Off Topic Conversation in Expert Tutoring: Waste of Time or Learning Opportunity?

Blair Lehman, Whitney Cade, & Andrew Olney
{balehman|wlcade|aolney}@memphis.edu
Institute for Intelligent Systems, University of Memphis, Memphis, TN 38152

Abstract. While many aspects of tutoring have been identified and studied, off topic conversation has been largely ignored. In this paper, off topic conversation during 50 hours of one-to-one expert tutoring sessions was analyzed. Two distinct methodologies (Dialogue Move occurrence and LIWC analysis) were used to determine the anatomy of off topic conversation. Both analyses revealed that the expected social talk occurred, but pedagogically-relevant talk emerged as well. These occurrences may reflect the discussion of more global pedagogical strategies. These findings suggest that off topic conversation may serve a useful purpose in tutoring and that further investigation is warranted.

1 Introduction

When it comes to tutoring, expert human tutors are widely acknowledged as the gold standard for producing learning gains. According to [1], accomplished tutors produce an effect size of 2.0 sigma (approximately 2 letter grades) when compared to learning gains found in a classroom scenario. Accomplished tutors also outperform tutors with domain knowledge but no pedagogical training [2] and Intelligent Tutoring Systems [ITSs; 3]. One assumes that expert tutors excel for a number of reasons. Expert tutors use pedagogical techniques that depart from those seen in previous work on novice tutors [4, 5]. It has also been hypothesized that expert tutors provide emotional support for their students [6] while increasing their feelings of competence [7]. The real question, unaddressed by current research, is this: how do expert tutors build the necessary rapport with students to produce high learning gains and boosted confidence?

Recent research proposes that this rapport building takes place during the non-pedagogical parts of a tutoring session [8]. While previous work on off-task behavior has shown a negative relationship with learning [9, 10], it is important to distinguish off-task behavior from off topic conversation. Off-task behavior involves such actions as gaming the system [9]. During off topic conversation, on the other hand, the student is still invested in the tutoring system and the divergence is still under the close supervision of the tutor. An investigation into a corpus of tutoring sessions collected by [8] suggests that rapport building occurs in these moments of "social talk." They contend that not only is social talk a considerable portion of their corpus (20%), but that it is repeatedly used for greetings, rapport building, social coordination, and general learning strategies. It is during social talk that the tutor will give "pep talks," potentially to increase both student motivation and self-confidence. They casually observed that those tutors who were the least effective in imparting knowledge also came up short in building positive rapport. Observations such as these deserve more systematic, follow-up attention. If social talk/off topic conversation is the source of rapport building, then it is certainly deserving of study. Rapport building helps to create good communication [11] and is viewed as very

important by students [12]. However, these observations were made in regards to a fairly contrived corpus; the tutors were graduate students, and the students were recruited as part of an experiment (though they were in a class on the domain topics). Additionally, these sessions occurred over the computer through an interface with video and chat capabilities. This raises an issue of ecological validity: do these observations generalize to natural tutoring sessions, and are they true of expert tutors as well?

To answer this question, an investigation into the practices of expert tutors in a natural tutoring setting is needed. The present study used a previously collected corpus of 50 expert tutoring sessions with ten expert tutors to investigate off topic conversation and rapport building. Though these sessions have been coded and examined in the past, off topic conversation in expert tutoring has been given little more than a cursory glance. However, we suspect that the interaction between student and tutor in these "off topic conversations" forms the basis of the student-tutor rapport, which will later allow the tutor to alternately encourage and critique the student without seeming disingenuous or harsh. We know that expert tutors give negative feedback where novice tutors never would [13, 14], but we also know that expert tutors help the student build confidence in themselves [7]. Perhaps the key to this paradoxical situation lies within the unstudied off topic conversation.

We approach this investigation using two distinct methods of exploring off topic conversation. The first of these methods seeks to establish a general layout of what is occurring during non-pedagogical modes through a manual coding scheme. This analysis will provide a broad picture of what is occurring during these discussions and reveal any potential pedagogical purpose to off topic conversation. We will also use the Linguistic Inquiry and Word Count tool (LIWC) [15] to get at the meaningful differences between Off Topic and Scaffolding. Both of these phases of tutoring are highly interactive, yet they greatly differ in purpose. While the purpose of Off Topic is assumed to be irrelevant conversations, Scaffolding has been shown to be the epitome of problem solving during these tutoring sessions [4]. By comparing these two different modes along linguistic feature dimensions, we can better establish the instructional and rapport-building components of Off Topic.

