Simulated RN | Virtual Healthcare Agent
M5 | William Morse

SUMMARY

The Simulated RN is a virtual healthcare agent designed to provide interactive patient education at the time of discharge in the post-partum unit of the hospital. The discharge process can have an impact on numerous factors associated with patient care and hospital-related outcomes. The discharge process influences many aspects of healthcare, including patient understanding of their care, patient satisfaction, the potential for adverse outcomes, hospital efficiency connected to bed availability as well as the overall quality of hospital services. The most common factor influencing the frequency of adverse events in healthcare relate to poor communication between hospital caregivers and the patient during the delivery of discharge instructions (Forster, Murff, Peterson, Gandhi & Bates, 2003).

The Simulated RN project will utilize a virtual healthcare agent that is designed to deliver face-to-face consultation with patients while upholding a natural appeal. The main design goal was to improve the discharge process by providing an interactive learning environment that enriches user tasks through communicative commonality (e.g., reduced communication barriers). It is predicted that reducing barriers in communication during the discharge instruction process will subsequently improve the acquisition of knowledge, patient comprehension and satisfaction.

To address concerns with literacy levels for some patients, and attending to the needs of the hearing-impaired, the interface was designed to support audio narration with written subtitles to accommodate users (Johnson, Sandford, & Tyndall, 2003). Furthermore, multisensory functions were needed to accommodate various user learning styles (Johnson, Sandford, & Tyndall, 2003), including subject matter content supported by visual aids (Koonce, Giuse, & Storrow, 2011). Two major driving

forces behind the prototype strategy centered on the psychological aspects related to high agency and high behavioral realism.

A preliminary version of the system was prepared with (2) main aims:

1. A high degree of functionality
2. Consistency with interaction

The prototype of the system was designed with these main features:

- 3D animated virtual healthcare agent
- Computer generated speech, including a demonstration in Spanish (SIDS section)
- Subtitles
- Scripted conversation intended to communicated workflow, including instructions
- Touch based user responses, using modal dialogue boxes
- Ability to pause workflow at anytime
- Educational content related to post-partum discharge instructions:
  - proper feeding of the infant,
  - urination patterns, bowel movements,
  - umbilical cord care, skin care,
  - genital care,
  - signs of illness,
  - prevention of sudden infant death syndrome,
  - car seat selection
  - vaccine information
- Optional instructional video (*only a placeholder*)

METHODS

Selection

Five associates, with varying degrees of relationship to the target audience were invited to participate in this study. Participants were from a generalized nonprobability convenience sample selected to represent the primary user (mothers of newborns) and secondary user (maternity experts) groups.

Participants

Five participants (1 male, 4 females) volunteered to partake in the usability study. One participant was a physician, two participants were RNs (1 from labor and delivery), and two participants represented mothers of recent newborns (having given birth within the past year). Therefore, two participants represented the primary user group (mothers of newborns), while three participants represented the

secondary user group (maternity experts). All participants were English speaking. The age of the participants included two participants between the ages of 20-29, two participants between the ages of 30-39 and one participant between the ages of 40-49. No control group was used for this study.

Procedures

A usability study was conducted to explore and assess the functionality of the Simulated RN prototype. The study was exploratory seeking to gather information about the prototype and to assess its preliminary design utility. The study required a user-based test to measure a set of representative tasks in the Simulated RN prototype interface. Each participant was scheduled for an individual 20-minute session and all sessions were conduct on the same day.

A controlled setting was used to conduct each session. The testing environment included a small office room with a desk and two chairs. On the desk sat the Simulated RN prototype interface installed on a Dell Inspiron 2320 all-in-one computer. The same volume and screen settings were used for all participants. Participants used the Dell Inspiron's touch screen for all user inputs, no keyboards or mouse were used.

The test moderator sat next to the participant during the test sessions. The test moderator introduced the session and provided the participants with a brief set of instructions. During the session, the frequency count and time measures were obtained automatically by the interface and were stored into a database for each participant. In addition, an accuracy measure was obtained after each user touch input to verify the participant's intention through a follow-up question asked by the test moderator. Participant verification was provided through verbal confirmation and recorded on paper.

