Be Open To Computer Based Coaching

Daniel Mäurer and **Karsten Weihe**
Departement of Computer Science
Darmstadt University of Technology
Hochschulstr. 10 64283 Darmstadt
{maeurer,weihe}@algo.informatik.tu-darmstadt.de

Anna Bruns
Institute of Psychology
Darmstadt University of Technology
Alexanderstr. 10 64283 Darmstadt
bruns@psych.tu-darmstadt.de

Abstract

Conversational agents are used for various purposes, such as marketing, e-learning and tutoring. However, they have not been used for personal coaching so far. Personal coaching can be used as a strategy to support professionals in transferring their newly acquired skills to everyday work after receiving a training. The aim of our research was to examine the usefulness of computer based coaching as a training transfer strategy. We also examined whether the user's openness is a key factor for the effectiveness of a virtual coach. We present a computer based coaching system specifically designed for training transfer. In a longitudinal experiment we tested our system against an online journal with regard to its effectiveness. We found some evidence that participants with high openness benefit more from computer based coaching than participants with lower openness, while openness has no influence on the effectiveness of the online journal. Our results suggest that computer based coaching can be effectively used as a training transfer strategy, but may not be equally effective for everyone.

1 Introduction

Conversational agents are used for marketing, e-learning and tutoring purposes, just to mention a few. Whereas users remain skeptical towards "talking" to a computer, there are still many unexplored application areas for conversational agents. Personal coaching, for example, is a promising field as the coaching maxim rather involves the guidance of individuals by stimulating questions than by specific instructions or correct answers. So far, conversational agents have not been extensively used for personal coaching purposes.

We present a computer based coaching system that is particularly designed to enhance training transfer. Training transfer, the application of newly acquired skills in everyday work, requires a maximum of support and reflection. This support and reflection can be provided by a personal coach (Olivero et al., 1997). However, personal coaching is very expensive and therefore only available to a limited circle of individuals in organizations. Computer based coaching, on the other hand, could support a much larger audience to successfully transfer their newly acquired skills from training to everyday work.

Our Contribution We have developed a conversational agent (a.k.a. dialog system) to support our computer based communication training. The dialog system was accessible for the participants during weekdays to support them with the application of acquired communication skills after a weekend training phase. This kind of computer based coaching is intended to support trainees to reflect on their process of goal accomplishment by targeted questions. Thereby, it does not understand the full content of the trainees' reactions to this questions in detail. Instead, the system classifies trainees' reactions as dialog acts (Stolcke et al., 2000) to keep track of the relevant information to successfully direct and control the dialog process.

Of course, our goal is not to replace real human coaches in general; just like Weizenbaum (1966) was sure that ELIZA is not a way to replace psychotherapists. Our vision is to introduce computer based coaching in situations where a personal human coach is simply not affordable or available.

Ever since, user acceptance has been a serious problem for conversational agents. In a professional setting such as training transfer, user acceptance is a key success factor. As the personality

Proceedings of the 17th Workshop on the Semantics and Pragmatics of Dialogue, December 16-18, 2013, Amsterdam, The Netherlands.

trait openness of the coachees is known to affect user acceptance (Devaraj et al., 2008), we expect that it can also have a huge impact on the transfer success as well. In line with this, we also expect that our results will stimulate a new direction of research on computer based coaching and the influence of personality traits.

Testing Effectiveness In order to test the effectiveness of our computer based coaching system, we present a longitudinal field experimental study that compares two training transfer support strategies after an online communication training: computer based coaching and online journals. This experiment is expected to reveal first results on the effectiveness of computer based coaching as training transfer strategy. Furthermore, we want to explore how openness and user acceptance interact with the effectiveness of computer based coaching. Our results can provide a basis for future developments of conversational agents in the field of coaching.

Related Work Many other work has been done in the field of conversational agents, but only few focused on coaching. The results closest to ours are found in the healthcare domain, where (Bickmore et al., 2005) present an agent with the role of a physical exercise advisor. Although their dialog manager is working in a similar way to the one we present, user contributions to the dialog are made primarily by selecting items from multiple-choice menus, whereas our system allows the user to answer in natural language at every time. While multiple-choice menus are a sufficient way to enforce the users motivation for physical exercising, we expect communication in natural language to be necessary for cognitive tasks such as training transfer. (Bickmore, 2003) focused on studying the effects of social conversation and the relation between the agent and the user in a artificial situation. He also found that the user personality and their trust in the agent were intercorrelated. In contrast, our study focuses on the outcome on the domain goal in a real world setting, namely the increase in their communication skills. SimCoach, a dialog based healthcare assistant, (Rizzo et al., 2011) focused on promoting access to domain specific information. (Conati et al., 2000) introduced a chat based tutor in an educational setting. This tutor aims to foster learning from examples and to provide feedback on self-testing examples.

