

Evaluating Social Attitudes of a Virtual Tutor

Florian Pecune¹, Angelo Cafaro¹, Magalie Ochs², and Catherine Pelachaud¹

¹ LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay

² Aix Marseille Université, CNRS, ENSAM, Université de Toulon, LSIS

Abstract. In this paper we evaluate a model of social decision-making for virtual agents. The model computes the social attitude of a virtual agent given its social role during the interaction and its social relation toward the interactant. The resulting attitude influences the agent's social goals and therefore determines the decisions made by the agent in terms of actions and communicative intentions to accomplish. We conducted an empirical study in the context of virtual tutor-child interaction where participants evaluated the tutor's perceived social attitude towards the child while the tutor's social role and relation were manipulated by our model. Results showed that both role and social relation have an influence on the agent's perceived social attitude.

1 Introduction

In order to improve the naturalness and the believability of virtual anthropomorphic agents, socio-emotional components should be considered when modeling their decision-making and behavior in a human-agent interaction. These components allow agents to express an emotional behavior and a social attitude relevant to the context of the interaction. According to Scherer [15], a social attitude is an *"affective style that spontaneously develops or is strategically employed in the interaction with a person or a group of persons, coloring the interpersonal exchange in that situation"*. The spontaneous aspect of social attitudes can be defined as the social relation between interactants. For instance, two people who like each other spontaneously tend to comply with the other's requests, thus showing a friendly attitude [7]. The strategic aspect can be governed by the interactants' social roles in their social context [8]. For example, in a restaurant (social context), a waiter (social role) is supposed to be nice and polite (social attitude) toward its clients, while a teacher might be more authoritative toward its students. We investigated the effects of the social relation of an agent compared to its social role on the perceived social attitudes that a user attribute to that agent. More specifically, is an agent liking its interactant *always* considered friendly regardless of its role during the interaction? Is an agent that wants to be authoritative always perceived as dominant?

In order to give a virtual agent the capability of reasoning about its role and its own social relation toward the user, and the capability of expressing an adequate in-context social attitude, we designed and evaluated an agent's model of social decision making that integrates spontaneous and strategic aspects of social attitudes.

2 Related Work

The most common dimensions used for representing an agent’s social relation toward another refer to the notion of *affiliation* (whether the agent likes or dislikes the other) and *status* (whether the agent has power over the other) [1]. In [5], authors evaluated the *status* and *affiliation* of their virtual recruiter through verbal cues and non-verbal behaviors. The social attitude of a virtual recruiter during an interview was also evaluated in [6]. In this work, recruiters conveyed *status* and *affiliation* through sequences of non-verbal signals. In [13], authors also used agent’s non-verbal behavior as a cue, as well as interruption strategies. Participants were asked to evaluate *status* and *affiliation* of two particular agents among a group of four discussing characters. Participants had to interact with a virtual museum guide in [3]. After a short interaction, they were asked to evaluate the *affiliation* of the guide. Here, the agent’s social attitude was defined by its amount of smiles, mutual gaze with the participant, and its proximity (whether it was standing close or far from the participant). These studies mainly focused on the verbal and nonverbal behavior of the agents, and social attitudes were evaluated by third parties. None of them, however, focused on the actions of the agents. Therefore, our main contribution is a study protocol designed to evaluate an agent’s social attitudes through sequences of actions.

3 SocRATES Model

The focus of this paper is the validation of SocRATES, our computational model of social attitudes. The purpose of this model is to build a virtual agent able to reason about its social role and its social relation towards its interactant, and thus select its actions accordingly. First, we compute the social attitude expressed by the agent according to its social relation toward its interactant and its social role. Then, considering its social attitude, the agent has two social goals: it wants its interactant to express (1) the same level of affiliation and (2) opposite level of status. Finally, the agent chooses its action according to importance given to its social goals and its task-oriented goals. A complete description of SocRATES can be found in [12].

Figure 1 shows a schematic representation of our model’s implementation. We first used PsychSim [10] for defining a set of actions and their influence on the states of the world for each agent. Social and situational goals were implemented as agent’s reward functions. When an agent plans its next action in PsychSim, it first evaluates the effect of each of its possible future actions on the different states of the world. Then it predicts the other agent’s expected actions and their impact on the world’s states. Then again, the agent will anticipate its reaction until a given *horizon* (i.e. number of steps). When the agent finishes its projection, it evaluates the overall effect of each sequence of actions according to its goals. Then, the agent selects the action with the highest expected utility. SocRATES enables each agent to take its decision according to its social attitude, while computing the influence of its actions on an other agent’s social relation.

