Building the Dream Team: Children's Reactions to Virtual Agents that Model Collaborative Talk

Joseph B. Wiggins
University of Florida
Gainesville, FL, USA
jbwiggi3@ufl.edu

Toni V. Earle-Randell
University of Florida
Gainesville, FL, USA
tearlerandell@ufl.edu

Dolly Bounajim
North Carolina State University
Raleigh, NC, USA
dbbounaj@ncsu.edu

Yingbo Ma
University of Florida

Julianna Martinez Ruiz
University of Florida

Ruohan Liu
University of Florida

Mehmet Celepkolu
University of Florida

Maya Israel
University of Florida

Eric Wiebe
North Carolina State University

Collin F. Lynch
North Carolina State University

Kristy Elizabeth Boyer
University of Florida

ABSTRACT

Intelligent virtual agents have tremendous potential for facilitating collaborative learning by modeling and reinforcing desirable collaborative practices. Despite recent work in this area, the extent to which intelligent virtual agents can facilitate improvements in the collaborative behavior of children is largely unknown. This study employed a wizard-of-oz study design and investigated elementary children's collaborative behavior after interacting with virtual agents. These agents model exploratory talk for upper elementary school dyads, such as asking higher-order questions and listening to their partners. The findings uncover associations between elementary learner dyads' positive changes in collaboration after agent interventions, the dyads' affective reactions to interventions, and their attentiveness to the agents. Our results also reveal associations between positive changes in collaboration and the timing of interventions: for example, earlier interventions had a higher occurrence of positive changes, and positive changes in collaboration typically happened within five seconds of interventions. The results suggest ways in which intelligent virtual agents may be used to promote effective collaborative learning practices for children.

CCS CONCEPTS

• **Social and professional topics → K-12 education**.

KEYWORDS

Pedagogical Agents, Collaborative Dialogue, Elementary School

IVA '22, September 6–9, 2022, Faro, Portugal

ACM ISBN 978-1-4503-9248-8/22/09…$15.00
https://doi.org/10.1145/3514197.3549683

ACM Reference Format:
Joseph B. Wiggins, Toni V. Earle-Randell, Dolly Bounajim, Yingbo Ma, Julianna Martinez Ruiz, Ruohan Liu, Mehmet Celepkolu, Maya Israel, Eric Wiebe, Collin F. Lynch, and Kristy Elizabeth Boyer. 2022. Building the Dream Team: Children's Reactions to Virtual Agents that Model Collaborative Talk. In *ACM International Conference on Intelligent Virtual Agents (IVA '22), September 6–9, 2022, Faro, Portugal.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3514197.3549683

1 INTRODUCTION

Collaborative learning is increasingly central in education, partly in response to collaborative skills becoming more crucial in the workplace [3]. Supporting collaborative discourse and collaborative learning is becoming a more prevalent application of intelligent virtual agents [33]. Collaborative learning involves two or more learners working interactively on a shared learning goal through information sharing and negotiation [6, 10, 23]. As a form of collaborative learning, *pair programming* is an effective pedagogical strategy for teaching programming skills in K-12 CS education [30]. However, simply pairing learners and assigning designated roles does not guarantee good learning outcomes [4, 15]. Learners benefit from being taught how to conduct collaborative dialogue [28].

In elementary grades (kindergarten through fifth grade in the United States, encompassing ages 5-11), recent studies have showcased how intelligent virtual agents can successfully promote communication skills [7] and individual learning in many disciplines, from algebra [16] and AI education [34] to foreign language learning [1, 11, 14]. However, not enough is known about partnering pairs of intelligent virtual agents with elementary learner dyads. Systematic surveys of the proceedings of the Intelligent Virtual Agents conference point to the need for exploring "human collaboration with one or more agents" [18]. More specifically, there is a need for understanding children's responses to the agents and how the children change behaviors after interacting with the agents.

The work presented in this paper addresses that need: we investigated how virtual agents facilitate improvement in the collaborative behaviors of upper elementary learner dyads. We conducted a Wizard-of-Oz (WoZ) study, where human wizards controlled a pair of virtual agents designed to promote strong collaborative behaviors. We investigated the interaction process between learner dyads and the agents through the following research questions: RQ1) What are upper elementary learner dyads' learning and attitude outcomes after working with intelligent virtual agents?; RQ2) To what extent do learners uptake positive collaborative behavior changes after virtual agents model those behaviors?; RQ3) What is the relationship between learners' uptake of positive collaborative behavior and their reaction to the virtual agents?; and RQ4) Does learners' uptake of positive collaborative behavior change as they interact with the agents over time?

To investigate these research questions, we conducted a Wizard-of-Oz study with two virtual agents in the form of learning companions. In this study, researchers observed the children's collaboration and triggered a set of agent dialogues to scaffold the children's collaborative process. We analyzed each dyad's affective state prior to, during, and after the agent interventions. Our findings reveal that improvements in collaborative behavior following the agents' interventions were associated with a dyad's affective state prior to the intervention (negative vs. non-negative) and attentiveness during the intervention; that the timing of the intervention was associated with positive changes in collaboration; and when improvements in collaboration happened, they occurred within five seconds of the interventions. By considering the factors that are associated with improvements in elementary school dyads' collaborative behaviors, intelligent virtual agents can be designed to provide feedback that improves young learner dyads' collaborative skills.