Finally, a post-session self-reporting qualitative survey was handed out to each participant to gather subjective data related to interface features. The self-reported measures asked the participants to respond using the rating scale.

Measures

User Study

The measurement criteria were based on Nielsen's Usability Engineering (1993) measureable goals: time to learn how to operate the system, speed of user performance, and the rate of errors made by users and the user's satisfaction with the system. All data will be used to summarize information related to the measureable goals in an attempt to determine system usability.

Quantitative Usability Measures

Quantitative measures were derived through these data sources:

1. Count frequency related to failure to make a selection;
2. Count frequency related to the accuracy of a selection (e.g., correct or incorrect intention of selection);
3. Time required to make a selection in seconds.

Qualitative Usability Measures

Qualitative measures will focus on attitudes toward the agent and satisfaction with the concept. These measures will be self-reported items used to assess overall satisfaction with the agent, ease of system use, and preference for human or agent. A survey instrument was created to capture responds to these items (see Appendix A, Simulated RN Survey).

Tasks

The tasks for this exploratory usability study are designed to answer the basic question of what are the main stumbling blocks for the user. The system begins with an introductory sequence, followed by instructions on how to use the prototype, including an evaluation of the user's readiness or ability to understanding the navigation process (see Appendix B, Task Flowchart).

Task 1 Description

The user will be given instructions by the virtual healthcare agent outlining the intended workflow, including how to provide user input and navigation. The first task is to determine whether the user understands the input instructions by answering the question—Are you ready to begin?

User Specific Task: Allow users to navigate the interface when they are prepared, or to locate desired information.

Required to Perform

The first task is touch-based requiring the user to identify the appropriate area within the user interface and to apply touch to initiate user input.

Success Criteria

The user must determine whether they are ready to begin. The input is recorded by the system. For example, the user may select either yes or no, either of which may indicate the user understands how to provide user input. Either response is considered valid. To further confirm the user's intention, an immediate follow-up question will be asked by the test moderator, "Did you intend to make that selection?" Only a participant response of yes will be considered as a success (accuracy, related to errors).

No Response Criteria

On this particular task, the user is given a 2-minute window to reply and if this time is exceeded, the system will automatically provide follow-up user instructions. The system is designed to reset its response clock and provide a second attempt. Failure to respond a second time will cause a full reload of the process, including introduction and instructions. Rationale: In real-world situations a failure to respond may indicate such things as the user is busy with something else (e.g., baby is crying, or family members are talking with the patient) and not necessarily a system failure. Therefore, the system has been built to consider this potential. For usability testing, failure to respond within the first 2-minutes will be considered a failure.

Maximum Time: 4-minutes (2-minutes per attempt; 2-attempts)

Task 2 Description

Educational content will load and the virtual healthcare agent will provide additional instructions and material. The Simulated RN is designed to monitor the progress of the user selections. The second task is oriented at determining a user's preference for feeding their infant. The user will be asked if they plan to breastfeed their baby. In this particular example, the users will either chose to breastfeed or formula feed.

User Specific Task: Offer various instructional opportunities for the user to explore.

Required to Perform

The second task is touch-based requiring the user to identify the appropriate area within the user interface and to apply touch to initiate user input.

Success Criteria

Success is not determined by the actual response, but the intention of the response, like in the previous task. The user will either select breastfeeding as their choice or default to formula feeding. Therefore, the user will make a selection and an immediate follow-up question will be asked by the test moderator, "Did you intend to make that selection?" Only a response of yes will be considered as a success.

No Response Criteria

On this particular task, the user is given a 5-minute window to reply and if this time is exceeded, the system will automatically provide follow-up user instructions. For example, the system is designed to say, "I haven't received your response, please make a selection."

The system is designed to reset its response clock and provide an endless number of attempts. The system is built to consider that the newborn mother may be busy with the baby or interacting with other people in the hospital room, therefore, the system waits and every 5-minutes attempt to initiate a user response. For usability testing, failure to respond within the first 5-minutes will be considered a failure.

Maximum Time: 5-minutes

Task 3 Description

The virtual healthcare agent will provide additional instructions, education and media content. The third task is another user preference. The Simulated RN will ask the question, "Would you like to watch an education video related to feeding baby?" The user will select either yes or no.