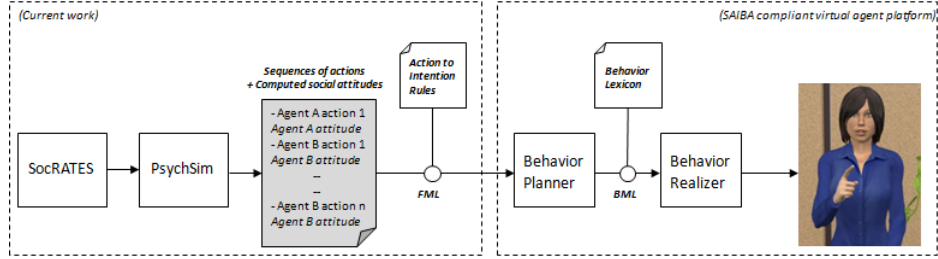


Fig. 1. Overview of the architecture: The left part of the image depicts SocRATES and its PsychSim implementation. The right part shows the SAIBA compliant virtual agent platform used to perform the generated sequences of actions and communicative intentions.

More details about this part of the implementation can be found in [12]. The outputs of SocRATES and PsychSim are actions and social attitudes generated for both interactants. For each interactant, both outputs are dynamically computed on a turn-by-turn basis and by taking into account the other interactant’s previous turn (i.e. action and attitude expressed). The set of actions for each interactant is transformed into a sequence of communicative intentions through a mapping from actions to FML [4] (described with more detail in Section 4.1). As shown in the second block of Figure 1, the produced communicative intents are accomplished through sequences of multimodal behaviors by relying on the Greta platform [11].

4 Experimental Design

We performed an evaluation study to measure the influence that agent’s initial social relation and social role have on its perceived interpersonal attitude by a user. Given the influence of *actual social relation* and *ideal social relation* on the agent decision making in SocRATES, we aimed at checking whether the sequences of actions computed by the model convey the expected social attitudes. Since many works have shown the influence of tutor’s social attitude on child’s motivation [20], we defined a scenario depicting a tutor-child interaction where an animated conversational agent plays the role of the tutor. The tutor’s situational goal is to make the child do its homework. Participants were asked to rate the perceived level of social status and affiliation of the tutor. As exploratory measure given by the social context, we also asked participants to rate the perceived level of performance of the tutor.

4.1 Stimuli

We created a series of tutor-child interaction videos. Since gender and/or visual appearance can influence users’ perceived social attitudes of an agent [9], in our study we controlled this aspect by using the same female character as a tutor

throughout all the videos. The child was represented by an androgynous still figure in order to avoid any gender or behavior biases. Figure 2 shows a screenshot from a video stimuli as shown to participants. The child’s still figure is shown on the left with a dynamic label underneath describing its current action. Neither speech nor behavior of the child were shown in order to steer the participants’ attention as much as possible on the tutor’s side. The animated virtual tutor was shown on the right. We used a synthesized voice accompanied by facial expressions, gestures, gaze behavior and head movements generated by the model described in [11] for expressing the tutor’s communicative intentions. A red square highlighted the agent (tutor or child) that had the turn (i.e. speaking or doing an action) during the interaction.



Fig. 2. A screenshot of the video used as stimuli in our study. The child is represented by a still figure on the left and the tutor is the animated virtual agent on the right.

The different videos were generated by systematically varying the tutor’s initial *actual social relation* and its *ideal social relation* using our model (see [12] for examples of the generated interactions). The sequences of actions were all different, however, the verbal and non-verbal behaviors used to represent each action were the same across the interactions (i.e. the mapping from action to communicative intention). We relied on [5] to represent the communicative intentions using a neutral verbal and non-verbal behavior. Thus, we identified a between-subjects variable that was the tutor’s initial *actual liking* (**T-InitialActualLiking**) towards the child. This variable has two levels: *Negative Actual Liking* vs. *Positive Actual Liking*. Since the tutor is able to influence the child’s goals, we fixed its initial actual power to a positive value. Then, once the initial *actual liking* was defined, we identified a within-subjects factor, the tutor’s **ideal social relation** (**T-IdealRelation**). Considering our scenario, we identified the three following levels: positive ideal power ($P+$), positive ideal liking ($L+$) and, positive ideal power and positive ideal liking ($P+L+$). In sum, we obtained six different videos. Figure 3 represents the evolution of the mean tutor’s social attitude computed by our model in the six different interactions. The scenario ended whenever the child finished its homework.