2 Expert Tutoring Corpus

In this study, we examined one-to-one, human-to-human expert tutoring sessions. Ten expert math and science tutors participated in 50 sessions in the current study. The tutors had to meet four criteria to be considered an expert: licensed to teach at the secondary level, five or more years of tutoring experience, employed by a professional tutoring agency, and highly recommended by school personnel. While one student was receiving tutoring in order to obtain a GED, the other student participants were in grades 7 to 12. All of the students were in academic trouble and actively sought out tutoring. Each student participated in a maximum of two tutorial sessions, while each tutor participated in between two and eight tutoring sessions.

Fifty, hour-long tutoring sessions were recorded with thirty-nine of the tutors' students. The tutoring locations were chosen by the tutor and the student. The experimenters were

merely observers of a natural tutoring session that would have taken place regardless of this study. Prior to any tutoring session, the tutors signed an informed consent and gave informed consent forms to the parents of prospective participant-students.

Each transcript has been coded using two coding schemes: the dialogue move and the dialogue mode coding scheme [4, 16, 17]. While a move is a smaller unit of meaningful speech (such as a word or short phrase), modes are longer, sustained, pedagogically-distinct phases of a tutoring session. Dialogue moves have been categorized into three pedagogically-relevant groups: Tutor Motivational Dialogue Moves, Tutor Pedagogical Dialogue Moves, and Student Dialogue Moves. There are eight dialogue modes in the coding scheme, but only four are of interest in this study: Scaffolding, Off Topic, Introduction, and Conclusion; [4]). These modes accounted for 48.58%, 4.35%, 2.60%, and1.85% of tutoring dialogue, respectively. For some of the analyses Introduction, Off Topic, and Conclusion are collapsed into one larger category of non-pedagogical modes (8.80% of tutoring sessions).

In preparation for the analysis of dialogue move distributions, non-pedagogical modes were isolated from the larger corpus. Proportional occurrence of student dialogue moves in non-pedagogical modes was then determined for each session. Proportions were based on the total number of dialogue moves in non-pedagogical modes in each session. Thus all instances of Introduction, Off Topic, and Conclusion in each session were combined for this analysis. Finally, the base rate proportional occurrence of each dialogue move within each session was determined to serve as a comparison.

Our next method of investigating off topic conversation involved the use of the Linguistic Inquiry and Word Count tool (LIWC), developed by [15]. In this study, we used LIWClite 7, which calculates the percentage of each document's words that fall into specific, predefined categories. Though this version of LIWC offers over 70 linguistic categories, only 16 were of interest in determining the nature of off topic conversation (see Table 3). These were selected because of their relevance to pedagogical and social dimensions of language. Only sessions including at least one instance of Scaffolding and Off Topic were included, leaving 30 sessions for this analysis. To prepare the transcripts for LIWC analysis, instances of Off Topic and Scaffolding were separated into two documents for each session. If a transcript had multiple types of each mode (e.g., three Scaffolding modes), the dialogue from those modes was compiled into one document. Those documents were then subdivided by speaker, making a maximum of four possible documents per session (Tutor-Scaffolding, Tutor-Off Topic, Student-Scaffolding, Student-Off Topic). Each document was capped at 1,000 words to control for verbosity (as Scaffolding is a frequently occurring, lengthy mode [4]). The resulting documents were then submitted to LIWClite 7 in order to find the means of every word category per document.

3 Results & Discussion

3.1 Comparison of dialogue move occurrence in off topic modes and base rate

For the dialogue move analyses, the non-pedagogical modes will simply be referred to as Off Topic. These proportions were then compared to the base rate occurrence for each dialogue move in each tutoring session. Base rate represents the average occurrence of each dialogue move throughout the entire tutoring session. Paired samples *t*-tests were conducted and significant differences were found. The means, standard deviations, *t*-values and effect sizes (Cohen's d) of each comparison can be found in Table 1 (tutor dialogue moves) and Table 2 (student dialogue moves).

