in life. Reminder systems, health apps and social-media blockers all function to help us change behavior in some way [5, 7]. However, such systems often blindly stick to a single goal, and do not truly understand the motivations behind our actions, nor the context in which we make our decisions. To help technology understand these motivations, values have been proposed [1]. Values represent the things we find important in life, and which guide our decisions [8]. Therefore, they have long been taken into account in system design [3]. However, to flexibly adapt to individual values, systems require values in the reasoning as well in the design. In recent years, a number of systems have attempted to model this reasoning by linking values to our choices in some way [2, 10]. Ideally, such work will lead to systems that can more flexibly adapt their decision making and take into account values in their reasoning [1].

Values are general, abstract concepts. However, for a system to use them, they need to be made concrete. They need to be linked to actions [10], or to choices [2]. Often, this is also done by transforming values into norms [3]. This concretization of values means that information needs to be added about the context in which they are applied. We identify two main types of context. Firstly, the individual needs to be taken into account, as people have different values, as well as different views on what a value means for them. Secondly, what type of choices or actions the value is applied to is relevant, values will take on different meanings in different domains.

The first type of context is the individual, which means that information about values should ideally come from them. The most obvious source for this information are the users themselves, but people have often not explicitly thought about values, or do not even fully understand the concept. Moreover, the conversational capabilities of many automated systems are not yet capable of this type of conversation. So this information is difficult for a system to obtain [6]. Therefore, most existing value-elicitation methods are based in human-human interaction [11], or are aimed at what values are important in general [9].

In order to make this elicitation of an individual's values easier, it is helpful to consider the second form of context, namely the action. Most systems have attempted to elicit values in general. But values can take on different meanings in different domains. For instance, safety might mean something different for choosing a car than for choosing a doctor. Similarly, the choice to go to work is motivated by different types of values than the choice to go to a party. This also means that we could use this type of context to narrow the conversation about values between a system and human.

If we want to know what value underlies a certain action for a specific individual, we could pose this as a question in which the user can pick from all possible values. However, this would mean a very large answer space. And as mentioned, the action probably also limits what values are most likely to underlie that choice. So it might be possible to use this context to limit the amount of possible values an individual has to pick from, for instance in the form of a pre-selection of the list of values. However, as we are interested in the individual's values, not just the most likely ones underlying a general action, it is also important to not limit the individual too much in what they can express by making this pre-selection too small. In this paper, we wish to explore whether it is possible to make elicitation easier in this way without limiting expression.

Thus, in this work we explore two things. Firstly, whether it is possible to make a pre-selection of values which are more likely to underlie a choice for a specific action. And secondly, whether a pre-selection like this makes it easier for users to select a value from a list while not limiting them in their expressive ability. In section 2 the first question is explored by means of an expert study. Section 3 explores the second question by means of a user study and 4 presents the results. A discussion and conclusion based on the findings can be found in section 5.

2 Value Categorization

In order to make value-selection easier, we propose to make a pre-selection based on the type of activity the value promotes. Our hypothesis is that different actions have different value types which often underlie them. For instance, the values which underlie people's choice to go to work are probably different from the one to watch a movie. In order to study whether such a pre-selection can be made and what it would be, an expert-study was performed. The goal of

[1] Delft University of Technology

this study was two-fold. Firstly, to see if there is agreement amongst experts in what categories of values are most likely to underlie the choice to perform a specific action. And secondly, if there is such agreement, what categories of values are most likely for what actions.

2.1 Participants

The study was conducted with 7 participants (71.4% male), recruited from research staff and PhD students of Delft University of Technology. All participants were familiar with or have worked on value-based topics. Average age was 33.4 (sd 7.2) and they had an average of 3.83 years (sd 4.41) of experience with value-based research.

2.2 Procedure

The participants were sent a survey along with instructions. The instructions defined value as used by Schwarz (1992) including a detailed description of each of the 10 value categories [8]. Participants were asked to consider 40 actions, and for each indicate which top three of value categories would be most likely to underlie a person's choice to perform those actions. The full list of actions can be seen in Table 1. These actions were selected in such a way that the list represented a diverse set of daily activities, and the authors felt all value categories were most likely to be covered at least once.