User Specific Task: Accommodate the users learning styles by providing them learning options.
Required to Perform

The third task is touch-based requiring the user to identify the appropriate area within the user interface and to apply touch to initiate user input.

Success Criteria

Like all previous tasks, the success is not determined by the actual response, but the intention of the response. The user will either choose to watch an instructional video or not. Therefore, the user will make a selection and an immediate follow-up question will be asked by the test moderator, "Did you intend to make that selection?" Only a response of yes will be considered as a success.

No Response Criteria

The same no response criteria used in Task 2 will be applied for Task 3.

Maximum Time: 5-minutes

RESULTS

Quantitative Results

The *Task Analysis Table* reflects participant data related to their touch interactions with the system interface. The metrics include frequency counts and measures of time in seconds. The participant distribution is the key factor of this table.

Task Analysis Table

	Task 1			Task 2			Task 3		
	Failure	Accuracy	Time	Failure	Accuracy	Time	Failure	Accuracy	Time
Participant 1	0	1	10	0	1	8	0	1	7
Participant 2	0	1	12	0	1	7	0	1	7
Participant 3	0	1	13	0	1	9	0	1	8
Participant 4	0	1	10	0	1	7	0	1	8
Participant 5	0	1	10	0	1	7	0	1	7
Totals	0	5	55	0	5	38	0	5	37
Mean (in sec)			11			7.6			7.4

The *Task Summary Table* reflects participant data related to their touch interactions with the system interface. The metrics include frequency counts and measures of time in seconds. The task distribution is the key factor of this table.

Task Summary Table

	Participants	Failure Count (failed/attempts)	Accuracy Count (correct/attempts)	Mean Time Required Selecting (in seconds)
Task 1	5	0 (0 of 5) 0%	5 (5 of 5) 100%	11
Task 2	5	0 (0 of 5) 0%	5 (5 of 5) 100%	7.6
Task 3	5	0 (0 of 5) 0%	5 (5 of 5) 100%	7.4
Mean	5	0%	100%	8.67

Task 1: All users were successful and accurate with their input and averaged 11 seconds to respond.

Task 2: All users were successful and accurate with their input and averaged 7.6 seconds to respond.

Task 3: All users were successful and accurate with their input and averaged 7.4 seconds to respond.

Qualitative Results

The *Participant Information Table* reflects participant demographic data. The metrics include frequency counts and scaled replies.

Participant Information Table

Gender	Male 1 (20%)	Female 4 (80%)		
Age	<20 0 (0%)	20-29 2 (40%)	30-39 2 (40%)	40-49 1 (20%)
User group	Primary (mothers of newborns within the past year) 2 (40%)	Secondary (maternity experts) 3 (60%)		

The *General Questions Table* reflects participant survey data. The metrics include frequency counts and scaled replies.

General Questions: Question 1 Table

1.	The Simulated RN was easy for me to use.	Agree Strongly 4 (80%)	Agree 1 (20%)	Do not agree or disagree 0 (0%)	Disagree 0 (0%)	Disagree Strongly 0 (0%)
	Gender	1 male 3 female	1 female			
	Age	2 20-29 1 30-39 1 40-49	1 30-39			
	User Group	2 primary 2 secondary	1 secondary			

All participants agree to an extent with the statement, the Simulated RN was easy to use.

General Questions: Question 2 Table

2.	Learning to navigate the Simulated RN was easy for me.	Agree Strongly 4 (80%)	Agree 1 (20%)	Do not agree or disagree 0 (0%)	Disagree 0 (0%)	Disagree Strongly 0 (0%)
	Gender	1 male 3 female	1 female			
	Age	2 20-29 1 30-39 1 40-49	1 30-39			
	User Group	2 primary 2 secondary	1 secondary			

All participants agree to an extent with the statement, learning to navigate the Simulated RN was easy for me.

General Questions: Question 3 Table

3.	In my opinion, using the	Agree	Agree	Do not	Disagree	Disagree

	Simulated RN would enhance the patient discharge process.	Strongly		agree or disagree		Strongly
		2 (40%)	2 (40%)	1 (20%)	0 (0%)	0 (0%)
	Gender	2 female	1 male 1 female	1 female		
	Age	1 20-29 1 30-39	1 20-29 1 40-49	1 30-39		
	User Group	2 primary	2 secondary	1 secondary		

4 of 5 participants agree to an extent with the statement, using the Simulated RN would enhance the patient discharge process. 1of 5 did not agree or disagree.