2 BACKGROUND

Intelligent virtual agents researchers have deeply explored agents' social and affective positioning and behaviors, but systematic reviews have pointed out that "multiple agents and user studies about perception or collaboration in multi-agent environments should be introduced and researched further" [18]. Agents that support the collaboration of human users typically aim to facilitate and promote conversation and curiosity [19]. Multi-agent systems, such a system which has multiple agents engage a user about their stress management [13], tend to emphasize non-verbal interactions to simulate group cohesion. They evaluate postures, gaze, and facial expression, attempting to understand which cues are the most important to the user [13]. Multi-agent systems can also present short interactions between multiple virtual agents to be perceived as more socially present and interesting [5]. Multi-agent systems have also demonstrated that they are more persuasive than single agents, and are more likely to cause behavior change [12]. While these result are promising, there have been very few virtual agents that were designed for promoting learners' collaborative skills.

Children in the elementary grades often lack the skills and experience necessary for collaborative learning, and without explicit support, they may encounter many barriers. When learning collaboratively, dyads must communicate incomplete and evolving thoughts without the confidence of concrete understanding. This type of talk is referred to as "exploratory talk": a learner thinking aloud by hesitantly voicing partly-formed ideas [17]. Exploratory talk is educationally productive, as it provides a means of forging understanding [2]. This form of communication is often not an overt effort in collaboration, but rather a learner attempting to hear how their own idea sounds. By engaging in exploratory talk, a learner makes themselves vulnerable to others' criticism, and doing so requires a degree of comfort with their peers [2]. To foster communities of inquiry in a classroom setting, teachers must facilitate discussion among dyads to encourage exploratory talk. This is a challenging task when many dyads are collaborating simultaneously. Instead of placing sole responsibility for this facilitation on teachers, it is critical to develop agents that are capable of promoting good exploratory talk behaviors.

The exploratory talk framework has shown promise in the context of pair programming practices that are used to introduce some elementary school learners to computer science. Rojas-Drummond and Zapata [22] utilized an exploratory talk framework while teaching fifth grade learners, engaging them in critical reasoning and promoting constructive knowledge building on peers' ideas. In a study examining upper elementary learners, Vandenberg et al. [29] found that during collaborative paired programming, learners given prompts that encouraged collaborative and exploratory talk differed both quantitatively and qualitatively in their interactions compared to the control group. They found that learners in the intervention group asked each other questions, presented alternative ideas, and justified their thoughts more frequently than learners in the control group [29]. While prompts were physically administered during their study, their work helps inform adaptive systems that can offer timely support to learners while also mitigating the burden on teachers. Some studies implemented Peer Assisted Learning Strategies (PALS), where two learners with different levels of achievements work in pairs and take turns assuming the roles of tutor and tutee [8]. These peer-assisted programs produced positive learning outcomes, especially for young learners [20].

However, facilitating effective collaboration in multiple groups of young learners in the same classroom is a challenging task. There is a high demand for teacher's time and resources, and facilitating collaboration takes time and attention that are difficult to achieve in everyday learning contexts [26]. Rojas-Drummond and Mercer [21] found it is rare for teachers to instruct learners in effective collaborative talk. To address these limitations of traditional educational settings, there is potential for intelligent agents to support elementary learners in their natural collaborative learning environments. Nonetheless, there is a limited amount of work that investigates the mechanics of how young learners interact with intelligent virtual agents, especially in the domain of collaboration. This is a missed opportunity because pedagogical agents have shown increased efficacy on younger learners compared to older ones [24], and it is well known that collaboration skills must be reinforced in younger learners to encourage their development [27, 35]. Recognizing the benefits of collaboration and the challenges of effectively enforcing these skills in the classroom, we aimed to examine how virtual agents can improve collaborative efforts and learner outcomes.

Inspired by these findings, we created interventions that encourage collaborative and exploratory talk during pair programming. We designed a pair of virtual agents to work alongside upper elementary school learners as they learned to program. The agents both modeled good collaboration, and prompted dyads to use good collaborative practices.

3 METHODS

This section describes the participants and study design, gives an overview of the learning environment and virtual agents, and presents the Wizard-of-Oz setup. This study was conducted as part of a larger project to investigate collaborative CS learning with virtual agents for upper elementary school children (fourth and fifth grade, generally ages 9-11). The study was conducted in an elementary school over six days in the Fall of 2021, and each study session lasted 50 minutes. Out of 35 children, 30 provided assent and parental consent for data collection, and of these children, 18 were female and 12 were male. The mean age was 10.6, with ages ranging from 9 to 11. In addition, 37% of learners reported having some prior coding experience at the beginning of the study.

On the first day of the study, learners began by individually completing a pre-assessment consisting of a 10-question coding knowledge quiz [29] and an 11-question computer science attitude survey, where learners stated their level of agreement with statements like "I would like to use coding to make something new" and "I believe I can be successful in coding" [28]. Following the pre-assessment, learners were guided through simple coding puzzles to familiarize them with some of the programming concepts that they would be using in the upcoming days, such as variables, conditionals, and loops[1]. On day two, learners explored the FLECKS environment built upon Netsblox[2]. The instrutor provided a demonstration on how the blocks worked, and were encouraged to start with an empty project and build sample code alongside the instructor with basic blocks such as motions and loops.

On days three, four, and five, the learners collaborated on the coding activities shown in Table 1, following the pair programming paradigm. The instructor described the roles of driver and navigator to the learners: the driver controlled the mouse and keyboard, and the navigator watched and helped the driver code through conversation. After twenty minutes, learners switched roles. Throughout the coding activity on days three, four, and five, the expert wizards observed learner behavior via Zoom and delivered relevant interventions to the learners through the virtual agents. Each day, when learners had completed the coding activity, they were asked to complete a short survey evaluating the activity and their partner satisfaction, along with a five-question coding quiz. On the sixth and final day of the study, learners completed a post-assessment identical to the CS knowledge and attitude pre-assessment.

