Off-Topic Recommendations

Elyon DeKoven
223 Avenue G, Redondo Beach, CA 90277
playful@dekoven.com

ABSTRACT

In order to make useful recommendations, the product and the person using the product need to focus on the same task. However, when working with a product, people may explore new product features or change their mind about what they want to do. Such behavior can confuse recommendation agents, unless people and their products communicate about the shift of focus.

This paper contributes to the design of recommendation agents that offer off-topic suggestions and consistently handle focus shifting. The paper describes key issues in the design of recommendations under focus shifting, and the three possible dialogue strategies for handling focus shifts. The results of an initial study on one of the strategies indicate trends in how people combine use of off-topic recommendations with focused action in the rest of the product interface. Study results also inform a model of focus shifting behavior, presenting a baseline for future research on the design of off-topic recommendations.

Keywords

Plan recognition, focus shift, dialogue design, recommendation agents

INTRODUCTION

Recommendation agents help people sift through options to come up with better alternatives. In order to make useful suggestions, an agent needs to have a good idea of what the user currently wants or is trying to do. However, at times people may explore unfamiliar product features or change their mind about what they want to do. Indeed, shifting focus can at times be the best thing to do.

People do not always tell their products when they shift focus. As (Lesh, Rich and Sidner, 2001) point out, "interruptions are part of natural collaborative behavior." An agent must be able to determine whether the user's next actions or selection should be interpreted within the same context as the previous one. Typically, at any point during product use there will be more than one reasonable interpretation of what the user is trying to do. Thus the product will rarely have a single 'best' guess.

When the agent is in doubt about the user's current intentions, guessing can lead to problems. For example, changing the temperature on a thermostat could be an attempt to save energy or increase personal comfort. In such cases, the thermostat could guess one or the other and offer advice, but if the guess is wrong the advice may have an adverse affect on the user's trust and use of the product. As Kuhme et al. (1993) say: "the provision of guidance has to be designed very carefully since wrong assumptions about the appropriateness of items can cause fatal problems. Obviously, misleading the user would be even worse than no guidance at all."

Guess, asking and waiting

Instead of risking a wrong guess, the system could ask the user to tell the system more explicitly about what the user is trying to do. Conversational recommendation agents bring the human in the loop to make better decisions. Recent work on such agents have examined *item* recommendations, such as comparing among restaurants (Thompson, Goker and Langley, 2004), and others have explored *action* recommendations, such as manipulating constraints to find available flights (Rich, Sidner and Lesh, 2001). The former relates more to the content of the domain, and the latter to the structure of the person-product communication. In both cases, the agent combines guessing with asking the user clarifying questions. However, such prompts can be an unwelcome intrusion, especially if the user is forced to respond to the prompt in order to continue.

	Guess	Ask	Wait
At best	Leads to efficient product usage and smooth task flow	Clarify what user wants to do, and how to do it	Avoid disturbing user
At worst	Leads to increase in user confusion and decrease in user trust	Annoy user if able to proceed without assistance	Withhold vital assistance when the user is lost

Table 1: Possible dialogue strategies and resulting user experience, in

In order to properly handle situations when a conversational agent is not completely sure about which way the user wants to go, agent design must address tradeoffs between *asking*, *guessing*, and *waiting* (see Table 1). In comparing asking versus guessing, Lesh et al. (2001) note that "The right balance depends, in part, on how often typical users unexpectedly shift their task focus and how often this intention is verbally communicated to the agent."

Towards designing the 'right balance', there is currently little empirical evidence concerning how often people shift focus. Lesh et al. (2001) offer a possible model as to how often people change focus with and without communicating is offered. They suggest that typical users' actions will be focused about 90% of the time, unexpected

Workshop: Beyond Personalization 2005
IUI'05, January 9, 2005, San Diego, California, USA
http://www.cs.umn.edu/Research/GroupLens/beyond2005

focus shifts (within-task) about 5% of the time, and interruptions about 5% of the time. If correct, this model implies about **one out of every ten** actions people take with their products would not fit with what they were doing up to that point. In order to give good recommendations, agents must gracefully handle focus shifting.