General Questions: Question 4 Table

4.	In my opinion, using the Simulated RN would improve patient comprehension.	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
		3 (60%)	2 (40%)	0 (0%)	0 (0%)	0 (0%)
	Gender	3 female	1 male 1 female			
	Age	1 20-29 1 30-39 1 40-49	1 20-29 1 30-39			
	User Group	2 primary 1 secondary	2 secondary			

All participants agree to an extent with the statement, using the Simulated RN would improve patient comprehension.

General Questions: Question 5 Table

5.	In my opinion, using the Simulated RN would improve patient satisfaction with the discharge process.	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
		2 (40%)	2 (40%)	1 (20%)	0 (0%)	0 (0%)
	Gender	2 female	1 male 1 female	1 female		
	Age	1 20-29 1 30-39	1 20-29 1 40-49	1 30-39		
	User Group	2 primary	2 secondary	1 secondary		

4 of 5 participants agree to an extent with the statement, using the Simulated RN would improve patient satisfaction with the discharge process. 1of 5 did not agree or disagree with the statement.

General Questions: Question 6 Table

6.	The virtual agent's movements appear natural	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
		0 (0%)	1 (20%)	1 (20%)	3 (60%)	0 (0%)
	Gender		1 female	1 female	1 male 2 female	
	Age		1 20-29	1 30-39	1 20-29 1 30-39 1 40-49	
	User Group		1 primary	1 secondary	1 primary 2 secondary	

3 of 5 participants disagree with the statement, the virtual agent's movements appear natural. 1 of 5 agree with the statement. 1 of 5 did not agree or disagree with the statement.

General Questions: Question 7 Table

7.	The virtual agent's speech was easy to understand	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
		0 (0%)	2 (40%)	0 (0%)	3 (60%)	0 (0%)
	Gender		2 female		1 male 2 female	
	Age		1 20-29 1 30-39		1 20-29 1 30-39 1 40-49	
	User Group		1 primary 1 secondary		1 primary 2 secondary	

3 of 5 participants disagree with the statement, the virtual agent's speech was easy to understand. 2 of 5 agree with the statement.

General Questions: Question 8 Table

8.	I was satisfied working with the virtual agent	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
		0 (0%)	4 (40%)	1 (20%)	0 (0%)	0 (0%)
	Gender		1 male 3 female	1 female		
	Age		2 20-29 2 30-39	40-49		
	User Group		2 primary 2 secondary	1 secondary		

4 of 5 participants agree with the statement, I was satisfied working with the virtual agent. 1of 5 did not agree or disagree with the statement.

General Questions: Question 9 Table

9.	If I had a choice between a human agent or a virtual agent, I'd select:	Human 3 (60%)	Virtual Agent 2 (40%)
	Gender	1 male 2 female	2 female
	Age	1 20-29 1 30-39 1 40-49	1 20-29 1 30-39
	User Group	1 primary 2 secondary	1 primary 1 secondary

3 of 5 participants, preferred a human agent to a virtual agent.

General Questions: Question 10 Table

And Why? (follow-up question to question 9)		
Participant 1	Human	I believe a human would sound clearer, or would be easier to understand.
Participant 2	Virtual	The virtual agent was great. I actually like the idea that she is not a person. I think that helps in getting the message out to a diverse population. Her movements can only improve, so I think I'd prefer the virtual agent.
Participant 3	Human	I realize this is in the early stages of development, but the virtual agent was at times difficult to understand. I think it has great potential, but I find the virtual agent distracting.
Participant 4	Virtual	I was comfortable with the virtual agent. I think an interactive human would be a little creepy, like human puppet. As a result, I think knowing that the virtual agent is not real satisfies the situation, computerized environment with computerized people.
Participant 5	Human	I have to believe a human would appear and sound more natural. However, I was ok with working with the virtual agent. I think more work on the sound and movements would vastly improve the concept.

