

Exploring the use of Online Simulations in Teaching Dialogue Skills

Psychology Learning & Teaching

2023, Vol. 22(1) 55–73

© The Author(s) 2022



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14757257221138936

journals.sagepub.com/home/plj**Michiel H. Hulsbergen** 

DialogueTrainer B.V., The Netherlands

Jutta de Jong 

University Medical Centre, Utrecht, The Netherlands

Maarten J. van der Smagt Experimental Psychology, Helmholtz Institute, Utrecht University,
The Netherlands

Abstract

To aid in teaching dialogue skills a virtual simulator called *Communicate!* was developed at Utrecht University. Teachers can build scenarios for students to practice dialogues with a virtual character. In two experiments ($n = 128$ and 133 , a year apart), we investigated if and how *Communicate!* can be an effective aid to study and practice dialogue skills, by comparing it to traditional “passive” learning tools, such as literature-study and a lecture, in an undergraduate psychology dialogue-skills course. Students were divided into four groups, two of which both read an article about conducting a bad-news dialogue and played a *bad-news-dialogue-scenario* (but in a different order), while the third group only played the scenario. The final group only read the article (expt. 1) or also attended a lecture on the topic (expt. 2). Playing a scenario improved performance on a different scenario played later. It increased the students’ reported engagement and motivation to learn about this topic, compared to reading the article, whereas their reported self-efficacy decreased, which may indicate a recognized learning need. It also increased the score on an MC-knowledge test on this type of dialogue. This suggests that online dialogue simulations aid studying (basic) dialogue skills, by providing flexible, authentic learning experiences.

Keywords

Simulation, dialogue skills, communication, active versus passive learning

Corresponding author:

Maarten J. van der Smagt, Director of Psychology Education at the Faculty of Social and Behavioural Sciences and Principal Fellow at the Centre for Academic Teaching of Utrecht University, Utrecht, The Netherlands.

Email: M.J.vanderSmagt@uu.nl

Introduction

Communication plays a pivotal role in many professions (e.g., Silverman et al., 2013), yet academic education, especially at the bachelor level, is predominantly aimed at developing theoretical knowledge. Even in programs in disciplines such as medicine and psychology that revolve around treating patients, advising others, or collaborating with colleagues, *verbal* communication skills training is rare (e.g., Deveugele et al., 2005). This is problematic, because it leaves students underprepared for their professional careers after university. In sectors ranging from care institutions to government organizations and professional service organizations, ineffective communication is linked to coordination problems, diminished innovative potential, suboptimal career paths, and even burnout (Buhler & Worden, 2013; Mastebroek et al., 2014), which may result in considerable costs for professionals and clients, organizations and society as a whole.

Learning to communicate starts in childhood and continues well after academic education. Traditionally, communication was to a large extent learned as part of a hidden curriculum that did not find its way into formal education (e.g., Hafferty, 1998; Windolf, 1981). Still, basic communication training, and advanced conversation methods may be a welcome addition to perhaps every curriculum, as communication is relevant in many fields (Lala et al., 2017). In healthcare psychology, at least in the Netherlands, verbal communication (or dialogue) skills are a prerequisite for admission to post-master's programs needed to acquire national registration for individual healthcare professionals, and therefore need to be taught at bachelor and/or master levels.

For many professional conversations, best-practice models exist, for example, for delivering bad news (Van der Molen et al., 1995), persuading others (Cialdini, 1984; Perloff, 2008), negotiating (Fisher et al., 1981) and motivating people (Miller & Rollnick, 2013) to name a few, but how to teach these to students? The predominant mode of instruction at academic institutions remains lecturing (Handelsman et al., 2004; Lee et al., 2020; Konopka et al., 2015). This form of "passive learning" is associated with lower performance in various fields, including science, technology, engineering, and mathematics (Freeman et al., 2014). Dialogue skills are eminently tacit knowledge, which require automated processes, including intuition and emotions. People's explicated cognitions rarely explain to the full what people actually do (e.g., Demetriou & Wilson, 2009). This means that to learn these skills, active (learner-centered) learning is required to be effective (e.g., Berkhof et al., 2011; van der Vleuten et al., 2019). Organizing an environment for active learning is however challenging for various reasons, related to what is learned, how learning takes place, and how we can provide a practical and affordable learning environment.

Dialogue training as tacit knowledge is exceptional, because of the dynamic circumstances in which skills are applied. Any conversation involves at least two parties with their own goals and motives, who set objectives and—either explicitly or implicitly—adapt their strategies based on different interpretations (Hargie, 2011). In addition, each student differs in onset skill levels and talents, which necessitates adaptive instruction. Meanwhile, how instructions land with students, remains largely hidden from view, as students may even shield off their mental processes to save face. Verbal communication is very personal.

This is where active learning through digital simulations may provide solutions. As people learn through experience and feedback, learning requires that we provide an environment that mimics the actual environment of skill application, at least to an extent that it provides a relevant challenging experience. In the current academic training environment, prevalent methods include professional training actors and roleplay exercises between students (e.g., Berkhof et al., 2011; Deveugele et al., 2005). However, each of these methods has its drawbacks. Professional training actors are expensive, which means training all students is unrealistic. Professional actors are thus mostly reserved

for assessment situations. Roleplay with fellow students is cost-efficient and does enable individual training. This form of practice however is suboptimal as knowing each other personally may hinder students' ability to get immersed in the roleplay and a lack of experience with the actual real-life situation means fellow students mostly fail to identify with what is at stake. This means counterplay may not include the complex emotions essential to the experience. So how can we make verbal communication training scalable? Technology might provide an answer. As digital flight simulators enable pilots to practice challenging circumstances endlessly in a safe learning environment, the same counts for simulations with many other complex skills (Aldrich, 2004). Simulation-based learning appears to be an effective means to facilitate the learning of a variety of complex skills in a variety of academic settings (Chernikova et al., 2020).

In medical education, where virtual patients are more prevalent, studies conclude that interactions with digital patients can be effective in the development of skills such as giving bad news and acquiring information (see Bosman et al., 2019 and Lee et al., 2020 for extensive reviews). Responsive simulations enable the student to apply present knowledge, and by providing direct feedback they inform students about what they already know and where they fall short. The case and the context of the simulation clarify the skills' importance, as they show the consequences of failure and success. Simulations thrive on motivation: the motivation to perform, and when one falls short, the motivation to learn (Aldrich, 2009).

To enable experiential (verbal) communication training with simulations for large groups of students, between 2013 and 2016 Utrecht University in the Netherlands developed a virtual simulator aimed at training dialogues. This simulator—*Communicate!*—created by the department of information and computer sciences, offers a virtual character that speaks by textual or spoken statements. Players interact with this virtual character (avatar) by choosing between answer options. Each choice leads to a different course of the conversation and is rewarded with a score, related to a learning objective. *Scenarios* for the simulator are created with an editor (see Figure 1, top panel), in which statements of the virtual character are matched to expressions of emotions and answer options, and include scores and written feedback to the student. As part of educational innovation, this platform is now used to train doctors, pharmacists, veterinarians, and psychologists (Jeuring et al., 2015). Although more platforms presently exist (e.g., Kron et al., 2017, for a platform that is at least visually similar), *Communicate!* adds the possibility for teachers themselves to build and adapt simulations. This means simulations can be continuously improved through use. By the time of this research, the bad news conversation model involved had gone through several rounds of improvements of both the best practices nodes, alternative answer options, and feedback.