Lesh et al. (2001) describe their discourse interpretation algorithm as being "optimized for users who seldom make unexpected focus shifts and, when they do, verbally communicate their intention roughly half the time." While this degree of focused behavior may be expected of 'typical' human collaborators, such an algorithm may not be optimal for first-time or occasional product use. Beginning users may not know how to communicate to the product about what it is they actually want to do. They may make frequent mistakes or randomly explore an interface. Even for people more experienced in using the product, during the course of solving open problems, such as finding a 'good enough' solution to saving energy while maintaining comfort, people might need to switch from one aspect of the problem to another, and from one strategy to another. Thus at times, the best recommendations the agent might have little or nothing to do with what the user has just done.

The rest of this paper examines the design of recommendation systems with respect to focus shifting and off-topic recommendations. Designing support for focus shifting requires three primary considerations: what the product should do when people go off topic, which off-topic items or actions to recommend, and design of the recommendation interface. This paper addresses all three of these issues in the context of action recommendation agents. First, three distinct conversational styles are identified and compared, and a recommendation interface, called the **S**om**e** Thi**n**gs To **Say** (SenSay) menu, is described. Then, results of a study are reviewed towards an empirical model of focus shifting in person-product collaboration.

FOCUS-SHIFTING AND CONVERSATIONAL STYLE

People may have different intentions when shifting focus with regards to whether or not they intend to return to the previous task. In human-human communication, different types of focus shifting are often indicated by linguistic and contextual cues (Grosz and Sidner, 1986). However, not all focus shifts are explicitly communicated.

For example, suppose two agents, say Maya and Reina, were doing something together, say playing marbles. Suddenly, Reina stands up and walks towards the kitchen. Maya might reasonably guess that Reina is hungry. Regardless, how would Maya know whether or not Reina plans on going back to the marbles, or whether she and Reina should clean them up? Choosing the right interpretation is necessary to ensure smooth communication.

Actually, choosing the right interpretation strategy to use depends on more than one focus shift; it is necessary to consider what happens over multiple sequential focus shifts. For example, suppose in the middle of eating, Reina runs off to play with marbles again. Should Maya think that playing with marbles is an interruption of eating as an interruption of playing marbles?

In order to maintain shared focus with the user, the agent can ask, guess or wait, as shown Table 1. The agent's choice hinges on how whether or not the user is done with the interrupted task. In this respect, waiting for more evidence would be the same as guessing that the user is indeed not done. The agent has thus only three options when the user tries to shift focus: ask the user, guess the user is done, or guess the user is not done. This leads to the following three interpretation strategies (see Table 2):

1) <u>Presumptive</u>: The first strategy in Table 2, *asking*, prevents the user from shifting focus until the user has stated whether or not they were done with the previous task. Such a strategy ensures the user is aware of the focus shift, and makes sure the system makes the right interpretation. The strategy is called 'presumptive' since the strategy implies the system knows what is in the best interests of the user, and hence the user will find it worthwhile to answer the system's questions. However, such questions could be confusing to the user if the user does not know whether or not to come back to the task. In the running example above, Reina may not be sure she wants to go back to playing marbles after she is done eating. This strategy also adds an extra turn at every focus shift for confirmation. As Constantine and Lockwood (1999, p. 257) say "Confirmations interrupt the progress of work and annoy users. Nearly all confirmations are unnecessary or ineffective." People might evaluate such extra effort as undesirable, and make them less likely to want to shift focus.

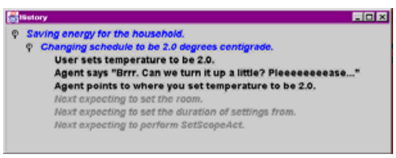
2) <u>Nested Interruptions</u>: The second strategy in Table 2, *guessing the user is not done*[1], creates many levels of

Strategy	Description	Agent	User	Analysis
Ask (Presumptive)	Prevent focus shift until user confirms done	Has user consent	May not be sure whether to shift	Adds extra turns and possibility of nested interruptions
Guess Not done (Nested interruptions)	Assume user will go back to previous task	No user consent	May not be aware of shift	Nested interruptions
Guess Done (Go with the flow)	Assume user is done with previous task	No user consent	May not be aware of shift	No interruptions

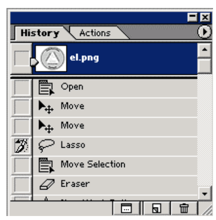
Table 2: Interpretation strategies when the user shifts focus.

nested interruptions. In the running marbles example, suppose in the middle of eating a cookie Reina comes back to the marbles. Adopting this strategy, Maya would think that Reina is not yet done eating the cookie, and thus that playing with marbles is an interruption of snacking, which itself was an interruption of playing with marbles. Such situations could complicate communication, and recovery from communication failures. While nested interruptions could also occur with the presumptive strategy if the user is not done, the guessing-not-done strategy does not have the advantage of the presumptive strategy of informing the user that a focus shift took place.