This paper is structured as follows. Section 2 will introduce coaching and the application of coaching as a training transfer strategy. In Section 3, we will introduce our computer based coaching agent. Section 4 will describe our experimental study on the effectiveness of our system. We will finish with the results of the study in Section 5 and an outlook on our future work.

2 Coaching

2.1 Solution Based Brief Coaching

Like many other disciplines, coaching has struggled with developing a common definition. For the scenario of a short training transfer dialog, we picked a goal-focused approach called *brief coaching* (Berg and Szabo, 2005). The primary method of solution based *brief coaching* is to support the coachee in defining goals and a suitable goal accomplishment strategy. A coaching session in *brief coaching* encompasses three stage phases:

- **Desired Future:** Defining a specific goal.
- **Changes in State:** Discuss recent past, look for indicators of changes in direction of desired state.
- **Experimental Phase:** Put into practice what has been discussed so far, agree on minor changes in everyday activity.

A general maxim of many coaching approaches is to regard the client as the expert for the relevant problem rather than seeing the coach as the expert for the client's problem. Therefore, coaching does not intend to give advice or push the client into any certain direction, but rather to ask targeted questions that help the 'expert' to get a new perspective and to develop his own solution. An ideal coach would do this by stating questions *only*.

2.2 Coaching as a Training Transfer Strategy

Training is successful, if training transfer was successful (Barnett and Ceci, 2002). Training transfer is defined as "...the degree to which trainees effectively apply the knowledge, skills, and attitudes gained in a training context to the job. For transfer to have occurred, learned behavior must be generalized to the job context and maintained over a period of time on the job." (Baldwin and Ford, 1988, p. 63).

However, at work, costs of failure and the pressure to meet deadlines typically hinder employees

from exploring new and alternative methods. Instead of further improving their skills, they tend to rely on existing and well-practiced methods (Ericsson et al., 1993; Haccoun, 1997). Transfer interventions are effectively used to increase the motivation of learners to use their newly acquired skills in their daily routine. In particular, literature strongly supports the use of the goal setting strategy (Burke and Hutchins, 2007). Comparative studies reveal that goal setting is superior to other post-training interventions in terms of increasing trainees' transfer performance (Wexley and Baldwin, 1986). As solution-based brief coaching (Berg and Szabo, 2005) is a particular form of goal setting, we suggest coaching to be a suitable alternative to conventional post training interventions (e.g. goal setting via online journals). Decisive superiority of coaching in comparison to other transfer strategies may lie in its ability to enhance participants' metacognitions (Grant, 2003). Metacognitions capture the planning how to best achieve a specific goal, monitoring the progress and the evaluation of the used strategies (Schraw and Moshman, 1995). First research attempts were able to show that managers who received personal coaching after a training intervention further increased their productivity during the coaching phase (Olivero et al., 1997). However, personal coaching would be far too expensive in order to provide it to a larger number of employees.

Our computer based coaching , on the contrary, could be an effective and economic alternative.

Openness We expect personality to affect the success of computer based coaching. "Openness to new experience" is one of the "Big Five" personality factors. It encompasses intellectual curiosity, preference for variety and the willingness to explore new ways (Costa and McCrae, 1992). Especially for complex and changing task conditions, openness has shown to significantly impact the effectiveness of training interventions (Herold et al., 2002).

Furthermore, openness has a significant positive influence on coaching success (Stewart et al., 2008) and certain components of user acceptance (Devaraj et al., 2008). In contrast to conventional transfer strategies (e.g. online journals), we expect user acceptance to be a key factor for the success of computer based coaching. Therefore, we argue that individuals with higher openness will benefit more from computer based coaching than individuals with low openness.