Table 1. Comparison of tutor dialogue moves within off topic modes and base rate

	Off Topic		**Base Rate**		***t*-value**	**Cohen's *d***
Motivational	*M*	*SD*	*M*	*SD*	*df = 49*	
Positive Feedback	0.035	0.037	**0.074**	0.037	-6.86*	-1.05
Negative Feedback	0.003	0.006	**0.009**	0.006	-5.58*	-1.11
Humor	**0.018**	0.033	0.009	0.014	2.14*	0.348
Repetition	0.005	0.012	**0.016**	0.016	-4.67*	-0.814
Solidarity Statement	**0.005**	0.013	0.001	0.002	1.80†	0.335
General Motivational	**0.016**	0.026	0.004	0.005	3.30*	0.613
Off Topic	**0.268**	0.109	0.068	0.039	14.46*	2.45
Pedagogical						
Forced Choice	0 .000	0.000	**0.002**	0.003	-4.26*	-0.852
Hint	0.001	0.006	**0.013**	0.009	-8.20*	-1.65
Preview	**0.007**	0.019	0.003	0.004	1.79†	0.341
Prompt	0.001	0.004	**0.020**	0.021	-6.33*	-1.28
Pump	0.000	0.002	**0.020**	0.005	-4.89*	-0.965
Paraphrase	0.000	0.002	**0.003**	0.003	-3.98*	-0.824
Simplified Problem	0.006	0.012	**0.034**	0.016	-10.25*	-1.95
Example	0.001	0.007	**0.006**	0.008	-3.63*	-0.586
Counter Example	0.000	0.000	**0.000**	0.001	-3.75*	-0.750
Provide Correct Answer	0.000	0.002	**0.011**	0.008	-9.16*	-1.81
Direct Instruction	0.069	0.069	**0.182**	0.045	-12.54*	-1.93
Comprehension Gauging Question	0.034	0.043	0.032	0.023	0.330	0.058

†= $p < .1$; * = $p < .05$. $d \approx .2, .5, .8$ indicate small, medium, and large effects, respectively [18]. Significantly larger mean values are italicized.

As was expected, student and tutor off topic dialogue moves along with other socially-focused dialogue moves occurred more frequently during the non-pedagogical modes [8]. Consistent with this expectation, problem solving and other pedagogically-focused dialogue moves occur outside of non-pedagogical modes. In particular, the absence of tutors asking questions, students answering questions, and tutors giving feedback demonstrates that non-pedagogical modes are truly a separate, distinct time of the tutoring session.

Table 2. Comparison of student dialogue moves within off topic modes and base rate

	Off Topic		Base Rate		*t*-value	Cohen's *d*
Answer	*M*	*SD*	*M*	*SD*	*df = 49*	
Correct	0.012	0.023	**0.056**	0.032	-8.77*	-1.583
Partial	0.004	0.012	**0.024**	0.026	-5.72*	-0.966
Vague	0.008	0.016	**0.019**	0.015	-4.10*	-0.706
Error-Ridden	0.001	0.004	**0.011**	0.009	-7.41*	-1.493
None	0.001	0.005	**0.006**	0.009	-4.76*	-0.666
Question						
Social Coordination	**0.015**	0.026	0.007	0.009	2.79*	0.425
Common Ground	0.005	0.013	**0.024**	0.019	-6.22*	-1.186
Knowledge Deficit	0.005	0.009	**0.011**	0.010	-3.85*	-0.665
Metacognition						
Misconception	0.000	0.002	**0.004**	0.004	-4.89*	-0.977
Metacomment	**0.030**	0.042	0.020	0.015	2.01*	0.321
Action						
Think Aloud	0.001	0.003	**0.006**	0.010	-3.26*	-0.643
Read Aloud	0.000	0.000	**0.003**	0.013	-1.93†	-0.385
Work Silently	0.001	0.005	**0.015**	0.019	-5.19*	-1.006
Off Topic	**0.215**	0.117	0.059	0.038	10.65*	1.791

†= $p < .1$; * = $p < .05$.

Those occurrences which break from these general patterns are of particular interest. The only pedagogical move to occur more frequently in non-pedagogical modes was preview. This could indicate that non-pedagogical modes are being used as a transition between new topics or problems. Preview could potentially be occurring within Introduction, which can be thought of as a preview to the entire tutoring session. For student dialogue moves, the higher occurrence of metacomment shows that non-pedagogical modes contain discussions of the student's knowledge. Many metacomments are delivered in response to tutor comprehension-gauging questions (i.e., "Do you understand?" "Okay?"). The fact that comprehension-gauging questions are not significantly occurring during non-pedagogical modes suggests that student knowledge and comprehension is being discussed in a different context.