Additional interface and dialogue mechanisms could alleviate these concerns, such as Collagen's segmented History window (Rich and Sidner, 1998) or the History window in Adobe PhotoShop 6.0, but might overly complicate a simple interface such as a thermostat. A second problem with this strategy occurs when actions



The segmented history window from Collagen shows dialogue topics in a nested fashion.



History window in Adobe PhotoShop 6.0.

Figure 1: Graphical representations of task focus.

taken during interruptions block performance of the interrupted task. For example, if Reina accidentally kicks all the marbles out the door on her way to the kitchen, it would no longer be possible to go back to the game. In the general case, competently handling such situations in collaborative agent design requires what are known as full causal models, containing preconditions, such as 'there are marbles with which to play', and postconditions, such as 'the marbles are back in their bag', specifying every possible situation in which the task is still possible to perform. Creating complete causal models can be difficult and time-consuming. In fact, one of the strengths of the planning algorithm in Lesh et al. (2001) is its ability to plan with partial models.

3) Go with the Flow: The third strategy in the table, *guessing the user is done*, does not add extra turns as does asking, and prevents nested interruptions. However, with this strategy there is also no context left around when the interrupting task is complete. For example, when Reina comes back to playing marbles after eating, Maya would not remember where they were in the marble game they were playing, nor that the marbles were still out and need to be put away. Moreover, this strategy has the potential disadvantage of increasing peoples' experience of lostness. Since the agent does not interfere with the user's task switching, the user may not be aware a switch has occurred. The user might want to know 'how do I get back to where I was?' Since there is no context left, the agent cannot answer the question.

SUPPORTING TASK FOCUS

As mentioned above, designing off-topic recommendations requires an understanding of agent dialogue strategy, recommendation interface design and human focus shifting behavior. The previous section described the first of these. Turning now to the second issue, there is little known about how to design an interface to help people usefully switch focus without confusing the product.

Existing efforts to design adaptive recommendations have focused on focusing the user, or at least confirming the user's current intentions. For example the 'intent interface' in (Miller and Hannen, 1999) gives the user a view on what the system thinks the user intends and lets the user override the interpretation, and the studies in (Sidner and Forlines, 2002; Freudenthal and Mook, 2003) presented sets of suggested things to say contributing to the user's current task. None of these efforts explicitly support the user in switching to a new task. They do not address situations when people might find value in switching between tasks (or subtasks) as part of normal problem solving, such as balancing heating comfort against energy costs, or planning a vacation the whole family could enjoy and afford.

A recommendation interface must support situations when people unexpectedly shift focus, and to encourage people to communicate about their current intentions. The **S**om**e** Thi**n**gs to **Say** (SenSay) menu addresses this challenge. An example is shown in Figure 3. As described in DeKoven (2004), SenSay contents are generated through generic rules based on the collaborative planning algorithms described in Lesh et al. (2001). The system updates a *focus stack* and *plan tree* by comparing actions to a task model. Simple rules walk the stack and tree to create an *agenda* of likely next steps.

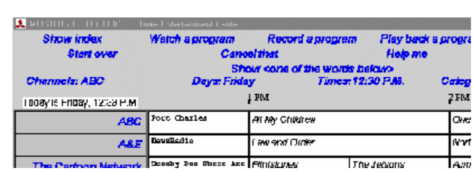


Figure 2: Two graphical interfaces for action recommendation. Left: Clippit agent from Microsoft Office XP; Right: part of a VCR interface described in (Sidner and Forlines, 2002).

