# A Disclosure Intimacy Rating Scale for Child-Agent Interaction

Franziska Burger, Joost Broekens, and Mark A. Neerincx

TU Delft, Interactive Intelligence
Delft, The Netherlands
http://ii.tudelft.nl

**Abstract.** Reciprocal self-disclosure is an integral part of social bonding between humans that has received little attention in the field of human-agent interaction. To study how children react to self-disclosures of a virtual agent, we developed a disclosure intimacy rating scale that can be used to assess both the intimacy level of agent disclosures and that of child disclosures. To this end, 72 disclosures were derived from a biography created for the agent and rated by 10 university students for intimacy. A principal component analysis and subsequent k-means clustering of the rated statements resulted in four distinct levels of intimacy based on the risk of a negative appraisal and the impact of betrayal by the listener. This validated rating scale can be readily used with other agents or interfaces.

**Keywords:** Long-term cHRI · self-disclosure intimacy · PAL project

## 1 Motivation

In a focus group conducted with diabetic children in 2012, it was found that children would like a companion robot to share their secrets with and to listen to them when they are sad [1].

According to Self Determination Theory (SDT) [3], successful establishment of a social bond between human and agent leads to sustained motivation both to interact with the agent and to engage in activities that the agent proposes. Such a bond could be established through increasingly intimate, reciprocal self-disclosures [4], that is the exchange of information about the self.

One of the key interests in human-human self-disclosure research has been the close link between disclosure and liking. For example, it was found that 6th grade children's liking of another child was influenced by that child's capacity to match the intimacy level of a disclosure while that of 4th graders was not [5].

To better study children's disclosure behavior when interacting with a virtual agent, we developed the Dyadic Disclosure Dialog Module (3DM) within the framework of the PAL project[1] and using a situated Cognitive Engineering (sCE) [2] approach. This, in turn, necessitated the development of a rating scale for intimacy of self-disclosure.

---

[1] http://www.pal4u.eu/

## 2 Intimacy Rating Scale

To design agent disclosure statements at various intimacy levels and to assess the depth of children's disclosures, a rating scale for disclosure intimacy was needed. For this, the following constraints were identified: (a) the scale should discretize the intimacy continuum, (b) each discrete level should have a clear definition, (c) the scale should have a minimum of three levels [6, Ch. 13], (d) the scale should be neither topical nor example-based. Upon reviewing the relevant child and adult literature on self-disclosure, no entirely suitable intimacy scale could be found. We therefore developed and validated the Disclosure Intimacy Rating Scale (DIRS).

As summarized in [7], intimacy of self-disclosure is directly related to vulnerability of the discloser. In a similar vein, it is argued in [8] that the social risk associated with disclosing determines the depth of disclosure. With each self-disclosure, we risk "social rejection [or] betrayal" [8, p. 180].

$$risk(\text{SD}) = risk(\text{SR}) + risk(\text{B}) \tag{1}$$

with SD := self-disclosure, SR := social rejection, and B := betrayal. Betrayal, here, describes the passing on of information by the recipient to third parties.

Risk can be formalized as the product of probability ($P$) and impact ($I$). If we further assume that social rejection does not occur at random but only follows if the disclosure is negatively appraised, we can approximate the risk of social rejection through the risk of negative appraisal:

$$risk(\text{SD}) = P(\text{NA}) * I(\text{NA}) + P(\text{B}) * I(\text{B}) \tag{2}$$

with NA := negative appraisal.

The probability of betrayal, $P(\text{B})$, can depend only on characteristics of and prior experiences with the disclosure recipient. It is therefore independent of the content and cannot be considered in the level definitions.

These considerations initially yielded six intimacy levels. Using these, a total of 6(*level*) x 3(*topic*) x 2(*valence*) x 2(*repetition*) = 72 statements were fabricated by the first author with the personality and biography of the ECA providing content and style information. To obtain a first validation of the scale, the statements were rated for intimacy by 10 university students (5 female, $M_{age} = 23$, $SD_{age} = 1.612$) on a six-point scale: only levels *0* and *5* were labeled with *not at all intimate* and *extremely intimate* respectively. We decided against asking adult participants to take on the perspective of a child (because results would be questionable in terms of validity) or to rate statements as if coming from a robot (because students are more critical towards the plausibility of a robot expressing emotions and a personality). The biography was hence slightly adapted to fit a 22 year-old student. Before rating, participants were asked to read a persona description of the student and instructions explaining self-disclosure. Intimacy was defined as: "the degree to which a statement reflects information about the self that is sensitive." Further, they were given one example disclosure for each level using a fourth topic. The intimacy levels of

the examples was not provided. Participants could thus get an impression of the covered range and the type of statements. Participants found the description of the student and the statements to be believable (the mean believability rating on a 5-point Likert scale was $M_{believability} = 4.3$). The inter-rater reliability was assessed using the two-way random intraclass correlation coefficient with the ten raters, yielding $ICC(2, 10) = .947$. Cronbach's alpha using all items was high with $\alpha = .948$. The Pearson correlation coefficient between the level of an item and the average rating it received across participants was determined to be $r = .85$. To check whether we would also find six intimacy levels back in the item pool, a principal component analysis was conducted on the ratings of all items. Using the point of inflexion as a cut-off criterion [9], four principal components explaining at least $10\%$ of the variance each and $67\%$ in total were revealed. *Four* was then used as the desired number of clusters in a k-means clustering algorithm. A post-analysis of the resulting item clusters afforded the four intimacy levels of the DIRS detailed in Table 1.