3.1 Learning Environment for Upper Elementary Collaborative Coding

The learning environment shown in Figure 1 was the basis of this study, which extended the Netsblox system to include two virtual agents designed to foster collaborative learning by modeling good collaboration and intervening when the learners were not displaying ideal practices. We provided a set of coding activities for the learners to work on together in this block-based learning environment.

The virtual agents were designed through a rigorous co-design process and positioned as near-peers to the upper elementary learners [31]. Our aim was to create agents that chime in with their own experiences and feel as though they could be the users' friends or classmates. The virtual agents are portrayed as being in the fourth grade and are framed as novice computer science learners. The agents make suggestions for improving collaboration, specifically the collaborative dialogue framework of exploratory talk [17]. The final dialogues were voice acted and played aloud while learners interacted with the system, and were also captioned on the screen.

[1]https://blockly.games/
[2]https://netsblox.org/

Table 1: Coding activities and surveys administered throughout the study.

Day	Activities
1	Pre-Assessment and Attitude Survey, Blocky Games
2	FLECKS environment interface overview, individual coding
3	WoZ study day 1, "The current code doesn't work, change it so that Alonzo moves in a square"
4	WoZ study day 2, "Change the code so that Kitty, Alonzo and Ladybug move around forever"
5	WoZ study day 3, "Change the code so that Alonzo, Kitty and Ladybug talk when they run into each other"
6	Post-Assessment and Attitude Survey

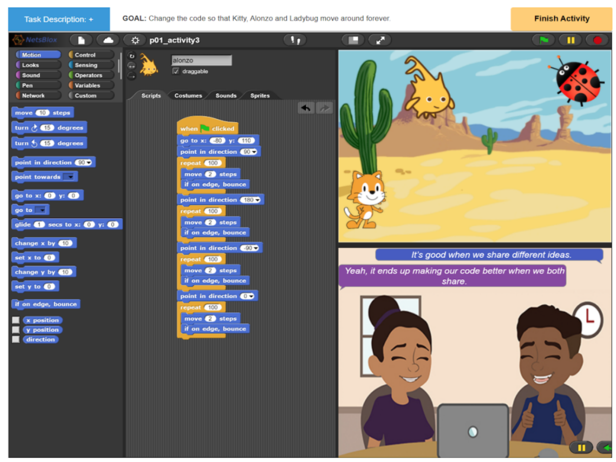

Figure 1: The interface of the FLECKS environment used in classrooms.

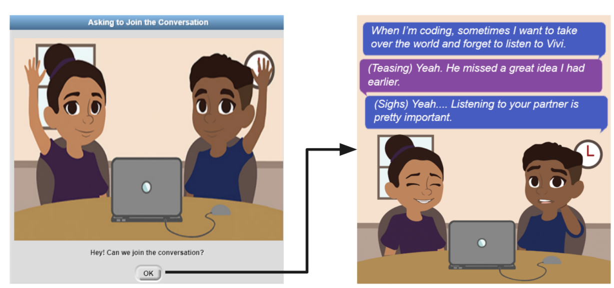

Figure 2: Each intervention begins with a pop-up asking "Hey! Can we join the conversation?" (Right) and once accepted, a vignette plays with animations and audio (Left).

3.2 Wizard of Oz Protocol

During the study, the virtual agents were controlled by human wizards in a WoZ design. The wizards observed both the learner dyad's screen and a webcam recording the learners' interactions. Guided by the protocol in Table 2, the wizards directed the virtual agents embedded in the FLECKS environment to present dyads with one of 12 possible vignettes. When the wizard initiated a vignette, learners would be presented with a pop-up alert (Figure 2), which the interface required them to accept, and then the virtual agents would play out the vignette associated with the scenario the wizard had selected. For example, a wizard recognized that Learners A and B were encountering a difficulty in collaboration. On the webcam feed, Learner A, the driver, was attentive and engaged while Learner B was disengaged and distracted. Learner A said "Learner B, can you actually help?", but Learner B said "I don't know what to do." From the learners' screen, the wizard was able to see that there was a disconnect between the learners' attentiveness to the activity, and that Learner A was working by themselves. The wizard identified this scenario as *not sharing ideas* and initiated the associated vignette. The wizards involved in this study were trained by iteratively watching a set of videos of elementary learner dyads learning computer science and voting on what wizarded moves should be taken. Disagreements between wizards were discussed within the entire group of wizards, and rules were refined.

Table 2: Wizards' intervention triggers and details.

Intervention Triggers	Details
Confusion about objective	After kids enter the activity, they look confused. This could happen even after reading the "GOAL". "I am lost, what are we doing?"
Floundering	Kids are making repetitive changes without advancing their solution.
Complete Impasse (Close to end of session)	Kids try to collaborate but run out of ideas. They reach a breaking point, and may ask for adult (facilitator, or teacher) help.
Inactive Coding	Not making any changes on the code for 2 minutes. If they are talking, it's unrelated to the task.
Inactive Talking	Kids are not talking at all for 2 minutes.
Distraction	Kids are disengaged due to an external distraction (cellphone, YouTube, etc).
Asking "Why" Questions	Kids are asking high order questions. ("Why?" "How?"). Play this to praise them.
Sharing Ideas	Kids are sharing ideas with each other, taking turns speaking. They wait for their turn to share. Play this to praise them.
Listening to Each Other	Kids (Driver) implement an idea ("Let's try that"), or acknowledge it. ("Yeah, but that won't work."). Play this to praise them.
Not Asking "Why" Questions	Kids are not asking why they chose this or that block. Play this if you want to encourage asking higher order questions ("Why?" "How?").
Not Sharing Ideas	Kids are not taking turns sharing ideas with each other. Play to encourage them to share ideas and consider their partner's.
Not Listening to Each Other	Offered ideas are not implemented or acknowledged. Play to encourage idea implementation and listening.