How does learning through such a platform compare to traditional communication teaching instruments such as written text and lectures? These well-tested and commonly used teaching methods clearly have value (e.g., French & Kennedy, 2017), and are designed to help students gain knowledge about conversation dynamics and behavioral patterns. This knowledge is assumed to help students set reasonable objectives in situations and evaluate the effect of a route of action. As compared to the passive “classical” learning instruments, the online *Communicate!* platform adds interactivity as the virtual character responds to choices with utterances and facial expressions that reveal how the character feels and how different routes of action influence this (see Figure 1, bottom panel; Jeuring et al., 2015). As players interact with the character, the answer options invite players to explicate their assumptions about what could be effective. Responses of the virtual character, therefore, become meaningful, as they either confirm or falsify these assumptions, which may either increase confidence or give rise to new problem definitions and hypotheses. To add theory to the interaction, the platform

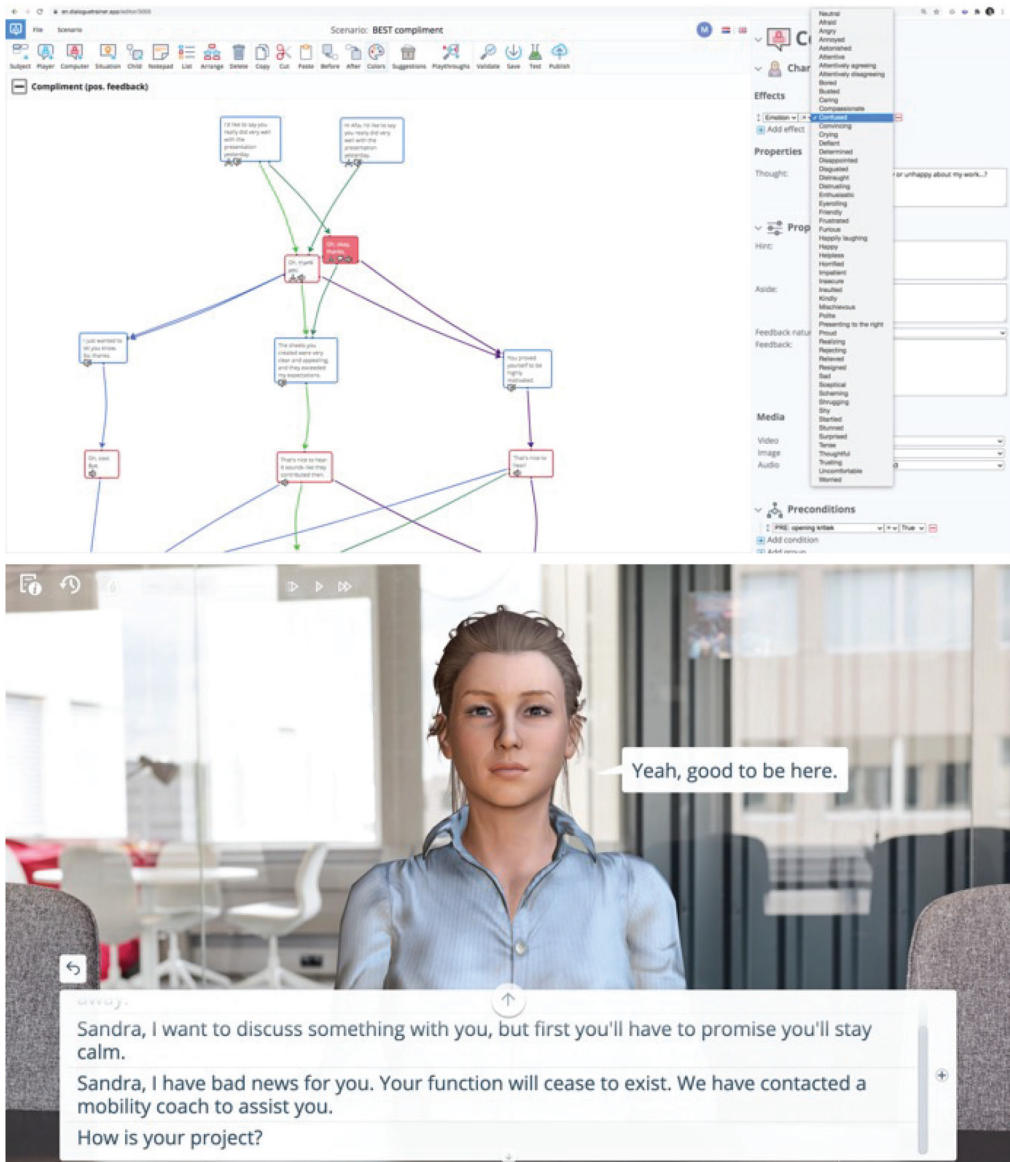


Figure 1. Impression of the editor (top) and user interface (bottom) of the current version of the DialogueTrainer platform (the successor to the *Communicate!* platform used in the experiments). The interface and editor are mainly the same, but the graphics are of higher quality in the more recent version. (For an impression of the current user experience see <https://www.dialoguetrainer.com/en/>).

provides written feedback in a report after a playthrough. In comparison to the more general theory conveyed through traditional teaching instruments, here the explanations provided are specific to choices users make, which we assume result in increased awareness of the relevance of theory.

The quality of a simulation is thought to depend partly on its realism (e.g., Aldrich, 2009). The interaction within the *Communicate!* platform is clearly distinguishable from an actual conversation. Students sit in front of a computer, which shows a character that is obviously animated. The interaction of selecting answer options is also notably different from uttering spontaneous responses. As was found in the research of other platforms, this does not necessarily disqualify this form of interaction, as evidence suggests users also learn with this approach (Lee et al., 2020). Since achieving objectives in the simulation depends on one's ability to apply conversation knowledge, it is expected that interaction with the platform is to some extent similar to having an actual conversation. As something is at stake and each decision results in an effect, we assume that emotions in the simulation, as promoters of interest (e.g., Frijda, 1986), are to some extent similar as well.

To assess the quality of the platform as a teaching instrument, and advance knowledge about the instructional design of dialogue simulations, we set out to explore the possible effects of the simulations (and platform) on learning as part of a real-life dialogue skills training course in a Psychology bachelor program in the Netherlands. We set up two related one-day experiments during subsequent iterations of the course, comparing the effect of using the *Communicate!* platform on possible changes in motivation toward the objective of the course (acquiring dialogue skills), immersion in the (topic of) the learning process, whether learning transferred to a different context (i.e., playing a different, but comparable, scenario on the same platform) and whether the use of the platform affected acquired theoretical knowledge.

Materials and Methods

Research Setting

The psychology curriculum in the Netherlands is divided into a three-year bachelor's program and a one-year master's program (with additional post-master education for healthcare psychologists). In the second or final year of the bachelor program, students participate in a 10-week "professional dialogue skills training course." The course is obligatory for most students, and one can choose between several different modules within this course depending on the specialization within the bachelor's program. In spring 2015 and 2016, one module of this course focused on bad news conversations and involved the use of the *Communicate!* platform.

Participants

All participants ($n = 128$ in 2015; $n = 133$ in 2016) were bachelor students specializing in Social and Organizational Psychology at the Utrecht University in the Netherlands (age range approximately between 19 and 24) of which about 80% were female.¹ Participation was voluntary and no course credits or monetary compensation was offered. Participation was offered as an extra training opportunity within the course. All those starting the session finished it also. Participants gave written consent, and the study was approved by the faculty ethics assessment committee of Social and Behavioural Sciences at Utrecht University.

Study Design

In both 2015 and 2016, the experiment was set up as a quasi-randomized controlled trial. Students were assigned to one of four groups (see Table 1). This assignment was as follows: each student was

Table 1. Experimental flow.