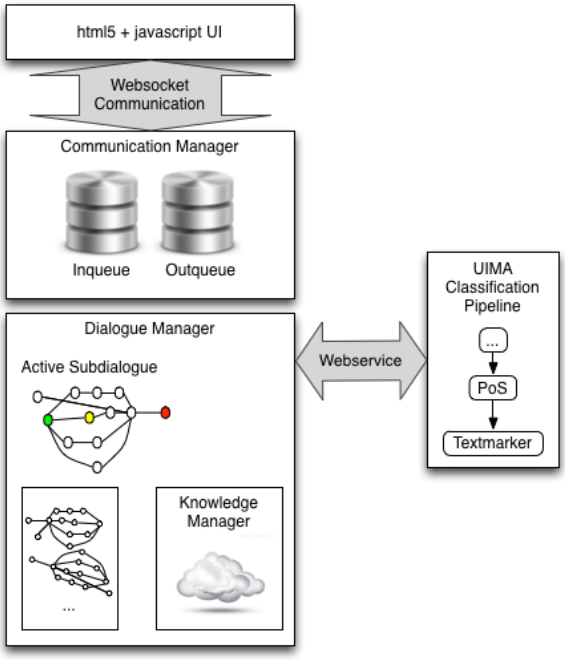

Figure 1: System Architecture

3 A chat based coaching system

In this section we will describe some of the details of our computer based coaching system. It is a mixed initiative system, which means that both the user and our system can take initiative turns to start a subdialog. A turn is a single utterance either by the user or the system. Although our system is technically capable of handling user initiatives as well, the coach is the one who usually takes the initiative in an ideal solution-based brief coaching process. Our conversational agent works with a high level dialog structure and local subdialogs. This high level structure follows the three stages model of solution-based brief coaching.

Additionally, we provide a set of transition rules that determine which subdialog is selected depending on the user's progress, for example:
if $worked\_on\_goals\_today?$
then $subdialog$: "$percentage?$"
else $subdialog$: "$what\_went\_wrong?$"

Our system can be divided into several components (Figure 1), namely *dialog manager*, *subdialog system*, *text classification* and *communication manager*. We will shortly describe them in the following:

Dialog Manager The main function of the dialog manager is to control the flow of the dia-

log. This means that it keeps track of the context knowledge, state of the conversation and all knowledge gathered from further sessions. Every single coaching question defined in the high level conversation structure is representing one subdialog. The system responses are constructed from a large set of pre-formulated sentences or sentence snippets. This also provides several alternative formulations for each possible system response.

After finishing a subdialog, a specific content analyzation component is extracting relevant changes in knowledge and forwards them to the knowledge manager. The analyzation component also decides whether the dialog manager should publish some feedback responses to the user.

Subdialog System The purpose of a subdialog is to extract a unique piece of information from the user. Each subdialog is based on a finite state machine. Transitions between the different states are triggered by user or system dialog acts. The users utterances are classified by a rule based classifier as described in the next paragraph. Each subdialog models a coaching question with all turns required to answer the question sufficiently. In an ideal case, this subdialog can be closed in two turns, for example a *y/n-question* can be simply responded with an *agreement* dialog act. A more complicated example would be the following which was taken from a real dialog (See appendix).
The coach is marked 'C', the human user with 'H'. In parentheses: The dialog act of the utterance.

...
C: *Ok. How would you rate yesterdays progress on your weekly goal? Did you fully reach your goal or only to some extent?* (question)
H: *Yesterday, one my supervisor talked to me about a mistake I made.* (statement)
C: *Go on.* (pull)
H: *I remained calm and analysed the situation. Based on my analysis I recognized that my mistake can be corrected.* (statement)
C: *And how would you rate your goal accomplishment in percent? (0 - 100)?* (question)
H: *I have reached my goal by about 80%.* (statement)
C: *Congratulations, well done.* (acknowledge)
...

In this subdialog the system asks for the amount of progress. With his first response, the user does not answer this question properly and responds with a loosely related statement instead. The system cannot handle the information correctly, but this is not obvious to the user. Instead, after another turn the system simply tries to restate the initial question and asks for the percentage of progress. When reaching the final state and analyzing the result, the system ends the subdialog with a feedback response.

Our system also provides fallback mechanisms in case of mismatches or unrecognized turns made by the user. An example for a fallback is a handler for counter questions, which suspends the current subdialog and resumes it afterwards. Another handler is implemented for problem recognition, for example too many turns in one subdialog or too many unexpected user responses.