4 DATA

Out of the 15 consenting dyads, only 9 dyads were present on each of the days and had video and audio recordings of their collaboration over the course of the study. These video and audio recordings were manually transcribed and prepared for tagging. This produced a dataset of dyad dialogue, virtual agent interventions, and learner surveys. We identified a total of 107 agent interventions, with a mean of 11.4 interventions per dyad (SD = 4.80).

To answer our primary research question of "To what extent do learners uptake positive collaborative behavior changes after virtual agents model those behaviors?", it was necessary to annotate the dataset. This analysis focused on dyads' responses to the virtual agents' interventions. Segments of video, screen recordings, and dialogue transcripts were extracted for analysis using the following process: each segment began at the moment learners were presented with the intervention pop-up alert, and ended 30 seconds after the intervention vignette finished playing. Our team reviewed the transcripts, video, and screen recordings for each wizard intervention during three different windows: (1) prior to accepting the intervention (as shown in Figure 2 Left), (2) during the intervention (as the vignette played as shown in Figure 2 Right), and (3) the 30 seconds following the intervention. We created a coding schema, revised that schema as we examined the data, and operationalized definitions, as shown in Table 3 and visualized in Figure 3. Non-Negative Affective Reactions and Negative Affective Reactions were annotated from the moment the agents suggested talking with the learner dyad until the learner dyad accepted (and thus, started) the intervention. Two annotators who were familiar with the study tagged an overlapping 20% of the data and were able to achieve near perfect agreement (kappa = 0.91). The annotators then proceeded to divide the remaining data and tag it independently.

Table 3: Description of learner dyads' immediate reactions to the agent interventions.

Immediate Response	Description	Frequency
Non-negative Affective Reaction	Accepting the agents' request to intervene neutrally or positively, "Maybe they can help"	83.7%
Negative Affective Reaction	Loud sighs, rolling eyes, "Noooooo!", "no, you can't join the conversation!"	16.3%

After accepting an intervention, a short vignette plays based on the option the wizard selected. During the interventions, annotators labeled the group behavior during the vignette as Attentive or Non-Attentive, as shown in Table 4 and visualized in Figure 3. Our annotators were able to achieve substantial agreement (kappa = 0.74). Finally, in the 30 seconds after an intervention vignette had played, we identified three codes for changes in collaborative behaviors, shown in Table 5: Positive Change, Neutral/No Change, and Negative Change. Our annotators were able to achieve substantial agreement on these labels (kappa = 0.69).

Positive Change in Collaborative Behaviors. Our agents were designed to facilitate positive change in upper elementary

Table 4: Description of learner dyads' attentiveness during the agent interventions.

During Intervention	Description	Frequency
Attentive	Looking at the agents, paused conversation, "thank you"	73.1%
Non-Attentive	Kept coding, starting off-topic conversations, scrolled away, insulting agents	26.9%

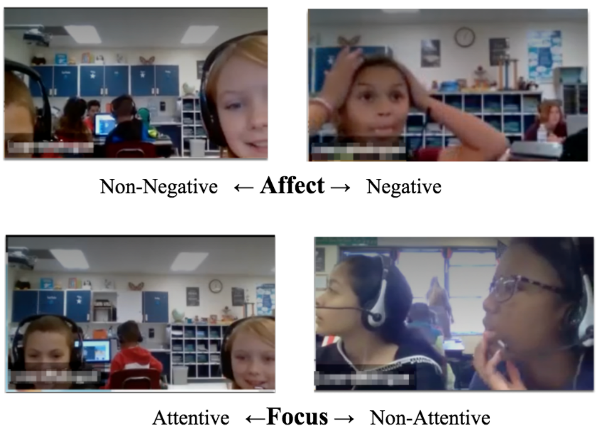

Figure 3: (Top) Partners having a non-negative affective reaction compared to partners having negative one; (Bottom) Partners being attentive while agents are giving an intervention compared to partners being non-attentive.

Table 5: Dyad behaviors within 30 seconds after the agent interventions.

30 Second Reaction	Description	Frequency
Positive Change	Uptake, improvement in collaboration (i.e., asking higher-order questions, sharing ideas, listening)	21.2%
Neutral	No noticeable change in collaboration	75.0%
Negative Change	Distracted, frustrated, agent focused off-task discussion (i.e., "They're listening to us!")	3.8%

dyads' collaborative behaviors. When tagging, annotators focused on three positive behaviors: asking higher order questions, sharing ideas, and listening to partner input. An example of uptake can be seen in Table 6. In the uptake of collaborative behaviors, the upper elementary learner dyads sometimes directly mimicked or paraphrased the dialogue proposed by the virtual agents, but always promoted more conversation.

Neutral or No Change in Collaborative Behaviors Neutral reactions were the most frequent reaction to the interventions, and were tagged when the elementary learner dyads showed no change in their collaborative behaviors after the intervention. Common reactions learners had in this scenario were light acknowledgement

Table 6: Excerpt of dyad's dialogue which exemplifies the integration of exploratory talk after intervention.