Course 1	2015	(n = 128)						
Group	n	Intervention 1	Q1	Intervention 2	Q2	Knowledge test	Performance test	
A (S▶A)	28	Scenario 1	✓	Article	✓	MC	Scenario 2	
B (A▶S)	35	Article	✓	Scenario 1	✓	MC	Scenario 2	
C (A only)	29	Article	✓	—	—	MC	Scenario 2	
D* (S only)	36	Scenario 1	✓	—	—	MC	Scenario 2	
Course 2	2016	(n = 133)						
Group	n	Intervention 1	Q1	Intervention 2	Q2	Knowledge test	Performance test	
A (S▶A)	31	Scenario 1	✓	Article	✓	MC	Scenario 2	
B (A▶S)	35	Article	✓	Scenario 1	✓	MC	Scenario 2	
C (A▶L)	30	Article	✓	Lecture	✓	MC	Scenario 2	
D (S only)	37	Scenario 1	✓	—	—	MC	Scenario 2	

Q1 is the first questionnaire after Intervention 1 and Q2 is the second questionnaire after Intervention 2.

*Due to technical problems the data from Group D in Course 1 (2015) could not be obtained.

part of one of four workgroups in the course. Students from the first workgroup were assigned to Group A, the second to Group B, and so on. Since allocation to a specific workgroup in the course during enrolment occurred on a random basis, the allocation to our experimental groups can be deemed random. Each group was presented with one or two instructive interventions about *giving bad news*, in a different order. In the first experiment (2015) Group A first played a scenario about delivering bad news and then read an article on the subject. Group B first read the article and then played the scenario. Group C only read the article and Group D only played the scenario.² In the second experiment—in 2016—the interventions for Groups A, B, and D were identical to the first experiment. The only difference was that Group C read the article, which was followed by a lecture on delivering bad news.

After each of the interventions, an online “motivation” questionnaire was filled out by each student. In addition, a set of “emotion cards” (e.g., Hulsbergen et al., 2019) was presented to gauge the emotions students experiences during the interventions. However, as the use of emotion cards was not yet validated at the time, the effects hereof were not further explored.

Each intervention, including filling out the questionnaires took about 1.5 h. Around 4 h after the start of the experiment, all participants took part in a short multiple-choice knowledge test about delivering bad news, which was followed by each participant playing a different bad news scenario.

Interventions

The article students read was a chapter from van der Molen et al. (1995, pp. 143–154), which is commonly used in the Dutch psychology curriculum. Both scenarios were based on the same chapter, in terms of phases in the conversation, alternative routes of action which relate to learning goals, and the effects of those routes of action. The scenarios presented students with a practical situation, namely having to refer a patient to an alternative therapist, which the patient is expected to find undesirable. The scenarios were based on theory from the article, combined with best-practice expertise from professionals drawn from interviews. This resulted in a believable course

of the interaction, including answer options in line with what players would say and believable responses from the virtual character. Scenarios were tested individually before implementation, to improve the matching of answer options to the students' learning goals.

Scenarios were "played" using the *Communicate!* platform, a simulation platform for dialogue training (see Jeurig et al., 2015 for a more detailed description). After reading an introduction, players engage in an interaction with a virtual avatar by choosing between various statements (see Figure 1, bottom panel). One of these statements is in line with a best-practice approach. Two or three alternative answer options are based on prevalent learning objectives as described in theory, in this case, van der Molen et al. (1995; see above). Each player's choice leads to a different response from the avatar and a different course of the conversation. As the platform enables scoring every choice, playthroughs also lead to different score profiles based on predefined parameters.

The lecture was based on the book chapter but added a perspective on the function of emotions during the five phases of processing bad news or mourning as described by Kubler-Ross (1969). These phases correspond to the phases that van der Molen et al. (1995) present as the best-practice approach.

Instruments and Outcome Measures

Questionnaire. The (5-point Likert-scale) questionnaire was aimed at six aspects of motivation:

Immersion and engagement (statements such as: "I felt the case situation was convincing"), *Usability* (statement such as: "I thought the method was easy to use"), *Motivation* to learn about the subject (statements like: "This learning method has aroused my curiosity in the subject matter"), *Task value* (example: "I think the skills taught are useful for me"), *Control beliefs* ("If I try hard enough, then I will be able to master 'bad news' dialogue skills") and "*Self-efficacy*" (example: "I am certain that I understand the dialogue method presented in this activity") taken from several validated questionnaires.

Immersion and Usability were translated and adapted from the *System Usability Scale* (Brooke, 1996) and the thereon-based questionnaire from Persson et al. (2014). Since the *Communicate!* platform presented an until then not encountered learning method in the program, we were interested in comparing how each learning/studying method appealed to students and whether it was deemed easy to use. The other scales were translated and adapted from the *Motivated Strategies for Learning Questionnaire* (Pintrich & de Groot, 1990; Pintrich et al., 1991). These scales were used to probe how the different learning methods influenced the way students experienced their learning process. Does it encourage students to learn more about the topic (*Motivation*), do students think is useful (*Task value*), do they feel in control during the process (*Control beliefs*), and is their confidence about what they have learned affected (*Self-efficacy*)?

For each participant, the average score on each of the scales on the motivation questionnaire was calculated. Cronbach's alpha was calculated for each scale and appeared reasonable (all $\alpha = .67$ or higher, except for control beliefs ($\alpha = .53$). See the Supplemental material for the questions and reliability per scale).

Performance Test. Because *Communicate!* is a digital instrument that logs users' interactions within the scenarios, scores can be calculated from these user interactions. Parameters in this specific simulation were based on theoretic constructs derived from the theory of van der Molen et al. (1995). To score playthroughs, theoretic constructs were operationalized from this theory as the following

parameters: clarity/transparency, empathy, tact/security, and “Method van der Molen.” The best practice derived from van der Molen’s description was scored as an optimal route on which players can maximize a score on the “Method van der Molen” parameter, as this construct was regarded to match the *best practice*. Consecutively, other choices in the simulation were scored based on assumptions about the described constructs, on which feedback to players was also based. Thus, a construct such as empathy was operationalized as a correct “understanding of the avatar’s needs,” as indicated by the player’s choice to either offer emotional support or provide information. Clarity/transparency related to being open toward the virtual character about the news, reasons, and consequences, and tact/security related to bringing the news across in a way that would minimize chances that aggression could turn against the player as a messenger. Scores on the parameters (possible maximum range –10 to 10) were added to each decision item (node) in the scenario, based on the above. To measure players’ effectiveness in the scenarios, the sum (across nodes) of the scores for the four parameters was automatically added up and converted into a single total percentage score. The highest score on the simulation required that the *best practice* was followed. Choices outside of the best practice led to a lower score on the “Method van der Molen” parameter but might still score on the other parameters.

Knowledge Test. The knowledge test consisted of 10 multiple-choice questions, based on the book chapter. On several questions, more than one answer was correct, such as: “Which emotions can you expect to encounter during a bad news talk?” of which all options were correct. For these questions, one point was given for each correct option and deducted for each wrong option. For the other (regular) multiple-choice questions, the correct answer yielded one point. This resulted in a maximum possible score of 17 points, which were converted into percentage scores. Cronbach’s alpha for the knowledge test overall was 0.71.

Statistical Analyses

Since the present study is exploratory in nature we conducted separate analyses on each motivation questionnaire scale, on game scores, and on knowledge test scores. All analyses were performed in JASP 0.16 (JASP Team, 2022). Where assumptions of normality and/or homoscedasticity were violated, bootstrap procedures were applied using JASP’s R module, based on Berkovits et al. (2000) and Spychala et al. (2020; code obtained from <https://nadinespy.github.io>).