Table 1: The four intimacy levels of the DIRS that resulted from the post-analysis.

| Risk | Definition | Example |
| --- | --- | --- |
| low | $P(NA)$, $I(NA)$, and $I(B)$ are low or zero: the discloser cannot be evaluated on the basis of the statement or the statement is very common-place. | "I have a lot of brothers and sisters." |
| moderate | $P(NA)$ is moderate, because statements are more opinionated, but $I(NA)$ and $I(B)$ are low. Negative appraisal can at best take the form of disagreement. The information cannot really be exploited, so that in the case of betrayal, no loss is to be expected. Includes preferences and opinions on activities and objects. | "I like online games in which you have to team up with other players." |
| high | Either $P(NA)$ is high and both $I(NA)$ and $I(B)$ are low (the content conflicts with the norms of the recipient but does not reflect on the character of the discloser), or $P(NA)$ is low but the content is of great significance to the discloser so that $I(NA)$ and $I(B)$ are high. Disclosures are emotional and may include evaluations of other people. | "I'm really disappointed that my sister will not try yoga with me. She already promised it twice but never followed through." |
| very high | $P(NA)$, $I(NA)$, and $I(B)$ are high, because the disclosure is at the core of the discloser's self-concept and could easily conflict with the norms of the recipient. In the case of betrayal, great emotional, physical, or material damage may ensue. Social stigmas, self-doubt, deep personal fears and secrets are accumulated on this level. | "Whenever I work really hard or I'm nervous, I start sweating like crazy. I can't get close to people then, because I'm really conscious of how I smell." |

## 3 Conclusion

3DM is intended to gain insights into how and how readily children in late childhood disclose to an artificial agent. Whether children absolutely, relatively,

or do not at all match the intimacy level of a robot's disclosure [10] being a main matter of interest. The DIRS is a supplementary instrument for 3DM to code and compare the intimacy levels of children's disclosures in response to agent disclosures. In an exploratory study using 3DM, 114 child-disclosures were rated by two independent raters using the DIRS. Raters agreed in 67 % of cases and deviated by 1 level in 27 % of cases. However, the disclosures were gathered in the field over the course of two weeks and children were found to mainly disclose on the lower two levels to the ECA (only in 26 % of disclosures was the mean rating of both raters larger than 1).

There are two main limitations to the DIRS. The first is that contextual information of the disclosure is unknown or ignored, and can only be estimated by the rater. As such, raters should have the same cultural background as the discloser. An additional limitation is that the DIRS has only been validated with adults, but is used with children in the PAL project. A next step is therefore to validate it with children of the target age group and using the original biography of the ECA.

In summary, we developed and validated the Disclosure Intimacy Rating Scale to rate statements for intimacy. This scale can be readily used across different human-robot interaction contexts.

## References

1. Baroni, I., Nalin, M., Baxter, P., Pozzi, C., Oleari, E., Sanna, A. and Belpaeme, T.: What a robotic companion could do for a diabetic child. In 2014 RO-MAN: The 23rd IEEE International Symposium on Robot and Human Interactive Communication, 936 – 941. IEEE Press (2014)
2. Neerincx, M.A. and Lindenberg, J.: Situated cognitive engineering for complex task environments. Ashgate Publishing Limited (2008)
3. Deci, E.L., and Ryan, R. M.: Overview of self-determination theory: An organismic dialectical perspective. In: Handbook of self-determination research, 3 – 33. University Rochester Press (2002)
4. Altman, I. and Taylor, D.: Social penetration theory. Holt, Rinehart &\ Mnston, New York (1973)
5. Rotenberg, K.J. and Mann, L.: The development of the norm of the reciprocity of self-disclosure and its function in children's attraction to peers. Child Development, 1349 – 1357 (1986)
6. Tardy, C.H.: A handbook for the study of human communication: Methods and instruments for observing, measuring, and assessing communication processes. Greenwood Publishing Group (1988)
7. Moon, Y.: Intimate exchanges: Using computers to elicit self-disclosure from consumers. Journal of consumer research 26(4), 323 – 339 (2000)
8. Omarzu, J.: A disclosure decision model: Determining how and when individuals will self-disclose. Personality and Social Psychology Review 4(2), 174 – 185 (2000)
9. Cattell, R.B.: The scree test for the number of factors. Multivariate behavioral research 1(2), 245 – 276 (1966)
10. Rotenberg, K.J. and Chase, N.: Development of the reciprocity of self-disclosure. The Journal of genetic psychology 153(1), 75 – 86 (1992)