Speaker	Utterance
Agent 1:	Teamwork is all about communication. If you two shared your ideas with each other, maybe it'd get the ball rolling!
Learner 1:	(To partner) What do you think we should do? Because I am trying to get him [sprite] to move.

(i.e., "Ok"), complete ignore (i.e., learners immediately refocused on their code), and least commonly, dismissive reactions (i.e., light sigh, quick joke, or light mocking). Dismissive behaviors were not considered negative changes in behavior in this analysis since they did not affect how the partners were treating each other or approaching the coding activity.

Negative Change in Collaborative Behaviors Finally, the least common outcome was a negative change in behavior. As shown in the excerpt in Table 7, these were scenarios in which the agent intervention began a derailment of the conversation and reduced learners' focus on their coding. This scenario is very undesirable, with the agents doing more harm than good. A peril of virtual agents is that they risk distracting learners more than aiding them [9, 24]. Designing virtual agents that strike the balance of being engaging without becoming an unhelpful distraction is a challenging task. This phenomenon may help explain why some dyads became frustrated with the agent, as they felt it was hindering their progress more than helping it. However, this was an unusual situation in our data.

Table 7: Excerpt of dyad's dialogue which exemplifies the derailment of their conversation after the agent intervention.

Speaker	Utterance
Agent 1:	Our teacher Ms. Diaz said that it's very important to ask "why" questions to each other.
Agent 2:	Yes. So we ask questions like: "Why do you think that?"
Learner 1:	Why do you think that [learner 2 name]? Why do you think they're doing that? Hmm, hmm?

5 RESULTS

First we address RQ1, "What are upper elementary learner dyads' learning and attitude outcomes after working with intelligent virtual agents?" While this exploratory study did not have a control group, we can still compare learners' CS content knowledge and CS attitudes from the beginning to the end of the study using paired-sample t-tests. Learners achieved significant improvements in CS Attitudes and CS Content Knowledge over the course of the six day study. On average, learners began the study with a CS Attitude score of 42/55 (SD = 6.46) and completed the study with 46.3/55 (SD = 6.26), paired samples test $t(29) = -3.438$, $p = .002$; $d = 0.64$. Their CS Content knowledge pre-test were, on average, 3.89/10 (SD = 1.69) and increased to 5.39/10 (SD = 1.87) on the post-test, paired samples test $t(28) = -5.116$, $p < .001$; $d = 0.97$. The learners also had a fairly positive view of the agents in their post-surveys, stating that "I learned how to share more ideas with my partner" and "they helped me and my partner ask why questions".

5.1 Uptake of Positive Collaborative Behaviors

Next, we present the changes that were observed after the virtual agents' intervention to answer RQ2: "To what extent do learners uptake positive collaborative behavior changes after virtual agents model those behaviors?" In our data, we saw learner dyads integrate positive collaborative behaviors following 21.2% of the interventions. This was much more frequent than negative changes following an intervention (3.8%), but still far less frequent than the neutral reaction (75.0%). In this dataset, there were no instances of uptake of positive collaborative behaviors with delays more than five seconds after the intervention, despite the taggers observing 30-second windows after the interventions.

Next, we investigate RQ3: "What is the relationship between learners' uptake of positive collaborative behavior and their reaction to the virtual agents? More specifically, we examine RQ3a: "Is learner uptake of positive collaborative behavior related to the learners' affective responses *immediately following* the virtual agents' modeling of those behaviors?" and RQ3b: "Is learner uptake of positive collaborative behavior related to the learners' attentiveness *during* the virtual agents' modeling of those behaviors?" To investigate these research questions, we computed the conditional probability of learner dyads' transitions between the three phases of intervention (Figure 4): immediate affective reaction to the alert, attentiveness during the vignette, and change in collaborative behavior in the subsequent 30 seconds.

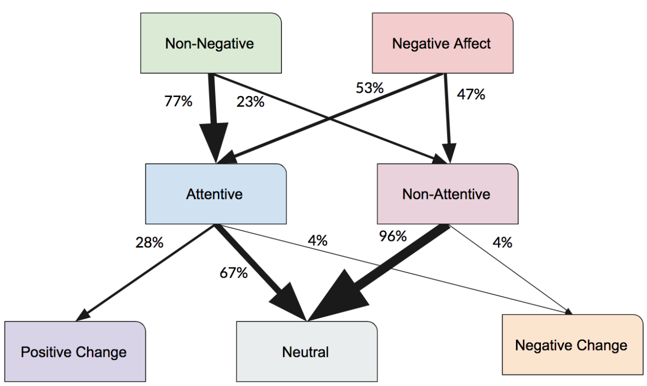

Figure 4: The probability of immediate reactions (Top) and the conditional probability of the behavior changes (Bottom).

When an intervention was sent to a dyad, there was an 83.7% chance of them accepting the agents' intervention without any negative affective reaction, and a 16.3% chance of a negative affective reaction such as loud sighing, eye rolling, or some negative utterance. When partners had a negative affective reaction to receiving the message, they were twice as likely to be Non-Attentive while the intervention was playing (47%) when compared to non-negative reactions (23%). However, it was more likely in both non-negative (77%) and negative (53%) affective reaction conditions for partners to be Attentive instead of Non-Attentive. During the 30-second window after an intervention, the most likely outcome was neutral or no change for both those partners who were Attentive (67%) and those who were Non-Attentive (96%). Both Attentive and Non-Attentive partners had an equal likelihood of a negative change in their behaviors (4% in both scenarios), but only partners who were Attentive were seen positively changing their collaborative behaviors (28%). Although not shown in this graphic, the likelihood of a positive change occurring when partners had reacted non-negatively to receiving the message was four times greater than when partners responded negatively (P(PositiveChange|Non-negativeAffect) = 24%; P(PositiveChange|NegativeAffect) = 6%).