Questionnaire. For the motivation questionnaire, we focus on the interaction between the group and the instance of the questionnaire (after intervention 1 or after intervention 2). This leaves Groups A (first scenario, then article) and B (first article, then scenario) in both experiments (2015 and 2016) and Group C (first article, then lecture) in 2016 only. First, we conduct a three-way mixed analysis of variance (ANOVA) on the effect of the intervention order, with a group (Groups A and B only) and year as between-subjects factors and questionnaire instance (1 or 2) as a within-subjects factor. If a year does not interact with a group, we subsequently collapse the data across the two experiments to gain more power in comparing these groups. To investigate the effect a lecture has as a second intervention, we then focus on 2016 only in a 2 (Group B vs Group C) \times 2 (questionnaire instance) ANOVA with a group as between and questionnaire instance as a within-subjects factor.

Performance Test. The performance test was analyzed similarly to the questionnaires, with a three-way mixed ANOVA on the effect of the intervention order, with group (Groups A and B only) and year as between-subjects factors and scenario instance (1 or 2) as a within-subjects factor. If a year

does not interact with a group we subsequently collapse the data across the two experiments to gain more power in comparing these groups. To investigate the effect of having a second intervention, we then focus on 2016 only in a 3 (Groups A, B, and D) \times 2 (scenario instance) ANOVA with a group as between and scenario instance as between factor.

Knowledge Test. The knowledge test was analyzed separately for each experiment. We used a one-way ANOVA for each year to compare the test scores between groups.

Results

We will describe our three outcome measures separately, first focusing on the responses to the motivational questionnaires. How do the different interventions influence the students' immersion, experiences, motivation to learn, and self-efficacy? Then, we will focus on a practical performance measure. Do the different interventions influence performance in playing a new *bad news* scenario? Last, we will focus on whether the different interventions influence the students' performance on a knowledge test on the subject.

Motivation Questionnaires

Each of the scales analyzed adhered to the assumptions for a regular ANOVA. We first conducted a $2 \times 2 \times 2$ mixed ANOVA with questionnaire instance (1 or 2; see Table 1) as a within-subject factor, and Group (A or B) and year (2015 or 2016) as between-subject factors, for each of the motivation questionnaire scales. For none of the six motivation scales, the year of the experiment interacted with the group or questionnaire instance and group (all $F \leq 3.15$, $p \geq .08$), so the data were collapsed across the two experiments. A mixed ANOVA was conducted with Group (A: S \blacktriangleright A and B: A \blacktriangleright S) as the between-subject factor and questionnaire instance (Q1 or Q2) as a within-subject factor. Group and questionnaire instances interacted for both usability scales and two of the motivation scales (see Supplemental Tables S1 and S2). We will zoom in on those scales by using simple main effects analyses.

Usability Scales. For both the *Immersion* and *Usability* scales simple effects analyses showed that for Group A (S \blacktriangleright A), scores do not differ between questionnaire instances (for both scales $F \leq 2.95$, $p \geq .09$; see Figure 2 and Supplemental Table S3). However, scores on these scales did increase significantly for Group B (A \blacktriangleright S) after playing the scenario (immersion: 3.23 [0.53] vs 3.74 [0.73], $F = 36.7$, $p < .001$; usability: 2.97 [0.49] vs 3.26 [0.65], $F = 12.28$, $p < .01$). At the start of the experiment students that read the article and those that played the scenario scored similarly on immersion. It was only when playing the scenario *after* reading the article that students felt more immersed in the scenario (left panel, Figure 2). On the other hand, students that played a scenario at the start of the experiment were more inclined to state the method of instruction was easy to use and this did not change (or only marginally) when reading the article afterward. Students that started by reading the article indicated that the scenario they played afterward was easier to use (right panel, Figure 2).

Motivation Scales. For both the *Motivation to learn* (left panel, Figure 3) and *self-efficacy* (right panel, Figure 3) scales, simple effects analyses showed scores for Group A (S \blacktriangleright A) did not differ between questionnaires filled out after the two interventions ($F \leq 3.13$, $p > .08$; Supplemental Table S4). Yet, the scores for Group B (A \blacktriangleright S) indicated an increased motivation to learn about the subject after playing a scenario, but *only* if an article on the subject was read first (3.56

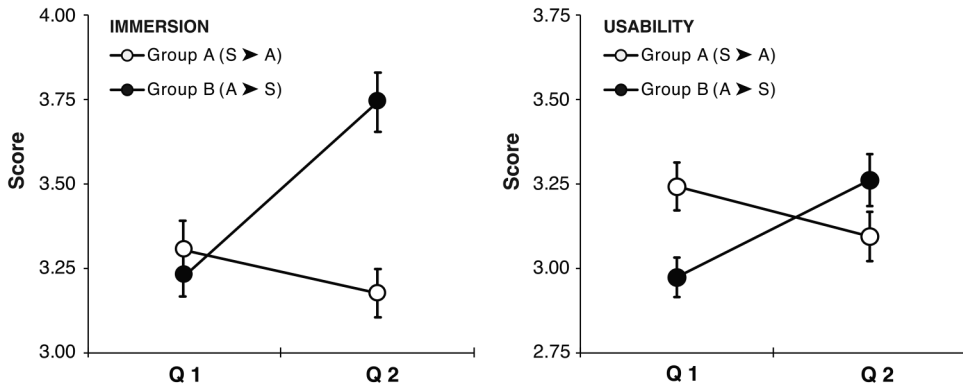


Figure 2. Scores on the immersion (left) and usability scale (right) after the two interventions. Q1 and Q2 denote the two instances in which the questionnaire was filled out. Open circles show the scores for Group A that played a scenario as the first intervention and read the article as the second intervention. Closed circles show the scores for Group B that started by reading the article and played the scenario as the second intervention. Error bars denote ± 1 SEM.

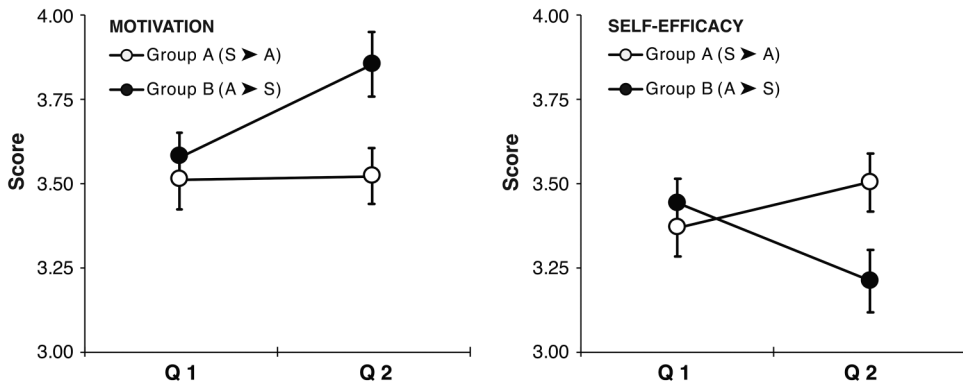


Figure 3. Scores on the motivation (left) and self-efficacy scale (right) after the two interventions. Q1 and Q2 denote the two instances in which the questionnaire was filled out. Open circles show the scores for Group A that played a scenario as the first intervention and read the article as the second intervention. Closed circles show the scores for Group B that started by reading the article and played the scenario as the second intervention. Error bars denote ± 1 SEM.