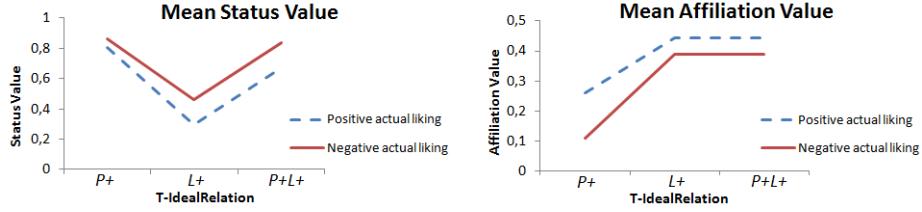


Fig. 3. The left figure represents the mean status value (y axis) according to the *ideal social relation* (x axis) for both positive *actual liking* (plain line) and negative *actual liking* (dashed line).

4.2 Measurements

We asked the participants to rate the perceived attitude of the tutor towards the child in terms of social status (**T-Status**) and social affiliation (**T-Affiliation**), and the perceived efficiency of its strategies (**T-Performance**).

For measuring **T-Status** and **T-Affiliation**, we adapted eight statements from the third person version of the InterPersonal Questionnaire (IPQ-R) [17], which is built on Wiggin’s circumplex representation of attitudes [19]. The IPQ-R questionnaire defines twelve interpersonal styles (i.e. duodecants) representing different blends of the two attitude dimensions. Each duodecant was linked to a sub-questionnaire containing six statements. For measuring status, we selected two statements from the *dominant* duodecant and two statements from the *deferent* duodecant. The statements for the *dominant* duodecant are: “Trying to control the child” and “Assertive toward the child”. The items adopted for the *deferent* duodecant are: “Avoiding imposing power over the child” and “Unauthoritative with the child”. For measuring affiliation, we selected two statements from the *warm* duodecant and two statements from the *aloof* duodecant. The statements for the *warm* duodecant are: “Warm toward the child” and “Taking strong interest in the child’s goals”. The items adopted for the *aloof* duodecant are: “Unfriendly toward the child” and “Uninterested in the child’s goals”.

For measuring **T-Performance** we selected the three most reliable items from the *efficacy for student engagement* factor found in the Ohio State teacher efficacy scale (OSTES) [18] according to their score. This questionnaire is used to measure teachers’ performance along three factors. The three selected statements are: “Helping the child to understand the importance of learning”, “Trying to motivate the child” and “Getting the child to believe s/he can do well”.

All answers were on a 7-points labeled Likert scale (anchors: 1. *Completely disagree* and 7. *Completely agree*).

4.3 Procedure and Participants

We recruited a total of 48 participants via academic mailing lists (24 in each group). 52% of the participants were between 18 and 30 years old, 48% were above 30. 60% were male, 40% were female. They had different cultural backgrounds, but the two most prominent groups were France (31%) and Portugal (15%). We ran this study online. Participants were first presented with a short demographic survey to know about their nationality, age and gender. Once the survey was completed, we randomly assigned the participants to a group according to the **T-InitialLiking** variable (positive vs. negative). Finally, we presented each stimulus as a video on a different web page with questions underneath by following a within-subjects repeated measures design and fully counterbalancing conditions' order as recommended by [2].

4.4 Hypotheses

Given Scherer's definition [15], we suppose that the tutor's initial *actual social relation* and its *ideal social relation* have both an influence on its perceived social attitude. Moreover, given that a tutor's social attitude has an influence on its performance, as suggested by [20] we suppose that the tutor's initial *actual social relation* and its *ideal social relation* have both an influence on its perceived performance. According to the social attitudes computed by our model and presented in Fig.3 we defined the following hypotheses:

H.1-Sta: The **T-IdealRelation** will have a main effect on perceived **T-Status**, in particular a tutor with a positive level of *ideal power* (P+) will be perceived with a higher status compared to a tutor with both positive *ideal power* and positive *ideal liking* (P+L+) that in turn will be perceived with a higher status compared to a tutor with a positive *ideal liking* (L+).