2.3 Measures

After the surveys were filled in, the anonymized data was aggregated. This was done by counting the frequency of each value category in the 1st, 2nd and 3rd places for each action. Then, first place was awarded a score of 4, second place a score of 2 and third place a score of 1 for each time it appeared in said place. The scores were summed up such that every value category received an overall score per action. This formula was chosen such that a first place was worth a little more than a third and second place combined, and the same as two second places combined. After this score was created, a threshold of 9 was chosen in order to determine which categories were most relevant for each action. All categories scoring 9 or over were marked as relevant. This threshold was chosen such that each action had at least has one value category above the threshold.

2.4 Results

Table 1 shows the full results, marking each value category's score for each of the included actions. The rightmost column shows the difference between the mean score and the maximum score per action. This number indicates how much agreement existed between experts, with higher numbers indicating more agreement. Furthermore, it shows which value categories were marked by the experts as being relevant (above the threshold of 9) in red/bold.

From Table 1 the average distance from the highest score to the mean was computed, which is 11.4 on average. This indicates that for many actions a value category exists which scores visibly better than the rest. After all, to get an overall score of 11, at least 3 of the 7 participants needed to have scored one particular category in at least 2nd place. To get this number as difference from the mean score, this means the majority of the 7 experts agreed on the highest scoring category. This consensus indicates that we might, indeed, use value categories that are in Table 1 to pre-select what values a user can choose from. However, more work is necessary to study if this pre-selection truly does not limit users in the expression of their values, as well as to know if it actually achieves its goal of making value selection easier.

3 User Study

The results from the expert study show the potential of using a pre-selection of possible values based on the action. The goal of this pre-selection would be to make it easier for users to indicate what values underlie decisions to perform actions. However, it is important that people do not feel this pre-selection limits their freedom of expression, as the pre-selection is not meant to push users into giving certain answers. To study these two aspects, an online between-subject user study was performed. Participants were asked what value would most likely underlie an action. Half were only shown the pre-selection to pick from, while participants in the other condition were shown the full list of values from Schwarz [8].

3.1 Participants

For this study, participants were recruited via Amazon Mechanical Turk. 297 started the survey, and 231 completed it. Of these 231, 64 did not answer the control question correctly and were, therefore, excluded. Of the 167 remaining, 8 filled in the survey twice, and the data of their second time was deleted, leaving 159. One final participant was excluded because they did not collect their payment, leaving us with 158 participants included in the initial analysis.

When looking at this initial data, we noticed that some of the participants had only clicked once on the pages with the questions, namely for going to the next page. This can be taken as evidence that they did not look at the full drop down list of values, just leaving the first, default answer in place. In some cases, this might just indicate that the default answer seemed correct, but some participants also did this for every question. In the end, it was decided to remove participants that had answered 10 or more questions within a second of seeing the page, as it would've been nearly impossible for them to have fully read a question in that time. The threshold of 10 was chose due to it being over half of the questions. This way 23 participants were removed. This made the final number of participants included in the analysis 135.

3.2 Procedure

The participants were asked to fill in a survey. The survey starting with some general information, followed by asking for informed consent of the participants. After obtaining consent the participants were placed in 1 of 2 conditions after which 19 questions were asked where the amount of answers was dependant on the condition the participant were in. The 19 questions were asked in random order where on each question the answers were also in random order. The survey concluded by asking the participants 5 questions on their experience completing the survey.

3.3 Measures

We measured the total time spent to complete the survey and the first click, last click, the total amount of clicks and time at which the questions was submitted. The difference between time of the first and last click was used to measure the time actually spent on each of the questions. This metric proved to be useful as some of the participants had

Table 1. Weighted numerical representation of action per value category.
Achievement(AC), Benevolence(BE), Conformity(CO),F Hedonism(HE), Power(PO), Security(SE), Self-Direction(SD), Stimulation(ST), Tradition(TR), Universalism(UN). First place is worth 4 points, second 2 and third 1. The mean to highest represents the difference from the highest to the average score. Highlighted in red/bold are the value categories higher or equal then 9, so marked as relevant for that action