[0.57] vs 3.86 [0.80], $F = 8.32$, $p = .005$). In contrast, the scores for this group on the *self-efficacy* scale decreased significantly after playing the scenario, when the article was read before (3.43 [0.61] vs 3.21 [0.78], $F = 4.59$, $p = .036$).

Comparison to a Lecture

To compare the differential effects of playing a scenario and listening to a lecture we subsequently compared Group B (A→S) with Group C (A→L) for the second experiment (2016) only (Figure 4

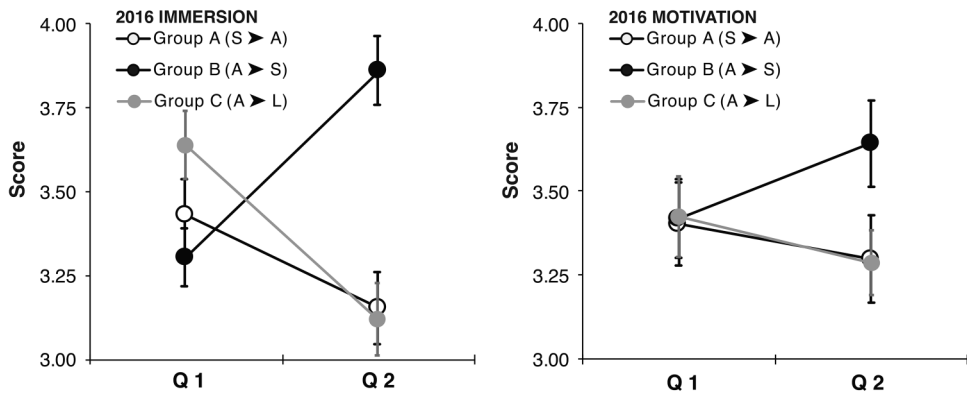


Figure 4. Scores on the immersion (left) and motivation (right) scale after the two interventions. Q1 and Q2 denote the two instances in which the questionnaire was filled out. Black circles show the scores for Group B that started by reading the article and played the scenario as the second intervention. Grey circles show the scores for Group C that started by reading the article and attended a lecture on the subject as the second intervention. For reference, open circles show the scores for Group A that played a scenario as the first intervention and read the article as the second intervention. Error bars denote ± 1 SEM.

and Supplemental Tables S5 and S6). Since we have about half the number of participants, it is not surprising that we only find significant interactions on a few scales: *Immersion*, *Usability*, and *Motivation*. Simple main effects analyses showed, not surprisingly an increase in scores on the usability scales immersion and usability for Group B (immersion: 3.31 [0.52] vs 3.86 [0.60], $F = 47.8$, $p < .001$; usability: 3.16 [0.53], 3.48 [0.60], $F = 7.21$, $p = .011$). Interestingly, the scores for Group C show a *decrease* over time (immersion: 3.64 [0.54] vs 3.12 [0.58], $F = 17.8$, $p < .001$; usability: 3.50 [0.47], 3.03 [0.57], $F = 18.09$, $p < .001$). In other words, listening to a lecture after reading the article slightly decreased the feeling of immersion and usability in this group, signifying disengagement. For the *Motivation* scale, simple main effects analyses showed that for Group B the scores increased after playing the scenario, while for Group C no such increase was apparent.

Performance Test

We analyzed whether scores on the second scenario (at the end of the session) were differentially affected by the different interventions. Since playing such a scenario may take a little practice, we were mainly interested in the possible improvement in scenario scores. Since the $2 \times 2 \times 2$ three-way mixed ANOVA (with questionnaire instance as within-subject factor, and group and year as between-subject factors) did not yield a significant interaction between year and group ($F[1,126] = 1.546$, $p = .261$), we collapsed the data for each group across years. As the scenario scores were not normally distributed (Shapiro–Wilk test, $p < .01$), we used bootstrapped analyses. The subsequent mixed ANOVA with Group (A: S → A and B: A → S) as between-subject factor and scenario instance (1 or 2) as within-subject factor revealed that students overall increased their score on the second scenario, but that the order of interventions interacted with the effect (main effect scenario instance $F[1, 128] = 22.49$, $p < .001$, $\eta_p^2 = 0.17$; interaction with Group $F[1,128] = 15.11$, $p < .001$, $\eta_p^2 = 0.11$; see Figure 5, left panel). Simple main effects analyses (Supplemental Table S7) show that students from Group A, who played a scenario first and then read an article

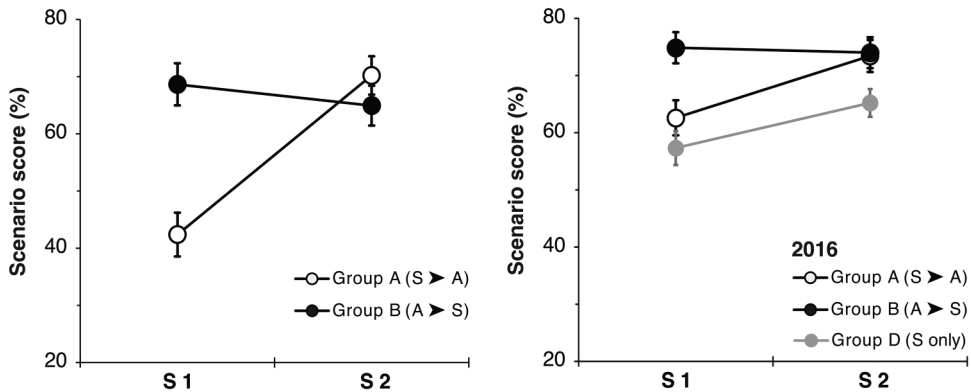


Figure 5. Performance scores on the scenarios averaged across years for Groups A and B (left), and for 2016 only to also include the performance of Group D (right). Error bars denote ± 1 SEM.

about it, significantly improved on the second scenario ($F = 32.78, p < .001$), while students from Group B, who *first* read the theoretical article and then played the first scenario did not show such an improvement ($F = 0.875, p = .353$). As can be seen in Figure 5 (left panel), the main difference here is that students in Group B who already had some theoretical knowledge (as a result of reading the article) scored better on the first scenario ($t = 4.227, p < .001$, Cohen's $d = 0.75$).

For students from Group D (2016 only), who only played two scenarios without reading the article, the score on the second scenario also improved significantly (see Figure 5, right panel and Supplemental Table S7).

Knowledge Test

For both experiments, the scores on the knowledge test adhered to the assumptions for a regular ANOVA. To compare the scores on the knowledge test, we performed a one-way ANOVA for each experiment (2015 and 2016). For the three groups in 2015, there was a significant effect ($F[2,82] = 27.82, p < .001, \eta_p^2 = 0.404$). Bonferroni-corrected post-hoc comparisons (see Supplemental Table S8) show that whereas the knowledge test performance did not differ between Group A and Group B, both groups outperform Group C (which only read the theoretical article) on the knowledge test (Figure 6; left panel). The same pattern emerges for the four groups in the 2016 experiment ($F[3,125] = 14.89, p < .001, \eta_p^2 = 0.263$). Here the difference is solely caused by a lower performance of students in Group D who only played a scenario once, sometime before the knowledge test (see Figure 6; right panel). Performance on the knowledge test did not differ between the students in the other groups.