H.1-Aff: The **T-IdealRelation** will have a main effect on perceived **T-Affiliation**, in particular tutors with a positive *ideal liking* (L+) or with both positive *ideal power* and positive *ideal liking* (P+L+) will be perceived with a higher affiliation compared to a tutor with a positive level of *ideal power* (P+).

H.1-Perf: The **T-IdealRelation** will have a main effect on perceived **T-Performance**, in particular tutor with both positive *ideal power* and positive *ideal liking* (P+L+) will be perceived with a higher performance compared to a tutor with a positive *ideal liking* (L+) that in turn will be perceived with a higher performance compared to a tutor with a positive level of *ideal power* (P+).

H.2-Sta: The **T-InitialActualLiking** will have a main effect on perceived **T-Status**, in particular tutor with (Negative Actual Liking) will be perceived with a higher status compared to a tutor with (Positive Actual Liking).

H.2-Aff: The **Initial Actual Liking** will have a main effect on perceived **T-Affiliation**, in particular tutor with (Positive Actual Liking) will be perceived with a higher closeness compared to a tutor with (Negative Actual Liking).

H.2-Perf: The **Initial Actual Liking** will have a main effect on perceived **T-Performance**, in particular tutor with (Positive Actual Liking) will be perceived with a higher performance compared to a tutor with (Negative Actual Liking).

5 Results

We conducted a two way repeated measures MANOVA (i.e. multivariate repeated measures analysis of variance) with T-InitialActualLiking as between-subjects factor and T-IdealRelation as within-subjects factor. The dependent measures were tutor’s status (**T-Status**), tutor’s affiliation (**T-Affiliation**) and tutor’s performance (**T-Performance**).

The MANOVA revealed two overall significant main effects of T-InitialActualLiking ($Wilks' Lambda = .57, F(3, 44) = 10.75, p < .001, \eta_p^2 = .44$) and T-IdealRelation ($Wilks' Lambda = .20, F(6, 41) = 26.30, p < .001, \eta_p^2 = .80$). The analysis also indicated a significant interaction effect ($Wilks' Lambda = .35, F(6, 41) = 12.23, p < .001, \eta_p^2 = .64$). Since the sphericity assumption was not violated, we performed a follow-up analysis that looked at univariate effects for each dependent measure with two-way Mixed ANOVAs. These analyses confirmed the significant main effects and interaction of T-InitialActualLiking and T-IdealRelation on our three measurements (all $p < .001$ and effect sizes were ranging from **.20** to **.72**).

This would suggest that **both** tutor’s initial *actual liking* of the child and its *ideal relation* have effects on our dependent measures. We conducted a post-hoc analysis by further analyzing the effects of the within-subjects factor, T-IdealRelation, by performing pairwise multiple comparisons with Bonferroni adjustments. Whereas for the between-subjects factor, T-InitialActualLiking, we ran a Simple Main Effects analysis. In Table 1 we report a summary of all means and standard errors (in parentheses) for the 3 dependent variables (DVs) as indicated in the table’s heading. The columns describe levels of the within-subjects factor (i.e. T-IdealRelation) corresponding to *positive ideal power* (P+), *positive ideal liking* (L+) and *positive ideal power and liking* (P+L+). The two rows correspond to the two levels of the between-subjects factor (T-InitialActualLiking). For instance, the mean value of the rated status was 5.65 for the interaction where the tutor had a positive *actual liking* and a positive *ideal power*. The follow-up simple main effect analysis revealed that all differences between the two groups were significant ($p \leq .001$) except for those when the tutor had positive ideal power (P+). Those means that did not significantly differ between groups are marked with “*” in the table.

We found that tutor with a positive *ideal power* was perceived with a lower affiliation and performance than tutor with positive *ideal liking* and tutor with both positive *ideal liking and power*, but tutor with both positive *ideal power and liking* (P+L+) was perceived with a lower affiliation and performance than a tutor with a positive *ideal liking* (L+). Thus, our hypotheses **H.1-Aff** and **H.1-Perf** are partially supported. We also found that tutor with a positive ideal power was perceived with a higher status than tutor with positive *ideal liking* and tutor with both positive *ideal liking and power*, but the perceived status was not significantly different between a tutor with a positive *ideal liking* and one with both positive *ideal power and liking*. Thus, **H.1-Sta** is partially supported. Participants rated the tutor with a positive *actual liking* with higher status, affiliation and performance than tutor with a negative *actual liking*. However, the