Text Classification: We decided to develop a rule based dialog act classifier. Other approaches, mostly for the English language, are machine learning or statistical approaches as presented in various publications (Stolcke et al., 2000; Marineau et al., 2000; Reithinger and Klesen, 1997). Machine learning was not an option because of non available suitable corpus for chat conversation in German language.

We implemented a UIMA[1] based classification pipeline using tokenization, lemmatization and a part-of-speech tagger for German language. Our set of classification rules at the end of the pipeline was implemented with TextMarker (Kluegl et al., 2009), now known as *UIMA ruta*.

Mixed Initiative Multi-turn Management: In order to improve the acceptance of the system we developed a communication manager protocol capable of multi-turn interactions. The protocol was specifically designed to simulate human chat behavior, for example that the user has the possibility to state more than one submission to the system:

...
C: *How are you today?*
H: *mhm...*
(5 sec pause)
H: *not that good.*
C: *Why is that?*
...

Similar to the user, our system is able to state more than one single utterance within a short amount

[1] UIMA stands for Unstructured Information Management Architecture and is a component architecture and software framework implementation for the analysis of unstructured content like text. See http://uima.apache.org for details.

of time. Therefore, our communication manager component works with two queue data structures. One queue is for the user utterances published since the system's last dialog act, called *incoming queue*. The other queue stores system utterances which the dialog manager has planned to publish to the user, but has not published yet. This one is called *outgoing queue*.

Once an utterance from the user is received, the dialog manager compiles a response and pushes it into the outgoing queue. These responses are prepared to be published under the given context. Depending on the question stated by the system and the new response to be posted, the system will determine a waiting time depending on the responses length and question initially stated and wait for that amount of time. While waiting, the user will be informed via the *typing* signal that the system is planning to say something. However, the user is still able to make another turn or add something to his/her last statement. If the user had something more to say, the system clears its outgoing queue with all the utterances the system planned to publish. The Dialog Manager recompiles its response based on the new state of information. Once the system is sure that the current response is valid under the context of all given answers by the user, it will publish its response and clears the ingoing queue with all the user's messages it had to respond to. With this protocol we improve naturalness and assure consistency in all system responses.

User Interface: Our chatcoach is presented via a web-based html 5 frontend. It was a very important requirement that our system is easily accessible. Our web frontend even enables support for mobile devices. The communication between the browser based client frontend and our server-based agent is handled by a publish/subscribe messaging framework. This gives the system the possibility to also display *typing* status messages which are very important for handling the message publication timing in a multiple user turn scenario. Every typing action of the user or of the system is immediately present to the other party. This enables a more lifelike user experience during the conversation.

Example Dialog The following dialog shows an example session of a workday translated from German.

C: *Hello. The last time we talked to each other you told me you couldn't work much on your goal. Did you come any closer in the meantime?*
H: *Yes, always whenever mistake occurs, I try to stay calm and analyse the situation.*
C: *Ok. How would you rate yesterdays progress on your weekly goal? Did you fully reach your goal or only to some extent?*
H: *Yesterday, one my supervisor talked to me about a mistake I made.*
C: *Go on.*
H: *I remained calm and analysed the situation. Based on my analysis I recognized that my mistake can be corrected.*
C: *And how would you rate your goal accomplishment in percent? (0 - 100)?*
H: *I have reached my goal by about 80%.*
C: *Congratulations, well done.*
C: *Why do you think you haven't accomplished more than that?*
H: *At first, it is always a shock and pretty difficult to stay calm. But when I take the time and analyse the situation, its getting better.*
C: *Is there anything that went particularly good, when you think about working on your goal yesterday?*
H: *Yes, that i remained calm although I was stressed.*
C: *What exactly was positive about that situation?*
H: *To stay calm and not get hectic*
C: *Ok good. Now that we have reflected what went well and not so well on your last workday is there something that you would like to change tomorrow?*
H: *No, I think its okay.*
H: *I'm starting to get a routine in such situations.*
C: *And what could help you reach your weekly goal completely?*
H: *I guess thats not really possible in the near future.*
C: *Are you satisfied with your actual goal accomplishment?*
H: *I think I am.*
C: *Sounds good.*
C: *And how much would you say does the work on your goal affect your communication skills? (strong to not at all)*
H: *I would say strong*
C: *Alright. I think that's it for today. Goodbye!*
H: *Bye*