5.2 Timing of Interventions and Uptake of Positive Collaborative Behaviors

In exit interviews with the wizards that controlled the virtual agents in this work, wizards remarked that the timing of the intervention seemed to be related to how accepting dyads were of the agents' suggestions. The wizards observed that dyads appeared more accepting of agents' suggestions earlier in lessons when the learners were not yet focused on coding. Similarly, the learners seemed more accepting of the agents on the earlier days of the study compared to the later days. There is some evidence, though not definitive, that the novelty of agents may impact outcomes after working with them [25]. To examine this phenomenon in our context, we investigate RQ4: "Does learners' uptake of positive collaborative behavior change as they interact with the agents over time?" More specifically, RQ4a: "Is there a difference in likelihood of uptake of positive collaborative behavior between the first half and second half of collaborative learning sessions?" and RQ4b: "Is there a difference in likelihood of uptake of positive collaborative behavior between the days of the study?"

To answer RQ4a, we divided each session into the first half of the learning session (i.e., the first 25 minutes) and the second half of the learning session (i.e., the last 25 minutes) and computed the frequencies of each tagged state for each half, as shown in Figure 5. While the immediate reactions that elementary learner dyads had to agent interventions were almost identical between the first (84% Non-negative Affective Reactions) and second halves of the session (82% Non-negative Affective Reactions), we see some larger differences between learners' attentiveness during interventions and the change in their collaborative behaviors. During interventions given during the first half of the sessions, learners were Attentive 78% of the time compared to the second half's 63%. The interventions that were given during the first half of the session were also more likely to lead to positive changes in collaborative behavior (25%) than those in the second half (14%). Negative behavior changes remained close to the same, with the first half having a 4% likelihood, and the second half 3%.

To answer RQ4b, we contrasted intervention reactions, attentiveness, and change in collaborative behavior across Days 3-5 (Figure 5, Bottom). The learner dyads had the least negative affective reactions to the agents on Day 3 at 10%, which increased to 24% on Day 4, and slightly decreased to 17% on Day 5. Similarly, dyads' attentiveness during the intervention was the highest on Day 3 at 83% and dropped to 66% on both Days 4 and 5. The changes in collaborative behavior followed a similar trend of positive behavior changes being the most likely on Day 3 at 26%, and then dropping to 18% on Day 4 and 17% on Day 5. Negative behavior changes also fell from Day 3 onward (7% → 3% → 0%) while the No change in behavior steadily increased (67% → 79% → 83%).

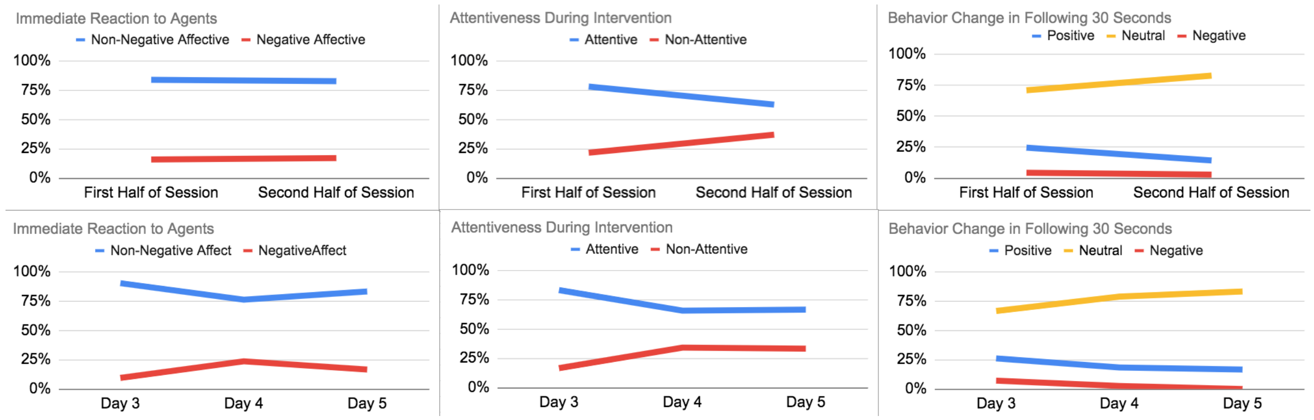

Figure 5: Differences in reactions and the conditional probability of the behavior changes in interventions given in first and second half of a session (Top) and interventions given in Days 3-5 (Bottom).

6 DISCUSSION AND IMPLICATIONS FOR IVAS

This study investigated elementary learner dyads' collaborative behavior changes and reactions to interventions from two virtual agents. Our analysis revealed that positive changes in collaborative behavior (asking higher-order questions, sharing ideas, and listening to partner input) following agent interventions only happened when dyads showed signs of actively listening to the agents during their interventions. Furthermore, positive changes in collaborative behavior were four times less likely when partners had a negative affective reaction to receiving the message. These relationships may be partly explained by the importance of the dyad's cognitive and affective states [32]. Furthermore, when positive changes in collaborative behavior occurred, they happened almost immediately after the intervention ended, with the majority happening in a 1-2 second time window and no instances being observed after five seconds. Similarly, our analysis demonstrates that certain periods during collaboration are more likely to result in the uptake of positive collaborative behaviors, such as the first day children are meeting the agents and early in the class. This finding suggests that these early periods are important, and decisions to intervene may need to be made quickly and in moments that learners are accepting of help in order to change behaviors.