DVs:	T-Status			T-Affiliation			T-Performance		
	P+	P+L+	L+	P+	P+L+	L+	P+	P+L+	L+
Positive	5.65*	4.05	3.95	2.96*	5.43	5.68	1.97*	5.48	5.73
Initial Liking	(±.19)	(±.22)	(±.20)	(±.27)	(±.24)	(±.25)	(±.23)	(±.25)	(±.26)
Negative	5.80*	5.45	5.28	3.09*	3.51	4.15	2.22*	2.70	3.44
Initial Liking	(±.19)	(±.22)	(±.20)	(±.27)	(±.24)	(±.25)	(±.23)	(±.25)	(±.26)

Table 1. Summary of means and standard errors in parentheses for the 3 dependent variables (DVs). The differences between the means marked with “*” (i.e. according to the Initial Actual Liking levels) were not significant. All comparisons among the within-subjects factor’s levels were significant ($p \leq .001$) except for the effects of $P+L+$ and $L+$ on tutor’s *Status*.

difference between T-InitialActualLiking for the two groups was not significant when the tutor had a positive *ideal power* (P+). Thus, hypotheses **H.2-Sta**, **H.2-Aff** and **H.2-Perf** are partially supported.

6 Discussion and Future Work

We found that participants were able to perceive the tutor’s social attitudes when our model was generating sequences of actions according to the tutor’s initial *actual liking* and its *ideal social relation*. More specifically, we found that both tutor’s initial *actual liking* and tutor’s initial *ideal social relation* had main effects on its perceived status, affiliation and performance.

As hypothesized, a tutor with a positive *ideal power* was perceived with a significantly higher status than a tutor with a positive *ideal liking* and both positive *ideal power* and *liking*, but with a significantly lower affiliation and performance (**H.1-Sta**, **H.1-Aff** and **H.1-Perf** partially supported). In our generated sequences of actions, tutors with a positive *ideal power* almost immediately switched off the child’s console for forcing it to work, thus possibly explaining the perceived low level of tutor’s affiliation. Moreover, tutors with positive *ideal power* preferred using coercive strategies more than explaining to the child the importance of doing its homework and thus a lower performance was attributed to them. However, a tutor with both positive *ideal power* and *liking* was perceived with a lower performance and affiliation than a tutor with only positive *ideal liking*. We believe that tutors with both positive *ideal power and liking* are more pragmatic, thus perceived as less friendly than tutors only aiming at increase a child’s liking towards themselves. Moreover, we think that the duration of the interaction was too short for participants in order to identify the child’s goals, therefore to judge the tutors’ performance. Moreover, since we did not model the content of the child’s exercises, it was difficult for participants to judge the quality or the correctness of their homework.

There was a significant interaction effect between the tutor’s initial *actual liking* and *ideal social relation* on tutor’s perceived status, affiliation and performance. In simple words, the initial *actual liking* had the effect of “amplifying” the outcomes on all dependent measures. Tutor’s with a positive *actual liking* were perceived with a significantly higher affiliation and performance than tutors with a negative *actual liking*, but they were perceived with a significantly lower status (**H.2-Sta**, **H.2-Aff** and **H.2-Perf** partially supported). We believe that when the tutor liked the child since the beginning it was displaying more *immediacy* towards the child, therefore increasing affiliation and rapport with it as described in [14]. Another possible explanation is that first impressions induced a cognitive bias which led participants to rate tutors trying to please the child from the beginning with a higher affiliation. Tutors with a positive *actual liking* might have been considering negotiating with the child to please it, which could have been considered as a sign of submission. However, tutors with positive *ideal power* were perceived with the same level of status, affiliation and performance, no matter their initial *actual liking*. One possible explanation is that tutors intending to be dominant immediately used coercive strategies (i.e. switching off the child’s console) without trying to explain to the child the importance of working.

Some future work should be considered. Concerning our model, we didn’t take interpersonal rigidity [16] into account when computing an agent’s social attitude as the mean of *actual* and *ideal social relation*. Interpersonal rigidity theory assesses that people with a high level of rigidity tends to maintain the same social attitude through the whole interaction (i.e. considering their *ideal social relation* as more important than their *actual social relation*), whereas people with a low level of rigidity are considered more flexible as they can adapt their attitude according to their interactant’s behavior (i.e. considering their *actual social relation* as more important than their *ideal social relation*). Thus, the tutor’s level of rigidity would change the importance accorded to its *ideal social relation* and its *actual social relation* when computing its social attitude. Regarding the evaluation, we consider evaluating our model in a context-free interaction, to check whether our results could be generalized outside a tutor-child scenario.