	Total M (SD)	Journal M (SD)	cbc M (SD)	1	2	3	4	5	6
Transfer Strategy	-0.08 (1.02)	1.00 (0.00)	-1.00 (0.00)	-					
Communication Skills (T1)	19.86 (4.83)	19.94 (5.07)	19.84 (4.81)	-.01	-				
Communication Skills (T2)	20.68 (6.09)	20.28 (5.42)	21.02 (6.79)	.06	.40*	-			
Communication Skills (T3)	22.57 (4.32)	22.43 (3.34)	22.68 (5.14)	.03	.44*	.49*	-		
Openness	2.54 (1.07)	2.42 (1.38)	2.64 (0.74)	.11	-.07	.10	.03	–	
Motivation	2.20 (0.51)	2.26 (0.60)	2.15 (0.43)	-.10	.33	.31	.22	.18	-
Login Frequency	4.23 (4.62)	2.50 (1.73)	5.71 (5.77)	.35^T	-.17	-.08	-.28	.19	-.06

Table 1: Means, Standard Deviations and Pearson Correlations. Note: *Computer Based Coaching vs. Online Journal* was contrast coded: Computer Based Coaching (cbc) = 1; Online Journal (journal) = -1. $^T p < .10, ^* p < .05$.

4 Experimental Study

Our study was conducted as a longitudinal field experiment with three measurement points. Participants were 26 alumni of TU Darmstadt (50% female) who voluntarily signed up for an online communication training program. On average, participants were 36 years old (SD = 10.89) and completed one to five communication trainings prior to this training.

Our study was conducted in two phases: A pretest on participants' Communication Skills (t1) was followed by a training phase that ended with a posttest on Communication Skills (t2). Immediately after the posttest, participants were randomly assigned to two different experimental conditions: Twelve participants were instructed to record their progress in an online journal on a daily basis over the course of one week. The remaining fourteen participants were instructed to use the computer based coaching also on a daily basis over the course of one week. The effects of the different transfer strategies were assessed in a follow-up test on participants' Communication Skills (t3) after the end of the transfer week.

Instructions after the posttest (t2) were identical in both experimental conditions (Computer Based Coaching and Online Journal) in the following aspects: On the first day (after the posttest at t2) participants were asked to define a specific goal they want to accomplish in the transfer phase. This goal should refer the improvement of communication skills. Positive examples for goals were provided. Participants were also asked to rate the feasibility, and to name potential promoters and inhibitors to accomplish their goal.

On the following four days participants were instructed to rate goal accomplishment and the (positive or negative) effect of this goal accomplishment on their communication skills. Furthermore we asked participants to name the specific promoters and inhibitors of goal accomplishment they faced during that day.

Instructions after the posttest (t2) were different for both experimental conditions (Computer Based Coaching and Online Journal) in the following aspect only: Whereas the Online Journal presented the instructions in a static form, the Computer Based Coaching presented the questions adaptively in the form of a dialog as described in Section 3.

4.1 Measures

Communication skills were assessed in a test at all three measurement points: in a pretest prior to the training phase (t1), in a posttest after the training phase and prior to the experimental manipulation of the Transfer Strategies (t2) and in a follow-up after the experimental manipulation of the Transfer Strategies (t3). The test consisted of three critical situations that were presented to the participants (i.e. 9 critical situations in total). Within 15 minutes, participants had to generate as many useful and original responses to the given situa-

tions as possible. Two independent experts rated the quality of the different responses on two dimensions (usefulness and originality) on an anchored 7-point Likert scale. Multiple responses of an individual participant were averaged per situation and dimension. A single Communication Skills Index was formed by multiplying scores on these two dimensions (Zhou and Oldham, 2001). A global Communication Skills Index per measurement point was aggregated across the three test situations. This elaborate procedure resulted in a good agreement between the ratings of the two independent experts (ICC .70 to .84).

Openness was measured by two items derived from (Rammstedt and John, 2007) Big Five Inventory-10. Participants rated themselves on both items ('I see myself as someone who has as few artistic interests.' and 'I see myself as someone who has an active imagination.') on a five-point Likert scale (1= disagree strongly to 5 = agree strongly). Both items were later aggregated to a global Openness score.