Promoted activity	AC	BE	CO	HE	PO	SE	SD	ST	TR	UN	Mean to highest
Act politely	2	**12**	**11**	0	1	1	4	0	4	7	7,8
Buy something	8	4	4	**11**	1	4	5	4	1	0	6,8
Care for someone	2	**20**	1	2	2	4	2	4	0	6	15,7
Celebrate holiday	0	0	2	**11**	4	5	2	5	**13**	0	8,8
Communicate	5	4	3	0	4	2	3	8	1	**12**	7,8
Compete	**10**	0	2	1	7	1	2	8	4	0	6,5
Cook	**9**	0	0	**13**	0	**10**	4	2	3	1	8,8
Create something (e.g. painting)	**10**	0	0	6	1	0	**11**	**12**	1	1	7,8
Decide what to do	4	0	0	1	**11**	0	**20**	4	0	2	15,8
Do something exciting	6	0	2	**16**	1	1	2	**14**	0	0	11,8
Drink	4	0	2	**20**	0	4	3	4	4	1	15,8
Eat	0	2	6	**12**	0	**14**	3	0	5	0	9,8
Enjoy art	0	0	0	**15**	2	4	3	**10**	2	6	10,8
Exercise	**13**	0	0	2	2	**11**	**9**	5	0	0	8,8
Exercise influence	4	7	0	1	**18**	0	4	8	0	0	13,8
Follow a ceremony	0	0	**11**	4	0	3	1	1	**20**	2	15,8
Follow the law	0	1	**22**	4	0	7	0	1	4	3	17,8
Help someone	1	**18**	2	0	4	2	4	1	0	**10**	13,8
Learn	8	0	2	1	2	2	**16**	6	0	5	11,8
Make decisions for others	8	5	0	1	**18**	0	5	0	3	2	13,8
Make money	**13**	0	0	**11**	8	8	5	0	0	0	8,5
Meditate	2	7	1	5	1	2	**16**	4	4	0	11,8
Perform (e.g. a play)	**11**	4	1	2	2	0	6	**15**	0	1	10,8
Plan your day	**10**	0	2	4	3	1	**18**	0	2	0	14
Play games	2	0	3	**11**	0	0	6	**14**	5	1	9,8
Pray	2	2	1	0	0	8	6	4	**17**	2	12,8
Protect others	0	**18**	4	0	5	8	0	0	1	6	13,8
Protect your belongings	1	0	4	2	5	**24**	2	0	1	2	19,9
Protect yourself	2	1	2	0	7	**20**	4	0	5	1	15,8
Read	0	4	1	4	2	1	**9**	**11**	0	**10**	6,8
Relax	0	6	1	**18**	0	8	6	1	0	2	13,8
Repair something (e.g. car)	**18**	2	0	5	4	1	1	**9**	0	2	13,8
Sleep	0	1	0	**11**	3	**16**	8	2	0	0	11,9
Spend time with family	0	6	4	8	1	3	0	6	**10**	4	5,8
Spend time with friends	0	5	2	**9**	4	5	6	**9**	1	1	4,8
Study	**10**	0	0	5	2	0	**12**	6	1	6	7,8
Take responsibility	2	**11**	0	2	**16**	2	5	0	1	3	11,8
Travel	1	0	2	**12**	0	0	**11**	**15**	0	1	10,8
Watch movies	2	0	1	**18**	0	0	2	**13**	4	2	13,8
Work	**11**	0	1	0	4	8	**11**	6	1	0	6,8

taken breaks over 10 minutes long before the first click on a question was made, so we could not look at total time spent on the page. The first 19 questions were regarding values, there the last 5 questions were about the participants' experience taking the survey. These 5 consisted of 4 questions about the difficulty of the survey, followed by 1 question asking if the participant was missing the option for the answer they wanted to give. The first 4 questions regarding difficulty of the survey used a 5-point Likert scale ranging from -2 (Extremely difficult) via 0 (Neither easy nor difficult) to 2 (Extremely easy). The last question regarding missing answer options used a 4-point Likert scale ranging from 1 (Only some of the questions) to 4 (All of the questions).

4 Results

The data was analyzed with R version 3.6.1 and the analysis was split into 3 parts. The first part is analyzing the time spent on questions about values. The second part is on the questions regarding difficulty of the survey. And the third and last part is on the perceived lack of answers to the questions of the survey.