For any reasonably complicated apparatus, such as a programmable thermostat, there could be many reasonable next steps the user could take. The decision concerning what to recommend rests with the agent. Typical recommendation systems hide the complexity and uncertainty from the user, and choose one best option from

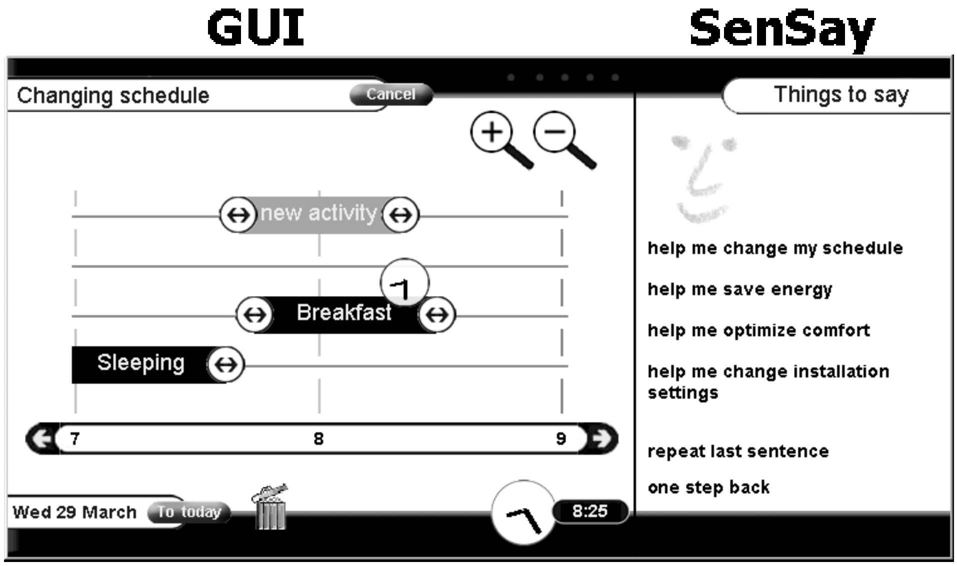


Figure 3: Example of a SenSay menu, in the case of a programmable thermostat, as described in Freudenthal and Mook (2003). The left-hand side of the interface is the basic graphical user interface (GUI), and the right-hand side contains the SenSay.

the agenda upon which to act next. In contrast, the SenSay presents the agenda as an ordered list of phrases the user can do or say to the system (through speech or directly pressing an item).

The SenSay can be viewed as a way to plug the user in to the system's reasoning process, giving the user means to explicitly communicate with the product about current intentions, in a way the system can understand. Since the SenSay allows the user to specify an intention, the system is not forced into making a single guess, nor does it have to ask as many clarifying questions. In effect, with the SenSay present the system can be satisfied with a certain degree of uncertainty. That is, the product can wait until it is more confident about what the user wants to do, while still being helpful.

A recommendation interface like the SenSay can make it easier for people to communicate their intentions to the thermostat. Furthermore, in order to support useful shifting of task focus, the SenSay can include suggestions about related tasks, even if not directly contributing to what the product thinks is the user's current task. By using the right rules for computing the agenda, the SenSay could present next steps for the current task, as well as ways to shift to completely different tasks. For example, while trying to increase home heating comfort by raising temperatures with a thermostat, the SenSay could include suggestions for saving energy by reducing temperatures.

It may be confusing or overwhelming to people if the SenSay were to include all possible focused and off-topic recommendations. It is a design decision as to which suggestions to include in a limited visual space like the SenSay. Finding the right balance requires knowing more about how people tend to shift focus when using a product, such as using a thermostat, to solve optimization problems, such as balancing comfort and costs.

TOWARDS A MODEL OF FOCUS SHIFTING

To reiterate, the goal of this paper is to discuss the design of recommendation systems to handle focus shifting, and offering off-topic suggestions. There is currently little empirical evidence related to how people might want or need to shift focus, and how they might communicate the focus shift with a product. There is even less known about how an interaction mechanism like the SenSay will affect the user's focus shifting and communication behavior.

As a starting point, Lesh et al. (2001) hypothesized that typically 90% of user actions are focused, 5% are focus shifting, and 5% are interruptions. However, there is little evidence upon which to validate this model. Quite likely, the model should be related to choices in agent design, as the different agent strategies described above might influence the results in different directions.

A case study was conducted in order to test this model (DeKoven, 2004). The full study incorporated tests of the SenSay as a multimodal (speech + touch) interface. Only those results relevant to modeling focus shifting and the utility of off-topic recommendations are reviewed here. Full results of the study, along with a more detailed analysis of SenSay item design and development, can be found in (DeKoven, 2004).