4.2 Control Variables

In our analysis, we wanted to see the "pure" effect of our Transfer Strategy(Computer Based Coaching and Online Journal) without the distortion of other influential factors. Therefore we controlled for several variables in our analysis that we expected to also have an influence on Communication skills at t3 apart from our Transfer Strategy:

First, we expected our participants to differ in their Communications Skills prior to the training (at t1) and prior to the experimental manipulation of the Transfer Strategy (t2). In order to eliminate in our analysis both the influence of prior Communications Skills and the effects of the training itself, we included Communication Skills at t1 and t2 as control variables into our analysis.

Second, we provided participants with access to the training chapters also after they had completed the posttest at t2. As further repetition of the training chapters may also cause a further improvement of participants' Communication Skills, we recorded the Login Frequency after the posttest and controlled for its influence in our analysis.

Third, we expected the participants' individual motivation to have an influence on the effectiveness of the training and possibly interfere with the effects of the different Transfer Strategies. Therefore, we assessed participants' initial motivation to sign up for our communication training on 15 items (e.g. "My main driver to participate in the training is because I want to improve my social skills") that covered five dimensions of motivators from technical aspects to career advancement. Participants rated their motivation on a five-point Likert scale (1= disagree strongly to 5 = agree strongly). All items were later aggregated to a global motivation score.

5 Results

5.1 Descriptives

Our analysis encompassed one dependent variable (Communication Skills at t3), two independent variables (Openness and Transfer Strategy) and four Control Variables (Communication Skills at t1, Communication Skills at t2, Login Frequency and Motivation). Means, standard deviations, and intercorrelations among all variables are presented in Table 1. Communication Skills are significantly correlated ($p < .05$) across the three measurement points. Furthermore, the Transfer Strategy and Login Frequency were significantly correlated ($p < .10$). As Computer Based Coaching was contrast coded with +1 and the Online Journal with -1 this positive correlation indicates that participants in the *Computer Based Coaching* condition had more logins after the posttest at t2 than participants in the Online Journal condition. All other variables did not differ significantly between the two experimental conditions.

5.2 Hypothesis Testing

We assumed that participants who are more open to new experience will benefit more from computer based coaching than participants who are less open to new experience.

We tested our assumption using hierarchical regression analysis. In the first step, we entered the Control variables (Communication Skills at t1, Communications Skills at t2, Login Frequency and Motivation). In the second step, we entered the moderator variable (Openness) and Transfer Strategy (*Computer Based Coaching vs. Online Journal*). The interaction term between Transfer Strategy and Openness was entered in the third step (Aiken and West, 1991).

To reduce multicollinearity, all variables were centred at their respective means.

Table 2 reports the test of our assumption: The Control variables entered in step 1 of the hierar-

	Step 1	Step 2	Step 3
Change in variance accounted for Communication Skills at t3 (ΔR^2)	.36*	.01	.14*
Communication Skills (t1)	0.26	0.26	0.12
Communication Skills (t2)	0.37^T	0.36^T	0.50*
Login Frequency	-0.21	-0.25	-0.33^T
Motivation	0.01	0.01	0.28
Transfer Strat.	—	0.09	0.08
Openness	—	0.05	0.26
Transfer Strat. x Openness	—	—	0.55*

Table 2: Test of Moderation Transfer Strategy x Openness on Communication Skills. Note: Values are standardized regression coefficients. All terms were centred prior to analysis.
^{T} : $p < .10$ *:$p < .05$.

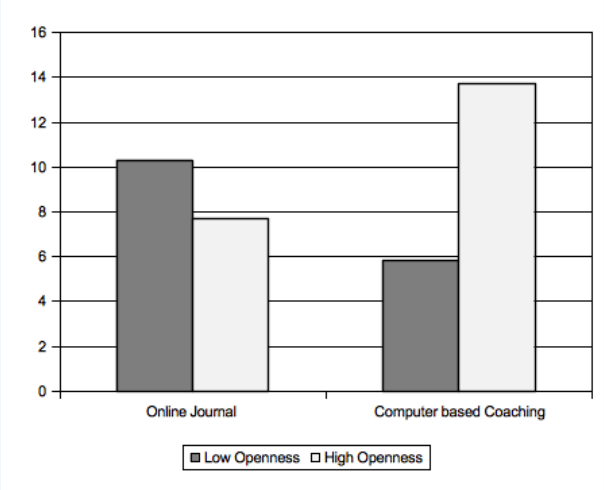

Figure 2: Gain in Communication Skills at t3 depending on Openness and Transfer Strategy after controlling for Communication Skills at t1, Communication Skills at t2, Login Frequency and Motivation.