Discussion

We investigated whether the use of an online simulation platform can aid in teaching dialogue skills in (in our case) undergraduate higher education. Rather than the more often used qualitative methods (such as focus groups), we organized one day of an existing dialogue skills course as an experimental setting with a quasi-randomized controlled trial design in two subsequent iterations of this course. In this design, we created three different types of outcome measures: experienced

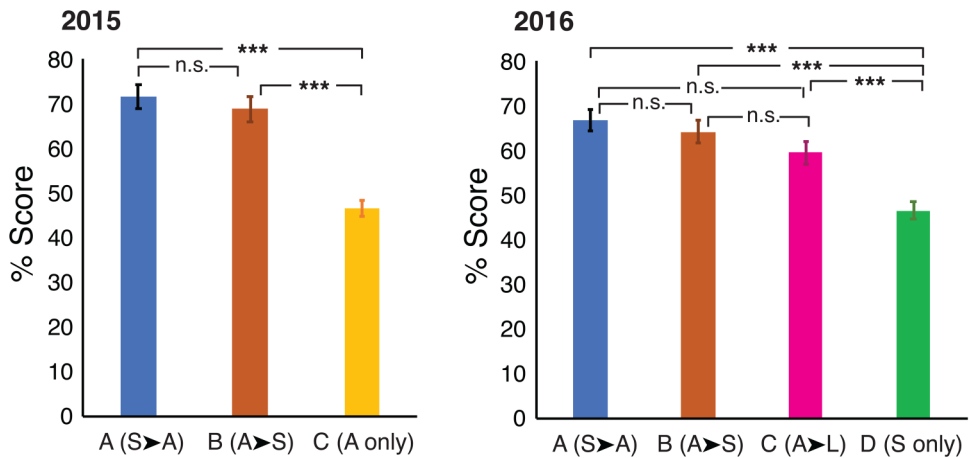


Figure 6. Scores on the knowledge test for the 2015 experiment (left) and the 2016 experiment (right). Groups are indicated on the x-axis. The y-axis indicates the average percentage of correct answers on the test. Error bars denote ± 1 SEM. Asterixes denote the level of significance (n.s. not significant, *** $p < .001$).

engagement and motivation by the students (operationalized by the questionnaires), a practical performance measure (operationalized as the score on a second scenario), and a theoretical performance measure (operationalized as the score on a knowledge MC test). On all these outcome measures, we observed an effect of using the online simulation platform, yet its precise role appears not as clear-cut.

Discussion of the Results

Playing a scenario using the platform resulted in a higher reported *Immersion* and engagement, *Usability* and *Motivation* to learn about delivering bad news in a dialogue, and a lower reported sense of *Self-Efficacy*, but only when the scenario was played *after* first reading about the theoretical underpinnings of such delivery (van der Molen et al., 1995). As is clear from Figures 2 and 3, comparing scores between initial interventions would yield no significant differences, except for the *Usability* scale (which may be perceived as a rather artificial construct for reading an article anyway). Apparently, the interplay between the two educational interventions determines the added value of the simulation in an instructional design. Would just any two educational interventions lead to such added value? Not necessarily since our results from the second (2016) experiment show that replacing playing the scenario by attending a lecture on the same subject does not lead to such added value (Figure 4). This lack of an increase in either usability scales or *Motivation* to learn about the subject is intriguing. The lecture was well perceived, with many students interacting with the lecturer before, during, and after the lecture. However, a lecture *is* mainly conveying theoretical knowledge, which was also covered in the article, and is considered a form of passive learning (e.g., Freeman et al., 2014), while a simulation is much more interactive and allows students to practice and make decisions individually, which challenges them and forces them to test their assumptions (i.e., active learning; see also Lee et al., 2020). Apparently, online simulation provides a form of authentic e-learning (e.g., Donovan et al., 1999; Herrington, 2006) that can help improve student engagement. This may also be apparent from students indicating a lower self-efficacy when

playing the scenario after reading the article. We speculate that this may reflect the scenario persuading the student of the inherent difficulty of bringing bad news and thus changing their perspective from unconsciously incompetent to consciously incompetent as a starting point for further learning, given that in our case the student already has obtained *some* theoretical knowledge by reading the article beforehand. We assume the instruction beforehand increases the student's commitment to performing, as apparently, a solution is possible and available. Opposite effects of simulations on self-efficacy when using simulations have also been found (Andrade et al., 2010), although there the first self-efficacy measure was taken *before* any intervention.

Similarly, the scores on a second scenario at the end of the experiment were higher than those on the first scenario, *except* when the first scenario was preceded by reading the article (Figure 5). When playing the second scenario, a student has the advantage of the experience of the first scenario, so when encountering situations that are similar in both scenarios (by design, since both were scenarios on delivering bad news) the student has encountered them before and has received feedback on handling those situations from the first scenario. However, the theoretical basis acquired from reading the article already improved scores in the first scenario to such an extent (compared to a novice student) that no additional improvement could be observed in the second scenario. This result also suggests that next to providing an authentic learning experience, simulation in *Communicate!*, with well-designed scenarios, may also be employed as an authentic assessment (e.g., Newmann et al., 1996) instrument.

Finally, the scores on the theoretical multiple-choice knowledge test (a traditional testing instrument) showed that time on task (e.g., Guillaume & Khachikian, 2009) probably accounted for the increase in scores on this test when in addition to reading the article a second intervention was added, be it playing the scenario or attending a lecture on the subject (Figure 6). Interestingly, only reading the article and only playing the scenario did not appear to differentially affect performance on the knowledge test. The simulations were not designed to teach students theoretical insights but to practice skills. Yet, the score increase for playing the scenario (after reading the article) was at least comparable to the increase when a lecture was attended. Moreover, although not directly compared as these manipulations were carried out for different experiments in different years, the scores on the knowledge test for the students that only read the article appear comparable to those for students that only played the scenario, which at least in part concurs with research showing "active" learning to improve study performance compared to "passive" learning (e.g., Freeman et al., 2014). This may also be one of the strengths of this platform, as it can provide written feedback during as well as after playing the scenario (Jeuring et al., 2015).

Limitations

Still, simulations on the *Communicate!* platform also clearly have limitations. Although the platform provides more visual input than a telephone call, the virtual agents are clearly animated. At the time of the experiments, characters had 14 possible facial expressions, compared to the perhaps 10,000 (Ekman, 2003) of a real human face. Also, the scenarios that were used in the experiments were purely text-based and scripted. Interaction by the student, which consisted of selecting answer options, is of course notably different from coming up with something to say. In addition, the answer options were limited to a maximum of four, as having to read more options may hinder flow (Csikszentmihalyi, 1982), which we expected would diminish immersion. Naturally, the platform is by now further developed (under the name *DialogueTrainer*) with more options and a better user experience,³ and also being investigated. For instance, new efforts to incorporate "open text input" as a more natural way to interact with the virtual character have

recently been explored (e.g., Lala, Jeurig, & van Geest, 2019). Nonetheless, there are clearly also benefits to offering predefined answer options based on theory and learning goals, which add to a very clear interaction. It would be interesting to investigate whether these enhancements would further increase the authentic experience and scores on our questionnaires' other outcome measures.

Our study also has limitations, the most important being that we cannot directly compare the effect of reading an article and playing the scenario, since technical problems prevented us from including the scenario-only condition in 2015. A considerable number of students who started with a scenario in 2015 (Group A) also scored very low on the first scenario (compare the left [aggregate] and right [2016 only] panels in Figure 5). In retrospect, we can only speculate on why this occurred specifically in this group. It was the first instance that *Communicate!* was used in this setting, and this group played the scenario in the same session as Group D, which had technical problems. Therefore, students in Group A may have been distracted by the problems in the other group and as a result, may not have taken the scenario as seriously as the others that came later. Note, however, that the pattern of results, though less extreme, is the same for this group in 2016.