This work has important implications for how educational intelligent virtual agents should be designed to better support collaborative learning and productive collaborative practices. Our findings highlight the importance of sending interventions at periods when dyads are in a receptive state (i.e., when they are more likely to have a non-negative affective reaction, and will be attentive while the intervention is playing) to facilitate positive changes in a dyad's collaboration. Given that elementary school dyads showed uptake of the primed exploratory talk behaviors within a five-second window, this small window could be used by intelligent virtual agents to quickly assess if the intervention was effective and if the intelligent virtual agent needs to try a different approach.

7 CONCLUSION AND FUTURE WORK

Intelligent virtual agents that support collaborative learning need to carefully navigate a multi-party dialogue, offering support only when it is likely to be helpful. The work reported here has suggested ways in which agents can be designed to improve the likelihood of learner uptake of desirable behaviors. We have reported on a Wizard-of-Oz study in which wizards controlled two virtual agents working with dyads of elementary school learners. The agents attempted to promote positive collaborative dialogue practices between partners as they were coding together. The findings revealed that improvements in collaborative behavior after the agents' interventions were: 1) only seen when elementary learner dyads were attentive, 2) far less likely if the elementary learner dyads had a negative affective response when the agents requested to join the conversation, 3) more frequent in agent interventions that happened in the first half of session than the second half of the session, 4) more likely on the first day than the second and third, and finally, 5) only seen within five seconds of the agents' intervention, if an improvement in collaborative behavior did occur.

The results provide many insights for the design of intelligent virtual agents and directions for future work. This work identified whether the agents' request to join the conversation caused a negative reaction, and whether the dyads' attentiveness during the intervention was associated with whether the dyads made improvements to their collaborative behaviors. The ability to identify these states could be integrated into a fully automated intelligent virtual agent, and would be greatly facilitated by the use of multimodal streams such as facial expressions, prosody, and proximity. Once these states can be recognized, intelligent virtual agents can approximate their interventions' likelihood of promoting improvements in the collaboration of learner dyads, and change strategies.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation through grant DRL-1721160. Any opinions, findings, conclusions, or recommendations expressed in this report are those of the authors, and do not necessarily represent the official views, opinions, or policy of the National Science Foundation.

REFERENCES

[1] Doris Luft Baker, Hao Ma, Paul Polanco, Jillian Marie Conry, Akihito Kamata, Stephanie Al Otaiba, Wayne Ward, and Ron Cole. 2021. Development and promise of a vocabulary intelligent tutoring system for Second-Grade Latinx English learners. *Journal of Research on Technology in Education* 53, 2 (2021), 223–247.

[2] Douglas Barnes. 2008. Exploratory talk for learning. *Exploring talk in school* (2008), 1–15.

[3] Marilyn Binkley, Ola Erstad, Joan Herman, Senta Raizen, Martin Ripley, May Miller-Ricci, and Mike Rumble. 2012. Defining twenty-first century skills. In *Assessment and teaching of 21st century skills*. Springer, 17–66.

[4] Stephen Brown and David Baume. 2022. 'Not another group activity!': Student attitudes to individual and collaborative learning activities, and some implications for distance learning course design and operation. *Innovations in Education and Teaching International* (2022), 1–10.

[5] Salam Daher, Kangsoo Kim, Myungho Lee, Ryan Schubert, Gerd Bruder, Jeremy Bailenson, and Greg Welch. 2017. Effects of social priming on social presence with intelligent virtual agents. In *International Conference on Intelligent Virtual Agents*. Springer, 87–100.

[6] Pierre Dillenbourg. 1999. What do you mean by collaborative learning?

[7] Maha Elgarf, Gabriel Skantze, and Christopher Peters. 2021. Once Upon a Story: Can a Creative Storyteller Robot Stimulate Creativity in Children?. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*. 60–67.

[8] Lynn S Fuchs, Douglas Fuchs, Sarah Kazdan, Kathy Karns, Mary Beth Calhoon, Carol L Hamlett, and Sally Hewlett. 2000. Effects of workgroup structure and size on student productivity during collaborative work on complex tasks. *The Elementary School Journal* 100, 3 (2000), 183–212.

[9] Ruth Garner, Rachel Brown, Sylvia Sanders, and Deborah J Menke. 1992. Seductive details and learning from text. *The role of interest in learning and development* (1992), 239–254.

[10] J Gerlach. 1994. Is this collaboration: New Directions for Teaching and Learning.

[11] Sivaldo Joaquim, Ig Ibert Bittencourt, Rafael de Amorim Silva, Patrícia Leone Espinheira, and Marcelo Reis. 2022. What to do and what to avoid on the use of gamified intelligent tutor system for low-income students. *Education and Information Technologies* 27, 2 (2022), 2677–2694.

[12] Reshmashree B Kantharaju, Dominic De Franco, Alison Pease, and Catherine Pelachaud. 2018. Is two better than one? Effects of multiple agents on user persuasion. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*. 255–262.

[13] Reshmashree B Kantharaju and Catherine Pelachaud. 2021. Social Signals of Cohesion in Multi-party Interactions. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*. 9–16.

[14] Sungwoo Kim. 2019. Playing with machine translation in language classroom: Affordances and constraints. *Multimedia-Assisted Language Learning* 22, 2 (2019), 9–28.