The study used an interface similar to that in Figure 3, translated into Dutch. The subjects were given a set of test tasks. Task order was random, except the last task. This task, called *Go Green*, was intended to force a situation in which subjects would need to balance comfort and costs to meet certain constraints.

In order to increase the likelihood of test participants switching focus, this experiment adopted the *Go with the Flow* strategy. One group of study participants used a SenSay containing focused and off-topic recommendations, and the other group used the same SenSay with only the focused items, in a between subjects design. With this setup, it is possible to examine focus-shifting behavior, with and without the presence of off-topic recommendations, under the *Go with the Flow* strategy.

Across both conditions, subjects generally used the SenSay during the *Go Green* task more than during other tasks. As verified in post-test interviews, subjects were inclined to use the SenSay when they did not know what to do with the GUI or were looking for better ways to do something. The results indicate that recommendations in the SenSay can be utilized effectively in complex tasks like saving energy while staying comfortable, but can be distracting when the user can complete the task more quickly in the rest of the interface.

The test subjects did use the off-topic recommendations on the SenSay, though sometimes the items appeared to be confusing at times. On the other hand, several subjects in the focused-SenSay condition were not in the energy saving dialogues at the right time, and, unlike subjects in the focus-shift-SenSay condition, could only get there by starting over or otherwise manually exiting all the subtasks.

	Degree of focus	**St. Dev.**
Predicted (Lesh et al. 2001)	90%	
Observed: All tasks	96%	1.23
Observed: Go Green	89%	2.65

Table 3: Observed proportions of focussed behavior.

That is, the off-topic recommendations served as useful short cuts for flipping between tasks and strategies.

Table 3 summarizes the observed frequency of focused actions. Across all test tasks and conditions, approximately 96% (st. dev. = 1.23) of subjects' actions were focused (cases 1a and 1b), much higher than the prediction in (Lesh et al., 2001). Looking just at actions taken during the *Go Green* task, the last and most difficult of the test tasks, only about 89% (st. dev. = 2.65) of subjects' actions were focused. Thus the predictions in (Lesh et al., 2001) might be more accurate for more difficult user tasks, or more experienced users.

DISCUSSION

In order for an interactive agent to recommend useful next steps and options it needs to keep track of what the user is doing. This gets more difficult if the user shifts focus.

As discussed in this paper, people do shift between tasks and between strategies while working on constraint-solving problems, such as when using a thermostat to balance heating comfort against saving energy. Recommendation agents can help people navigate to better answers, but only if people and their products are focused on the same task and plan. This paper has discussed how to incorporate focus shifting into recommendation agent design.

The primary design question addressed in this paper is how to best balance guessing, asking and waiting when the agent is unsure of the user's current intention. This paper compared three possible response strategies an agent can adopt when the user shifts focus. In particular, the *Go with the Flow* strategy appeared the strongest of the three in terms of allowing for smooth task transitions. It was also the most likely strategy to engender a feeling of lostness. In the study reviewed here, subjects did not indicate strong feelings of lostness. That is, the agent being loose did not lead to the user feeling lost. Moreover, the study subjects used the off-topic SenSay items to quickly switch between alternate strategies. Thus we can say that *Go with the Flow* did in fact support the user in flexibly redirecting the dialogue, via the off-topic recommendations.

These study results need to be understood in terms of the study participants. Most of the participants were young and well educated (university students). Many of the older subjects used in pilot testing in particular had more difficulty completing test tasks. Given the results and designs of other similar studies (Freudenthal and Mook, 2003; Sidner and Forlines, 2002), more agent utterances or visual feedback confirming the focus shifts could have significantly helped subjects find their way with the SenSay. More studies are needed comparing all three strategies with respect to usefulness of off-topic recommendations.

Central to future research in this area is a baseline model of human focus shifting when using a product to solve constraint problems. The study reviewed here is the first to provide empirical evidence towards a predictive model of focus shifting such as that presented in Lesh et al. (2001). The model in that paper appears more correct for experienced usage and/or complicated tasks. More studies are needed to confirm this result. In particular, the same SenSay with a Presumptive or Nested Interruption agent might not lead to the same results. Longitudinal studies would be useful for tracking these relationships. Combining these results with user modeling (such as in Rickel et al., 2002) could lead to creating agents that adaptively alternate between the three strategies.