In addition, we chose to incorporate the experiment in a real-life classroom setting. Students that took part in the experiments were enrolled in a course on dialogue skills and therefore possibly interested in the topic already. This also meant that we were not able to include a very large number of participants in each experiment and experimental group, diminishing our power somewhat. The effects on the performance measure (second scenario) and knowledge test are interesting but do not give any information on whether any of the interventions aid retention more than another beyond a single day, which admittedly is not a very realistic or useful time frame for retention in real-life education. However, anecdotal observations from later years in this and other programs do appear to indicate at least some retention: students still recall to break the news right away and not avoid an unavoidable confrontation. This concurs with studies described by Freeman et al. (2014), but also Lee et al. (2020) for the use of simulations in communication training especially, which show positive long-term effects of active compared to passive learning.

Further Research

Communicate! and the later *DialogueTrainer* platform can also serve as a tool for (education) research. We identify four areas of interest. The effects of communication simulations on learning and motivation, that is, *what do players learn from simulations and how do simulations impact attitudes in conversations?*, is one we started investigations on in the present article. How to improve the (technology of) scenarios and user experience in such simulations is an obvious, and highly interrelated, second area of interest (e.g., Lala, Jeurig, & van Geest, 2019; Lala et al., 2019). But one may also use the platform to investigate the models underlying several communication types and scenarios (e.g., Lala et al., 2017), or even use these types of simulations as a highly standardized and controlled environment in social psychology research, that is, to study social interactions.

Conclusion

Notwithstanding the above-mentioned shortcomings, we have demonstrated that simulations such as those provided by *Communicate!* may have added value for teaching basic dialogue skills. It is relatively cost-efficient, as scenarios can be used over multiple years, students can practice multiple times, and teacher feedback is automated (part of the scenario). The interaction takes place on a

computer with a virtual avatar and is therefore not entirely realistic, especially visually. However, the simulation challenges students to test their assumptions about conversations. As the avatar's response informs players about the validity of their assumptions, the interaction is comparable to a real conversation. In recent years, in part due to the Covid-19 pandemic more and more consultation, also in the psychological domain, is provided through online interaction, and E-Health is increasingly common (e.g., Andersson, 2009). *Communicate!* (or *DialogueTrainer*) mimics these circumstances better than a real-life consultation and therefore may also provide a valuable E-Health training platform.

Acknowledgments

We would like to thank Dr. Richta IJntema for gracefully including the presented experiments in her course on dialogue skills and for many fruitful discussions, and Sofia Barocca for her comments on an earlier draft of this manuscript.



Declaration of Conflicting Interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: The first author is currently the CEO of DialogueTrainer B.V.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Maarten J. van der Smagt  <https://orcid.org/0000-0003-2772-6429>
Michiel H. Hulsbergen  <https://orcid.org/0000-0003-0592-5312>
Jutta de Jong  <https://orcid.org/0000-0002-5066-5970>

Supplemental Material

Supplemental material for this article is available online.

Notes

1. To ensure maximum privacy and anonymity we did not ask for gender, nor age. However, the presented numbers are what is typical for this course.
2. Unfortunately, In 2015 there were technical problems that prevented group D from completing the scenario. Therefore, data from this group was not included in the analyses.
3. For example, in the current version of *DialogueTrainer* up to 60 emotions can be graphically conveyed

References

- Aldrich, C. (2004). *Simulations and the future of learning: An innovative (and perhaps revolutionary) approach to learning*. Pfeiffer Books.
- Aldrich, C. (2009). *The complete guide to simulations and serious games: How the most valuable content will be created in the age beyond Gutenberg to Google*. John Wiley & Sons.
- Andersson, G. (2009). Using the internet to provide cognitive behaviour therapy. *Behaviour Research and Therapy*, 47(3), 175–180. <http://doi.org/10.1016/j.brat.2009.01.010>
- Andrade, A. D., Bagri, A., Zaw, K., Roos, B. A., & Ruiz, J. G. (2010). Avatar-mediated training in the delivery of bad news in a virtual world. *Journal of Palliative Medicine*, 13(12), 1415–1419. <https://doi.org/10.1089/jpm.2010.0108>

- Berkhof, M., Rijssen, H. J. V., Schellart, A. J. M., Anema, J. R., & Beek, A. J. V. D. (2011). Effective training strategies for teaching communication skills to physicians: An overview of systematic reviews. *Patient Education and Counseling*, 84(2), 152–162. <https://doi.org/10.1016/j.pec.2010.06.010>
- Berkovits, I., Hanckock, G. R., & Nevitt, J. (2000). Bootstrap resampling approaches for repeated measure designs: Relative robustness to sphericity and normality violations. *Educational and Psychological Measurement*, 60(6), 877–892. <https://doi.org/10.1177/00131640021970961>
- Bosman, K., Bosse, T., & Formolo, D. (2019). Virtual agents for professional social skills training: An overview of the state-of-the-art. In P. Cortez, L. Magalhães, P. Branco, C. Portela, & T. Adão (Eds.), *Intelligent technologies for interactive entertainment. INTETAIN 2018. Lecture notes of the institute for computer sciences, social informatics and telecommunications engineering* (vol. 273, pp. 75–84). Springer. https://doi.org/10.1007/978-3-030-16447-8_8
- Brooke, J. B. (1996). SUS: A ‘quick and dirty’ usability scale. In P. W. Jordan, B. Thomas, I. L. McClelland, & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.
- Buhler, P. M., & Worden, J. D. (2013). *Up, down, and sideways: High-impact verbal communication for HR professionals*. Society for Human Resource Management.
- Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F. (2020). Simulation-based learning in higher education: A meta-analysis. *Review of Educational Research*, 90(4), 499–541. <https://doi.org/10.3102/0034654320933544>
- Cialdini, R. B. (1984). *Influence: The psychology of persuasion*. William Morrow and Company.
- Csikszentmihalyi, M. (1982). Towards a psychology of optimal experience. In L. Wheeler (Ed.) *Annual review of personality and social psychology* (vol. 3, pp. 13–36). Sage Publishing.
- Demetriou, H., & Wilson, E. (2009). Synthesising affect and cognition in teaching and learning. *Social Psychology of Education*, 12(2), 213–232. <https://doi.org/10.1007/s11218-008-9075-x>
- Deveugele, M., Derese, A., de Maesschalck, S., Willems, S., van Driel, M., & de Maeseneer, J. (2005). Teaching communication skills to medical students, a challenge in the curriculum? *Patient Education and Counseling*, 58(3), 265–270. <https://doi.org/10.1016/j.pec.2005.06.004>
- Donovan, S., Bransford, J., & Pellegrino, J. W. (1999). *How people learn: Bridging research and practice*. National Academy of Sciences.
- Ekman, P. (2003). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*. Henry Holt and Company.
- Fisher, R., Ury, W., & Patton, B. (1981/2011). *Getting to yes: Negotiating an agreement without giving in*. Random House.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410–8415. <https://doi.org/10.1073/pnas.1319030111>
- French, S., & Kennedy, G. (2017). Reassessing the value of university lectures. *Teaching in Higher Education*, 22(6), 639–654. <https://doi.org/10.1080/13562517.2016.1273213>
- Frijda, N. H. (1986). *The emotions*. Cambridge University Press.
- Guillaume, D. W., & Khachikian, C. S. (2009). The effect of time-on-task on student grades and grade expectations. *Assessment & Evaluation in Higher Education*, 37(3), 251–261. <https://doi.org/10.1080/02602930903311708>
- Hafferty, F. W. (1998). Beyond curriculum reform: Confronting medicine’s hidden curriculum. *Academic Medicine*, 73(4), 403–407. <https://doi.org/10.1097/00001888-199804000-00013>
- Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., Gentile, J., Lauffer, S., Stewart, J., Tilghman, S. M., & Wood, W. B. (2004). Scientific teaching. *Science*, 304(5670), 521–522. <https://doi.org/10.1126/science.1096022>
- Hargie, O. (2011). *Skilled interpersonal communication: Research, theory and practice*. Routledge.
- Herrington, J. (2006, October 13–17). Authentic E-learning in higher education: Design principles for authentic learning environments and tasks. World conference on E-learning in corporate, government, healthcare,