Participant Comments

"The hand gestures or movements are a little creepy at times."

"Her pronunciation of a few words was difficult to understand."

"The text size could be a little larger, I found that the text to be small in relationship to the rest of the material."

"She is not wearing an ID badge."

"I think with some fine tuning this could really be useful."

DISCUSSION

The prototype performed well and provided a preliminary level of usability, validated through evidence gathered in the user-based test. All five participants successfully completed all 3 tasks without failure or error. In addition, all participants agreed to some extent that the system was easy to use and navigate.

The mean time for all tasks averaged 8.67 seconds to complete, beginning with an average 11 second time to response in task 1, to a 7.6 second average on task 2 and to a 7.4 second average on task 3. The time required to complete tasks incrementally trended downward, perhaps indicating that the users were getting more familiar with how to operate the system. The effects of prolonged system use could indicate an increase in task efficiency.

In terms of Neilson's Usability Engineering measureable goals, the time to learn how to operate the system, speed of user performance and rate of errors the data indicates that the system's functionality supports the user's tasks. However, the Simulated RN Survey highlighted user's satisfaction with the system suggested that there is more work to be done.

The virtual healthcare agent performed well, four of five participants claimed they were satisfied working with the virtual agent. Furthermore, four of five participants agreed to some extent that the Simulated RN would improve patient satisfaction. In terms of the high agency and high behavioral realism, three of five participants had difficulty understanding the computer generated speech and three of five disagreed that the movements appeared natural. Participants were asked, if they had a choice between a human agent and a virtual agent, three of five participants chose a human. There was a direct correlation with these three participants, each citing problems related to the quality of speech.

The main implications of the prototype related to the body language and movements (high behavioral realism), and most of all, the quality of the computer generated speech (high agency). All of the areas studied are critically dependent on communication. Surprisingly, despite the apparent deficiency in the quality of speech, it had no influence on user performance. This particular area is of significance and requires more emphasis in future design and testing.

In one sense, the data suggest the high agency and high behavioral realism required more attention in future development, but all participants agreed that the Simulated RN would improve patient comprehension and four of five participants suggested it would enhance the patient discharge process. As one participant stated, "I think with some fine tuning this could really be useful."

In conclusion, the prototype process and usability study aimed to determine required changes in the structure of the design. These results show the Simulated RN prototype design has feasible utility as a product. However, as systems are characterized by increasing levels of automation, the virtual agent's activities are increasingly driven by the user's response to the system. The structure of task performance to user satisfaction will require more work to improve the overall usability. Limitations of this study include a smaller sample size. In addition, the use of a convenient sample may provide some face validity, but could also skew results and the interobserver reliability. An empirical investigation will be needed in the future to demonstrate the effectiveness of the proposed design on other aspects related to patient acquisition of knowledge, comprehension and satisfaction.