The time spent on the questions on values was analysed by using the mean time spent per question. The Shapiro-Wilk normality test was used, indicating that the data was not normally distributed ($W = 0.77$, $p < 0.01$). Therefore the Wilcoxon rank sum test with continuity correction was used, indicating that a significant difference exists between conditions in the amount it took for people to answer what value was most relevant ($W = 3068$, $p < 0.01$).

Difficulty was tested with four questions. In order to create a single difficulty score, the questions had their internal cohesion tested using Cronbach's alpha and were found to be internally cohesive (α.83). The Shapiro-Wilk normality test shows the data was not normally distributed (W = 0.95, $p < 0.01$). Therefore the Wilcoxon rank sum test with continuity correction was used, showing significant difference in the answers on questions regarding the difficulty of the survey between the two conditions (W = 1394.5, $p < 0.01$).

The question regarding freedom of answers was analysed separately. On average, people indicated that they could answer as they wished for 'most of the answers' (3) for both conditions (all answers: M=2.95, SD=0.73, pre-selection: M=3.01, SD=0.86). As the data was not normally distributed (Following Shapiro-Wilk W = 0.79, $p < 0.01$), the Wilcoxon rank sum test with continuity correction was used, showing no significant difference between the two groups regarding their experience of missing answers (W = 2112.5, $p = 0.455$).

5 Discussion and Conclusion

The results show that participants that received the pre-selection spent significantly less time on average per value question, implying that it was easier to select an answer from the pre-selection. This was probably partly because there are less answers to consider, but could also be because people already had had an answer in mind and it would take less time to find their answer. Overall this means that the survey with pre-selected answers was less of a time investment, and that it was potentially easier to complete. This implication is supported by the results from the questionnaire, which also show that the participants that received the pre-selection found the survey significantly easier to complete. One concern with only presenting people with a pre-selection would be that it limits people's freedom of expression. However, our results show no significant difference in the amount of times people wanted to pick a value which was missing from the list. Note that the average score of both conditions indicated that they were able to find their value for 'most of the actions'. Therefore, we found no evidence that making a pre-selection lead to people feeling restricted in their expression.

5.1 Contributions

Values are abstract concepts, but when a system needs to use them, they need to be seen in the context of both the individual and what actions they are applied to. In this work, we use the context of these actions to inform us about what values are most likely, in order to more easily elicit values from an individual. More specifically, this study shows that it is possible to present a pre-selected list of values to participants based on the context of the action it is applied to. This pre-selected list makes the process of picking underlying values faster and easier to perform, without it affecting the freedom of expression perceived by participants. This is important as this technique can be used by systems to learn what values underlay an individual's choice to perform an action. In this way, values can be used by system's to adjust their advice and decision making processes, and to align better with their users. Values form a large part of the moral context in which people make decisions, so it is important that we take steps to allow systems to understand these better [1].

5.2 Limitations

Firstly, our pre-selection was based on a limited number of expert participants. Although our results indicate that this was a good pre-selection, we do not assume full consensus on what this should look like. For a fully validated pre-selection of what value type corresponds to what action, more work would need to be done. However, our main intention was to study whether such a pre-selection was even possible in the first place and we believe this smaller sample was enough to show that this is indeed the case. Secondly, the questions about difficulty and perceived amount of missing answers used self-reported data for the analysis. We do not fully know to what extent people truly found it more difficult because of the long list, or because the selection made values easier to think about. Moreover, the results with respect to freedom of answers were all relatively high, which might indicate a ceiling effect. Although we did not find that a pre-selection limited people's perceived freedom in choice, this might be because they simply could not think of anything else. However, when presented with a full list some people might still pick things which were not in the pre-selection. As we did not show the same people both the full and the pre-selection lists, a direct comparison like this was not possible.