chical regression analysis already accounted for 36% of the variance in the Communications Skills at t3. In particular, Communications Skills at t2 significantly predict Communications Skills at t3 ($\beta = .37, p < .10$). When we entered the Transfer Strategy (Online Journal vs. computer based coaching) and Openness in step 2 no additional variance in the Communications Skills at t3 was explained. However, when we entered the interaction between Transfer Strategy and Openness in Step 3 additional 14% of the variance in the Communications Skills at t3 could be explained. In sum, a total of 50% of the variance in the Communications Skills at t3 can be explained by using this set of variables. Among the variables three significant predictors were identified: The interaction between Transfer Strategy and Openness significantly predicted Communication Skills at t3 together with the Control variables Communication Skills at t2 ($\beta = .50, p < .05$) and Login Frequency ($\beta = -.33, p < .10$).

The plot of the relationship between Transfer Strategy and Openness is presented in Figure 2 and supported our Hypothesis: Participants who are more open to new experience benefit more from the Computer Based Coaching-condition than participants who are less open to new experience. The simple slope analysis revealed this difference to be significant ($p < .05$). In the Online Journal-condition, the effect seemed to be reversed. However, the simple slope analysis revealed this difference not to be significant ($ns.$).

5.3 Quality Evaluation:

Classification error rates did not vary significantly between the high and the low openness group. Fatal classification errors, such as mistaking a *disagreement* for an *agreement*, were not observed during our study. One of the shortcomings of our system was its deficient handling of counter questions. We counted four dialogs where the user aborted the conversation. In three of those conversations, the situations that caused the dialog to fail were initiated by user questions or false-positive questions.

6 Conclusion and future work

Our results suggest that computer based coaching effectively helps the participants to further increase their communication skills after a training intervention. According to (Shawar and Atwell, 2007), the best method to evaluate a conversational agent is to measure whether it achieves the service or task it was intended to. In this respect, our system performed quite well.

However, the participants' success largely differ with regard to their openness: Participants with high levels of openness benefit more from computer based coaching than participants with low levels of openness. In contrast, openness for experience does not seem to influence the effectiveness of online journals. This implies that computer based coaching is probably not suitable for everyone. Therefore, future work on similar research questions should take into account the influence of personality and background. It may be advisable to consider the users personality and background in order to avoid biased results in similar studies.

Of course, the dialog system will be further improved and is planned to be used for other application scenarios, for example decision coaching. Our future work includes building a German chat

corpus with dialog act tags. We are planning to use it for further evaluation and improvements of the dialog act classifier.

References

[Aiken and West1991] Leona S Aiken and Stephen G West. 1991. *Multiple regression: Testing and interpreting interactions*. Sage.

[Baldwin and Ford1988] Timothy T Baldwin and J Kevin Ford. 1988. Transfer of training: A review and directions for future research. *Personnel psychology*, 41(1):63–105.

[Barnett and Ceci2002] Susan M Barnett and Stephen J Ceci. 2002. When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological bulletin*, 128(4):612.

[Berg and Szabo2005] Insoo Kim Berg and Peter Szabo. 2005. *Brief coaching for lasting solutions*. WW Norton New York.

[Bickmore et al.2005] Timothy Bickmore, Amanda Gruber, and Rosalind Picard. 2005. Establishing the computer-patient working alliance in automated health behavior change interventions. *Patient Education and Counseling*, 59(1):21 – 30.

[Bickmore2003] Timothy Wallace Bickmore. 2003. *Relational agents: Effecting change through human-computer relationships*. Ph.D. thesis, Massachusetts Institute of Technology.

[Burke and Hutchins2007] Lisa A Burke and Holly M Hutchins. 2007. Training transfer: An integrative literature review. *Human resource development review*, 6(3):263–296.

[Conati et al.2000] Cristina Conati, Kurt Vanlehn, et al. 2000. Toward computer-based support of meta-cognitive skills: A computational framework to coach self-explanation. *International Journal of Artificial Intelligence in Education (IJAIED)*, 11:389–415.

[Costa and McCrae1992] Paul T Costa and Robert R McCrae. 1992. Four ways five factors are basic. *Personality and individual differences*, 13(6):653–665.