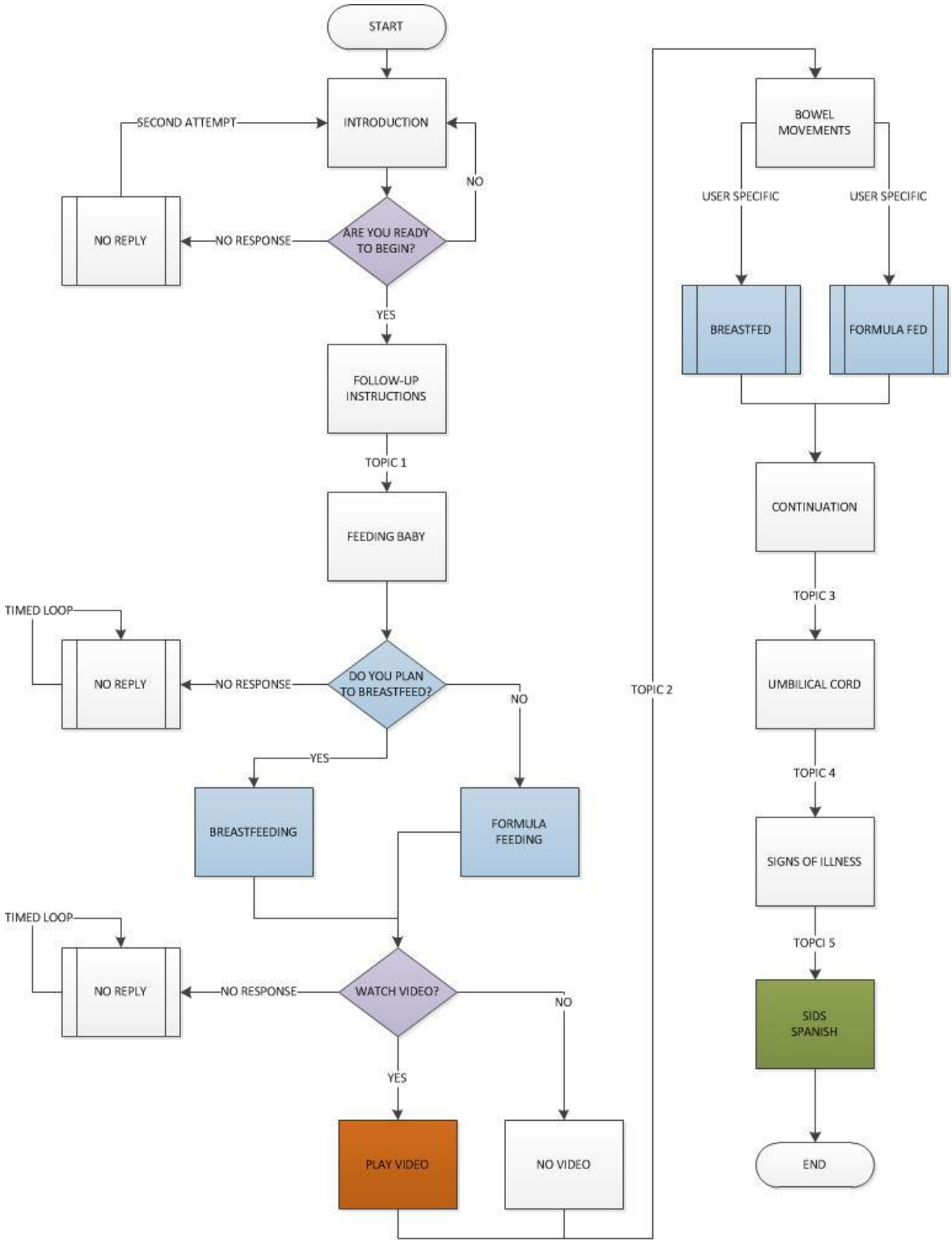
APPENDIX A

Simulated RN Survey

Participant Information					
Gender	Male	Female			
Age	<20	20-29	30-39	40-49	
User group	Primary (mothers of newborns within the past year)	Secondary (maternity experts)			
General Questions					
1. The Simulated RN was easy for me to use.	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
2. Learning to navigate the Simulated RN was easy for me.	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
3. In my opinion, using the Simulated RN would enhance the patient discharge process.	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
4. In my opinion, using the Simulated RN would improve patient comprehension.	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
5. In my opinion, using the Simulated RN would improve patient satisfaction with the discharge process.	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
6. The virtual agent's movements appear natural	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
7. The virtual agent's speech was easy to understand	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
8. I was satisfied working with the virtual agent	Agree Strongly	Agree	Do not agree or disagree	Disagree	Disagree Strongly
9. If I had a choice between a human agent or a virtual agent, I'd select:	Human	Virtual Agent			
10. And Why? (follow-up to question 9)					
Comments:					

APPENDIX B

Task Flowchart



REFERENCES

Forster, A., Murff, H., Peterson, J., Gandhi, T., Bates, D. The Incidence and Severity of Adverse Events Affecting Patients after Discharge from the Hospital. Annals of Internal Medicine, 138, (2003).

Johnson, A., Sandford, J., Tyndall, J. (2003) Written and verbal information versus verbal information only for patients being discharged from acute hospital settings to home. Cochrane Database Syst Rev. 2003; (4): CD003716.

Koonce, T., Giuse, N., Storrow, A. (2011) A pilot study to evaluate learning style-tailored information prescriptions for hypertensive emergency department patients. J Med Libr Assoc 99 (4) October 2011.