5.3 Future Work

Firstly, this paper focused on a pre-selection on values for ease of use. At the moment, you need to have the pre-selection for each specific action. To be able to scale up to any arbitrary set of actions it would be worthwhile to explore the existence of a groupings of actions that share the same values. The possibility exists that values can be extrapolated, making it easier for the system to scale in the amount of actions. Secondly, this paper only looks at actions to narrow down a pre-selection of possible underlying values. However, in indicating what value underlies an action, more contextual factors might play a role. Things like time of day, weather and surrounding actions might be relevant. But a good starting point for taking into account more context might also be social situation. Social norms are highly dependent on our values, so whether we perform an action with friends or with colleagues might change what value underlies it [4]. More work is necessary to see whether such additional context factors would allow for better pre-selections of values. Finally, this paper assumes that the answers filled in by the participants in the surveys are representative of their beliefs. However, talking about values is difficult, and so is verifying whether what people say about their values matches with what they actually value in practise. Therefore, it would be interesting to see to what extent the answers given in the survey coincide with the values that the participants actually hold.

5.4 Conclusion

Values are increasingly being incorporated in technology, but their elicitation remains difficult. In this work, we explore whether it is possible to make value elicitation for specific actions easier by presenting people with a pre-selection containing only those values most relevant to that action context. In an expert study, we found that there is indeed some consensus on what value categories are most likely to correspond to an action. This indicates that it is indeed possible to make a pre-selection of most relevant values based on the actions that are looked into. Additionally, in a user study with such a pre-selection we found that it made it easier for people to choose the most likely underlying value for an action, without diminishing their perceived freedom of choice. These results are important for the process of value elicitation and through that of value-based reasoning, which is becoming more important in today's society where we increasingly interact with technology on a personal level.

Acknowledgement: This work is part of the research programme CoreSAEP, with project number 639.022.416, which is financed by the Netherlands Organisation for Scientific Research (NWO).

REFERENCES

[1] *Ethically Aligned Design - A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2*, The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems., 2017.

[2] S. Cranefield, M. Winikoff, V. Dignum, and F. Dignum, 'No pizza for you: Value-based plan selection in BDI agents', in *International Joint Conference on Artificial Intelligence*, (2017).

[3] Batya Friedman, Peter H. Kahn Jr., and Alan Borning, *Human-Computer Interaction and Management Information Systems: Foundations Advances in Management Information Systems, Volume 5 (Advances in Management Information Systems),*, chapter Value Sensitive Design and Information Systems, 348–372, M.E. Sharpe, 2006.

[4] Ilir Kola, Catholijn M. Jonker, and M. Birna van Riemsdijk, 'Modemodel the social environment: Towards socially adaptive electronic partners', in *International Workshop Modelling and Reasoning in Context (MRC), Held at FAIM*, (2018). AAMAS/IJCAI Workshop on Modeling and Reasoning in Context.

[5] Eleonora Milić, Dragan Janković, and Aleksandar Milenković, 'Health care domain mobile reminder for taking prescribed medications', in *ICT Innovations 2016*, eds., Georgi Stojanov and Andrea Kulakov, pp. 173–181, Cham, (2018). Springer International Publishing.

[6] Alina Pommeranz, *Designing Human-Centered Systems for Reflective Decision Making*, Ph.D. dissertation, Delft University of Technology, 2012.

[7] Danielle E. Schoffman, Gabrielle Turner-McGrievy, Sonya J. Jones, and Sara Wilcox, 'Mobile apps for pediatric obesity prevention and treatment, healthy eating, and physical activity promotion: just fun and games?', *Translational Behavioral Medicine*, **3**(3), 320–325, (2013).

[8] Shalom H Schwartz, 'Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries', in *Advances in experimental social psychology*, volume 25, 1–65, Elsevier, (1992).

[9] Shalom M. Schwarz, Gila Melech, Arielle Lehmann, Steven Burgess, Mari Harris, and Vicki Owens, 'Extending the cross-cultural validity of the theory of basic human values with a different method of measurement', *Journal of Cross-Cultural Psychology*, (2001).

[10] M.L. Tielman, C.M. Jonker, and M.B. van Riemsdijk, 'What should I do? Deriving norms from actions, values and context', in *International Workshop Modelling and Reasoning in Context (MRC), Held at FAIM*, (2018). Under revision at the AAMAS/IJCAI Workshop on Modeling and Reasoning in Context.

[11] Ibo van de Poel, *Translating Values into Design Requirements*, chapter Philosophy and Engineering: Reflections on Practice, Principles and Process, Springer, 2013.