To what degree should an action recommendation interface support people in shifting focus? Clearly, the modalities and manners of expressing focus shifts impact the frequency of shifting focus. The SenSay discussed in this paper is only one such way to present off-topic recommendations. More design research is needed to better understand the impact of interface design, such as the SenSay, on use of action recommendations in general, and off-topic recommendations in particular.

While more studies are needed to verify the results discussed in this paper, there does appear to be support for the following propositions:

- People do shift focus when solving constraint problems. Moreover, they shift focus for different reasons as they get more experienced and as the tasks get more difficult.

- Graphical interfaces such as the SenSay help people in communicating with a product, at least initially (see Sidner and Forlines, 2003).
- Off-topic recommendations can be useful for problem solving as well as presenting unfamiliar product capabilities. However, they may be more useful for people more experienced with regular focused interface usage.
- Choosing strategies for action recommendation agents must be done in tandem with designing the rest of the interface.

As discussed in this paper, focus shifting is fine, and can even be good, as long as it is communicated. Interfaces for conversational recommendation agents need to motivate people to tell their products what they want to do. The main goal in the line of research leading up to this paper has been to design products to better help people tell the product what they want to do. This has been called the Help Me Help You principle (DeKoven, 2004).

Underlying this research has been the SharedPlans model of human-human collaboration (Grosz and Sidner, 1986), and the discourse interpretation algorithm based on SharedPlans defined in (Lesh, Rich and Sidner, 2001). Unlike most current recommendation agents that are based on fact databases (e.g. information about restaurants), collaborative agents use a task model to determine which actions to recommend. Future work should combine these two lines of work (as was attempted in Rickel et al., 2002) with the design research described in this paper, towards collaborative recommendation agents that can help people with both facts and actions.

ACKNOWLEDGEMENTS
This work was conducted at and supported by the Delft University of Technology. Thanks are due to David Keyson and Jans Aasman for their assistance in focusing the study, and to Marc de Hoogh and Tanja Veldhoen for assistance in preparing and running the study.

REFERENCES
Constantine, L. L., & Lockwood, L. A. D. (1999). *Software for Use: A Practical Guide to the Models and Methods of Usage-Centered Design*. New York: ACM Press.

DeKoven, E. (2004), *Help Me Help You: Designing Support for Person-Product Collaboration*, Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands.

Freudenthal, A., & Mook, H. J. (2003). "The Evaluation of an Innovative Intelligent Thermostat Interface: Universal Usability and Age Differences." *Cognition, Technology and Work*, *5*, 55-66.

Grosz, B. J., & Sidner, C. L. (1986). "Attention, intentions, and the structure of discourse." *Computational Linguistics*, *12*(3), 175-204.

Kuhme, T., Malinowski, U., & Foley, J. D. (1993). *Adaptive Prompting* (No. GIT-GVU-93-05). Atlanta, GA: Georgia Institute of Technology.

Lesh, N. B., Rich, C., & Sidner, C. L. (2001, July). *Collaborating with Focused and Unfocused Users Under Imperfect Communication.* International Conference on User Modelling, Sonthofen, Germany.

Miller, C. A., & Hannen, M. D. (1999). *User Acceptance of an Intelligent User Interface: A Rotocraft Pilot's Associate Example.* Intelligent User Interfaces (IUI'99), Redondo Beach, CA.

Rich, C., & Sidner, C. L. (1998). "COLLAGEN: A Collaboration Manager for Software Interface Agents." *User Modeling and User-Adapted Interaction, 8*(3/4), 315-350.

Rich, C., Sidner, C. L., & Lesh, N. B. (2001, Winter). "COLLAGEN: Applying Collaborative Discourse Theory to Human-Computer Interaction." *Artificial Intelligence Magazine*, *22*, 15-25.

Sidner, C. L., & Forlines, C. (2002, September). *Subset Languages for Conversing with Collaborative Interface Agents.* International Conference on Spoken Language Processing (ICSLP'02), Denver, CO.

Thompson, C.A., Göker, M. H. & Langley, P. (2004). "A Personalized System for Conversational Recommendations." *Journal of Artificial Intelligence Research*, *21*, 393-428.

Rickel, J., Lesh, N., Rich, C., Sidner, C.L. & Gertner, A. (2002, June). Collaborative Discourse Theory as a Foundation for Tutorial Dialogue. In *Proc. Sixth International Conference on Intelligent Tutoring Systems*, pp. 542-551, Springer.