[Devaraj et al.2008] Sarv Devaraj, Robert F Easley, and J Michael Crant. 2008. Research notehow does personality matter? relating the five-factor model to technology acceptance and use. *Information Systems Research*, 19(1):93–105.

[Ericsson et al.1993] K Anders Ericsson, Ralf T Krampe, and Clemens Tesch-Römer. 1993. The role of deliberate practice in the acquisition of expert performance. *Psychological review*, 100(3):363.

[Grant2003] Anthony M Grant. 2003. The impact of life coaching on goal attainment, metacognition and mental health. *Social Behavior and Personality: an international journal*, 31(3):253–263.

[Haccoun1997] Robert R Haccoun. 1997. Transfer and retention: Let's do both and avoid dilemmas. *Applied Psychology*, 46(4):340–344.

[Herold et al.2002] David M Herold, Walter Davis, Donald B Fedor, and Charles K Parsons. 2002. Dispositional influences on transfer of learning in multistage training programs. *Personnel Psychology*, 55(4):851–869.

[Kluegl et al.2009] Peter Kluegl, Martin Atzmueller, and Frank Puppe. 2009. Textmarker: A tool for rule-based information extraction. In Christian Chiarcos, Richard Eckart de Castilho, and Manfred Stede, editors, *Proceedings of the Biennial GSCL Conference 2009, 2nd UIMA@GSCL Workshop*, pages 233–240. Gunter Narr Verlag.

[Marineau et al.2000] J. Marineau, P. Wiemer-Hastings, D. Harter, B. Olde, P. Chipman, A. Karnavat, V. Pomeroy, A. Graesser, and the TRG. 2000. Classification of speech acts in tutorial dialog. In *Proceedings of the workshop on modeling human teaching tactics and strategies at the Intelligent Tutoring Systems 2000 conference*, pages 65–71.

[Olivero et al.1997] Gerald Olivero, K Denise Bane, and Richard E Kopelman. 1997. Executive coaching as a transfer of training tool: Effects on productivity in a public agency. *Public personnel management*, 26:461–470.

[Rammstedt and John2007] Beatrice Rammstedt and Oliver P John. 2007. Measuring personality in one minute or less: A 10-item short version of the big five inventory in english and german. *Journal of Research in Personality*, 41(1):203–212.

[Reithinger and Klesen1997] Norbert Reithinger and Martin Klesen. 1997. Dialogue act classification using language models. In *EuroSpeech*. Citeseer.

[Rizzo et al.2011] Albert Rizzo, K Sagae, E Forbell, J Kim, B Lange, JG Buckwalter, Josh Williams, TD Parsons, P Kenny, David Traum, et al. 2011. Simcoach: an intelligent virtual human system for providing healthcare information and support. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*, volume 2011. NTSA.

[Schraw and Moshman1995] Gregory Schraw and David Moshman. 1995. Metacognitive theories. *Educational psychology review*, 7(4):351–371.

[Shawar and Atwell2007] Bayan Abu Shawar and Eric Atwell. 2007. Different measurements metrics to evaluate a chatbot system. In *Proceedings of the Workshop on Bridging the Gap: Academic and Industrial Research in Dialog Technologies*, pages 89–96. Association for Computational Linguistics.

[Stewart et al.2008] Lorna J Stewart, Stephen Palmer, Helen Wilkin, and Marie Kerrin. 2008. The influence of character: Does personality impact coaching success. *International journal of evidence based coaching and mentoring*, 6(1):32–42.

[Stolcke et al.2000] Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol Van Ess-Dykema, and Marie Meteer. 2000. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational linguistics*, 26(3):339–373.

[Weizenbaum1966] Joseph Weizenbaum. 1966. Eliza - a computer program for the study of natural language communication between man and machine. *Commun. ACM*, 9(1):36–45, January.

[Wexley and Baldwin1986] Kenneth N Wexley and Timothy T Baldwin. 1986. Posttraining strategies for facilitating positive transfer: An empirical exploration. *Academy of Management Journal*, 29(3):503–520.

[Zhou and Oldham2001] Jing Zhou and Greg R Oldham. 2001. Enhancing creative performance: Effects of expected developmental assessment strategies and creative personality. *The Journal of Creative Behavior*, 35(3):151–167.