[15] Colleen M Lewis. 2011. Is pair programming more effective than other forms of collaboration for young students? *Computer Science Education* 21, 2 (2011), 105–134.

[16] Arcanjo Miguel Mota Lopes, José Francisco de Magalhães Netto, Ricardo AL de Souza, Andreza Bastos Mourão, Thais Oliveira Almeida, and Dhanielly PR de Lima. 2019. Improving students skills to solve elementary equations in k-12 programs using an intelligent tutoring system. In *2019 IEEE Frontiers in Education Conference (FIE)*. IEEE, 1–9.

[17] Neil Mercer. 2002. *Words and minds: How we use language to think together*. Routledge.

[18] Nahal Norouzi, Kangsoo Kim, Jason Hochreiter, Myungho Lee, Salam Daher, Gerd Bruder, and Greg Welch. 2018. A Systematic Survey of 15 Years of User Studies Published in the Intelligent Virtual Agents Conference. In *Proceedings of the 18th international conference on intelligent virtual agents*. 17–22.

[19] Bhargavi Paranjape, Yubin Ge, Zhen Bai, Jessica Hammer, and Justine Cassell. 2018. Towards automatic generation of peer-targeted science talk in curiosity-evoking virtual agent. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*. 71–78.

[20] Cynthia A Rohrbeck, Marika D Ginsburg-Block, John W Fantuzzo, and Traci R Miller. 2003. Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of educational Psychology* 95, 2 (2003), 240.

[21] Sylvia Rojas-Drummond and Neil Mercer. 2003. Scaffolding the development of effective collaboration and learning. *International journal of educational research* 39, 1-2 (2003), 99–111.

[22] Sylvia Rojas-Drummond and Margarita Peon Zapata. 2004. Exploratory talk, argumentation and reasoning in Mexican primary school children. *Language and education* 18, 6 (2004), 539–557.

[23] Jeremy Roschelle and Stephanie D Teasley. 1995. The construction of shared knowledge in collaborative problem solving. In *Computer supported collaborative learning*. Springer, 69–97.

[24] Noah L Schroeder, Olusola O Adesope, and Rachel Barouch Gilbert. 2013. How effective are pedagogical agents for learning? A meta-analytic review. *Journal of Educational Computing Research* 49, 1 (2013), 1–39.

[25] Noah L Schroeder and Chad M Gotch. 2015. Persisting issues in pedagogical agent research. *Journal of Educational Computing Research* 53, 2 (2015), 183–204.

[26] Harriet R Tenenbaum, Naomi E Winstone, Patrick J Leman, and Rachel E Avery. 2020. How effective is peer interaction in facilitating learning? A meta-analysis. *Journal of Educational Psychology* 112, 7 (2020), 1303.

[27] Jennifer Tsan, Jessica Vandenberg, Collin Lynch, Eric Wiebe, and Kristy Elizabeth Boyer. 2021. Examining Collaborative Support for Block-Based Programming with Upper Elementary Students. In *Proceedings of the 14th International Conference on Computer-Supported Collaborative Learning-CSCL 2021*. International Society of the Learning Sciences.

[28] Jessica Vandenberg, Arif Rachmatullah, Collin Lynch, Kristy E Boyer, and Eric Wiebe. 2021. Interaction effects of race and gender in elementary CS attitudes: A validation and cross-sectional study. *International Journal of Child-Computer Interaction* 29 (2021), 100293.

[29] Jessica Vandenberg, Arif Rachmatullah, Collin Lynch, Kristy Elizabeth Boyer, and Eric Wiebe. 2021. The Relationship of CS Attitudes, Perceptions of Collaboration, and Pair Programming Strategies on Upper Elementary Students' CS Learning. In *Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 1*. 46–52.

[30] Xuefeng Wei, Lin Lin, Nanxi Meng, Wei Tan, Siu-Cheung Kong, et al. 2021. The effectiveness of partial pair programming on elementary school students' computational thinking skills and self-efficacy. *Computers & education* 160 (2021), 104023.

[31] Joseph B. Wiggins, Jamieka Wilkinson, Lara Baigorria, Yingwen Huang, Kristy Elizabeth Boyer, Collin Lynch, and Eric Wiebe. 2019. From Doodles to Designs: Participatory Pedagogical Agent Design with Elementary Students. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children*. ACM, 642–647.

[32] Beverly Woolf, Winslow Burleson, Ivon Arroyo, Toby Dragon, David Cooper, and Rosalind Picard. 2009. Affect-aware tutors: recognising and responding to student affect. *International Journal of Learning Technology* 4, 3-4 (2009), 129–164.

[33] Fatma Gizem Karaoglan Yilmaz and Ramazan Yilmaz. 2019. Impact of pedagogic agent-mediated metacognitive support towards increasing task and group awareness in CSCL. *Computers & Education* 134 (2019), 1–14.

[34] Won Jin Yu, Jun Hyeok Jang, Joong Min Ahn, Dae Ryoon Park, In Hwan Yoo, Young Kwon Bae, and Woo Yeol Kim. 2019. Design of artificial intelligence education program based on design-based research. *International journal of advanced smart convergence* 8, 4 (2019), 113–120.

[35] Zarifa Zakaria, Jessica Vandenberg, Jennifer Tsan, Danielle Cadieux Boulden, Collin F Lynch, Kristy Elizabeth Boyer, and Eric N Wiebe. 2022. Two-computer pair programming: Exploring a feedback intervention to improve collaborative talk in elementary students. *Computer Science Education* 32, 1 (2022), 3–29.