7 Acknowledgement

This work was partially performed within the Labex SMART (ANR-11-LABX-65). It has also been partially funded by the French National Research Agency project MOCA (ANR-12-CORD-019).

References

1. Argyle, M.: Bodily communication. Methuen, New York, 2nd ed. edn. (1988)
2. Bradley, J.V.: Complete counterbalancing of immediate sequential effects in a latin square design. *Journal of American Statistical Association* 53(282), 525–528 (1958)

3. Cafaro, A., Vilhjálmsón, H.H., Bickmore, T., Heylen, D., Jóhannsdóttir, K.R., Valgárdhsson, G.S.: First impressions: Users' judgments of virtual agents' personality and interpersonal attitude in first encounters. In: *Intelligent Virtual Agents*. pp. 67–80. Springer (2012)
4. Cafaro, A., Vilhjálmsón, H.H., Bickmore, T., Heylen, D., Pelachaud, C.: Representing communicative functions in saiba with a unified function markup language. In: *Intelligent Virtual Agents* (2014)
5. Callejas, Z., Ravenet, B., Ochs, M., Pelachaud, C.: A computational model of social attitudes for a virtual recruiter. In: *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. pp. 93–100 (2014)
6. Chollet, M., Ochs, M., Pelachaud, C.: From non-verbal signals sequence mining to bayesian networks for interpersonal attitudes expression. In: *Intelligent Virtual Agents*. pp. 120–133. Springer (2014)
7. Cialdini, R.B., Goldstein, N.J.: Social influence: Compliance and conformity. *Annu. Rev. Psychol.* 55, 591–621 (2004)
8. Goffman, E.: Presentation of self in everyday life. *American Journal of Sociology* 55, 6–7 (1949)
9. Krämer, N.C., Karacora, B., Lucas, G., Dehghani, M., Rüter, G., Gratch, J.: Closing the gender gap in stem with friendly male instructors. on the effects of rapport behavior and gender of a virtual agent in an instructional interaction. *Computers & Education* (2016)
10. Marsella, S.C., Pynadath, D.V., Read, S.J.: Psychsim: Agent-based modeling of social interactions and influence. In: *Proceedings of the international conference on cognitive modeling*. pp. 243–248 (2004)
11. Niewiadomski, R., Bevacqua, E., Mancini, M., Pelachaud, C.: Greta: an interactive expressive eca system. In: *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*. pp. 1399–1400. International Foundation for Autonomous Agents and Multiagent Systems (2009)
12. Pecune, F., Ochs, M., Marsella, S., Pelachaud, C.: Socrates: from social relation to attitude expressions. In: *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*. pp. 921–930. International Foundation for Autonomous Agents and Multiagent Systems (2016)
13. Ravenet, B., Cafaro, A., Biancardi, B., Ochs, M., Pelachaud, C.: Conversational behavior reflecting interpersonal attitudes in small group interactions. In: *International Conference on Intelligent Virtual Agents*. pp. 375–388. Springer (2015)
14. Richmond, V., McCroskey, J., Hickson, M.: *Nonverbal communication in interpersonal relations*. Allyn and Bacon, 6th ed. edn. (2008)
15. Scherer, K.R.: What are emotions? and how can they be measured? *Social science information* 44(4), 695–729 (2005)
16. Tracey, T.J.: Interpersonal rigidity and complementarity. *Journal of Research in Personality* 39(6), 592–614 (2005)
17. Trapnell, P.D., Broughton, R.H.: The interpersonal questionnaire (ipq): Duodecant markers of wiggins' interpersonal circumplex (2006), unpublished data, The University of Winnipeg
18. Tschannen-Moran, M., Hoy, A.W.: Teacher efficacy: Capturing an elusive construct. *Teaching and teacher education* 17(7), 783–805 (2001)
19. Wiggins, J.S., Trapnell, P.D.: A dyadic-interactional perspective on the five-factor model. In: Wiggins, J.S. (ed.) *The five-factor model of personality: Theoretical perspectives*, pp. 88–162. Guilford Press, New York, NY, USA (1996)
20. Wubbels, T., Opdenakker, M.C., Den Brok, P.: Let's make things better. In: *Interpersonal relationships in education*, pp. 225–249. Springer (2012)