- and higher education (ELEARN), Honolulu, Hawaii, USA, pp. 3164–3173. <http://researchrepository.murdoch.edu.au>
- Hulsbergen, M. H., Dop, P. Y., Vernooij, J. C. M., & Burt, S. A. (2019). Teaching slaughter: Mapping changes in emotions in veterinary students during training in humane slaughter. *Journal for Veterinary Medical Education*, 46(1), 128–137. <https://doi.org/10.3138/jvme.0617-075r>
- JASP Team (2022). JASP 0.16 [Computer software].
- Jeuring, J. T., Grosfeld, F., Heeren, B. J., Hulsbergen, M., Ijntema, R. C., Jonker, V., Mastenbroek, N. J. J. M., van der Smagt, M. J., Wijmans, F., Wolters, M., & van Zeijts, H. (2015). Communicate!—A serious game for communication skills. In G. Conole, T. Klobucar, C. Rensing, J. Konert, & E. Lavoué (Eds.), *Design for teaching and learning in a networked world—10th European conference on technology enhanced learning, EC-TEL 2015, Toledo, Spain, September 15–18, 2015, proceedings* (pp. 513–517). Springer International Publishing.
- Konopka, C., Adaime, M., & Mosele, P. (2015). Active teaching and learning methodologies: Some considerations. *Creative Education*, 6(14), 1536–1545. <https://doi.org/10.4236/ce.2015.614154>
- Kron, F. W., Fetters, M. D., Scerbo, M. W., White, C. B., Lypson, M. L., Padilla, M. A., Gliva-McConvey, G. A., Belfore, L. A., West, T., Wallace, A. M., Guetterman, T. C., Schleicher, L. S., Kennedy, R. A., Mangrulkar, R. S., Cleary, J. F., Marsella, S. C., & Becker, D. M. (2017). Using a computer simulation for teaching communication skills: A blinded multisite mixed methods randomized controlled trial. *Patient Education and Counseling*, 100(4), 748–759. <https://doi.org/10.1016/j.pec.2016.10.024>
- Kubler-Ross, E. (1969). *On death and dying: What the dying have to teach doctors, nurses, clergy and their own families*. Scribner Book Company.
- Lala, R., Jeuring, J., van Dortmont, J., & van Geest, M. (2017). Scenarios in virtual learning environments for one-to-one communication skills training. *International Journal of Educational Technology in Higher Education*, 14, 17. <https://doi.org/10.1186/s41239-017-0054-1>
- Lala, R., Jeuring, J., & van Geest, M. (2019). Scaffolding open text input in a scripted communication skills learning environment. In A. Liapis, G. Yannakakis, M. Gentile, & M. Ninaus (Eds.), *Games and learning alliance. GALA 2019. Lecture notes in computer science* (vol. 11899, pp. 169–179). Springer. https://doi.org/10.1007/978-3-030-34350-7_17.
- Lala, R., Jeuring, J. T., Van Geest, M., Ruseti, S., Dasclu, M., Van Dortmont, J., Gutu-Robu, G., & Hulsbergen, M. H. (2019). Enhancing free-text interactions in a communication skills learning environment. In K. Lund, G. P. Nicolai, E. Lavoué, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *A wide lens: Combining embodied, enactive, extended, and embedded learning in collaborative settings* (vol. 2, pp. 877–878). International Society of the Learning Sciences.
- Lee, J., Kim, H., Kim, K. H., Jung, D., Jowsey, T., & Webster, C. S. (2020). Effective virtual patient simulators for medical communication training: A systematic review. *Medical Education*, 54(9), 786–795. <https://doi.org/10.1111/medu.14152>
- Mastenbroek, N. J., Demerouti, E., van Beukelen, P., Muijtjens, A. M., Scherpbier, A. J., & Jaarsma, A. D. (2014). Measuring potential predictors of burnout and engagement among young veterinary professionals; construction of a customised questionnaire (the Vet-DRQ). *Veterinary Record*, 174(7), 168. <https://doi.org/10.1136/vr.101761>
- Miller, R. W., & Rollnick, S. (2013). *Motivational interviewing: Helping people change* (3rd ed.). The Guildford Press.
- Newmann, F. M., Marks, H. M., & Gamoran, A. (1996). Authentic pedagogy and student performance. *American Journal of Education*, 104(4), 280–312. <https://doi.org/10.1086/444136>
- Perloff, R. M. (2008). *The dynamics of persuasion; communication and attitudes in the 21st century* (3rd ed.). Lawrence Erlbaum Associates.
- Persson, J., Dalholm, E. H., Wallergård, M., & Johansson, G. (2014). Evaluating interactive computer-based scenarios designed for learning medical technology. *Nurse Education in Practice*, 14(6), 579–585. <https://doi.org/10.1016/j.nepr.2014.05.004>

- Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33–40. <https://doi.org/10.1037/0022-0663.82.1.33>
- Pintrich, P. R., Smith, D. A. F., García, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies questionnaire (MSLQ)*. University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning.
- Silverman, J., Kurtz, S., & Draper, J. (2013). *Skills for communicating with patients* (3rd ed.) Radcliffe Publishing Limited.
- Spychala, N., Debener, S., Bongartz, E., Müller, H. H. O., Thorne, J. D., Philippen, A., & Braun, N. (2020). Exploring self-paced embodied neurofeedback for post-stroke motor rehabilitation. *Frontiers in Human Neuroscience*, 13, 461. <https://doi.org/10.3389/fnhum.2019.00461>
- van der Molen, H. T., Kluytmans, F., & Kramer, M. (1995). *Gespreksvoering: Vaardigheden en modellen (in Dutch)* (pp. 143–154). Wolters-Noordhoff.
- van der Vleuten, C., van den Eertwegh, V., & Giroldi, E. (2019). Assessment of communication skills. *Patient Education and Counseling*, 102(11), 2110–2113. <https://doi.org/10.1016/j.pec.2019.07.007>
- Windolf, P. (1981). *Berufliche sozialisation: Zur produktion des beruflichen habitus*. Ferdinand Enke Verlag.

Author Biographies

Michiel H. Hulsbergen, MSc, Psychology and Management Consultancy. Michiel Hulsbergen is a Psychologist and CEO of DialogueTrainer, which was founded in 2016 to distribute the *Communicate!* simulation platform outside of Utrecht University. Michiel contributed to the development of the *Communicate!* platform from 2013 on, at the Faculty of Social and Behavioural Sciences at Utrecht University. His research focuses on simulations, emotions, and conversation models, to advance education and organizations.

Jutta de Jong, MSc, is working as a data management consultant at the University Medical Centre (UMC) Utrecht, the Netherlands.

Maarten J. van der Smagt is a professor of Psychological Literacy in Higher Education at the Faculty of Social and Behavioural Sciences and a principal fellow at the Centre for Academic Teaching, at Utrecht University. He is the education director of Utrecht University's psychology bachelor and one-year master programs and teaches at both undergraduate and graduate (MSc and PhD) levels. His research focuses on perceptual and cognitive processes and diverse topics in higher education.