However, the strongest occurrences by far were still tutor (d = 2.45) and student (d = 1.79) off topic dialogue moves. Given that these dialogue moves serve as a catchall for any topic outside of the tutoring topic (e.g., algebra), it is difficult to truly determine what occurs during non-pedagogical modes. While it was casually observed that these dialogue moves ranged from after school activities to study strategies, the exact proportion of each is currently unknown. There are two options for expanding our analysis of off topic dialogue moves. One is to create a new coding scheme that makes finer distinctions in off-topic. The other option, which we discuss next, is to use a text analysis tool to look for text-level features that might show what's going on inside off topic

3.2 LIWC analysis of off topic modes

The comparisons between the Off Topic and Scaffolding modes along LIWC dimensions were done using a series of paired *t*-tests. The means (values in % of words), standard deviations, t-values and effect sizes (Cohen's d) of each comparison can be found in Table 3. The analysis was conducted on both tutor (T) and student (S) contributions during the tutoring sessions. These same comparisons were also made using the Wilcoxon's signed-rank test, as a normal distribution of scores cannot be assumed. However, those results very closely mirror the results of the paired *t*-test, and so only the paired *t*-test comparisons are presented here. A Bonferroni correction was not used in this analysis; as this is exploratory research that will be used to orient future research, the authors felt that a conservative correction would result in a loss of critical, if minor, information. In sum, the results here seem to indicate that every significant category difference favored the Off Topic mode, with the exception of when students use ACHIEVEMENT and FUTURE words.

Table 3. Occurrence of LIWC category words

LIWC Category	T/S	Off Topic		Scaffolding		*t*-value	Cohen's d
		M	*SD*	*M*	*SD*	*df = 29*	*(sig only)*
Social Process	T	**11.15**	3.09	7.75	1.66	6.272*	1.37
	S	**8.25**	4.73	4.87	2.43	4.148*	0.9
Positive Emotion	T	5.41	2.68	4.83	2.43	1.082	
	S	**6.54**	5.56	4.54	2.34	2.046*	0.47
Insight	T	**3.00**	1.40	2.06	0.67	3.480*	0.86
	S	2.75	1.75	1.62	0.68	3.029	
Tentative	T	**3.10**	1.70	1.91	0.64	3.874*	1.08
	S	**2.68**	2.19	1.60	0.80	2.562*	0.65
Work	T	**2.90**	2.81	1.10	1.00	3.289*	0.86
	S	2.70	2.70	2.09	3.19	0.801	
Achieve	T	1.02	1.14	0.95	0.55	0.313	
	S	0.52	0.72	**1.89**	1.98	3.682*	0.92
Leisure	T	0.78	2.74	0.23	0.20	1.115	
	S	**0.50**	0.95	0.15	0.32	1.859†	0.5
Home	T	**0.30**	0.68	0.04	0.08	2.013†	0.53
	S	0.24	0.89	0.01	0.03	1.412	
Nonfluencies	T	**1.51**	1.00	1.11	0.81	1.849†	0.44
	S	3.89	5.97	4.14	2.56	0.236	
Future	T	1.13	0.86	1.23	0.70	0.452	
	S	0.74	1.27	**1.35**	1.24	2.124*	0.49

† = $p < .1$; * = $p < .05$.

In general, it may be said that there is more to off topic conversation than simple socializing and "time wasting." Instead, there is a complex dynamic within the Off Topic mode, where the tutor and student are achieving a balance of work-related discussion and subtle socializing. Though off topic conversation seems to be an unlikely place for WORK

words to arise, the Off Topic mode contains significantly more WORK words than does Scaffolding. At first, this seems counterintuitive; however, this difference may be due to the way in which work is talked about in these two different modes. In Scaffolding, work may not be discussed on a superficial level, as this is where work is actually performed. Off Topic may be a place to discuss work on a superficial level, without content. The authors of LIWC list examples of WORK words being things like "class" and "graduate," so perhaps Off Topic conversation is a place where the student and tutor talk generally about schoolwork and homework. This could be supported by the significantly elevated use of HOME words in Off Topic as opposed to Scaffolding. Perhaps HOME is being discussed in the context of homework. Word categories that would hint at a more social use of HOME words, like FAMILY and FRIEND words, do not significantly differ from Off Topic to Scaffolding. In fact, those two categories make up less than 1% of the words in both modes. This suggests that HOME words are not being used in a social context, but rather, in work-related discussion. This is similar to the results of [8]; in that work, tutors engaged in "social talk" with their students, which involved discussing learning strategies. Our tutors may be doing something similar during their "social talk" (Off Topic). It may be the case that expert tutors use off topic conversation to discuss more general studying strategies.

This may also explain why TENTATIVE words and NONFLUENCIES occur significantly more in Off Topic than they do in Scaffolding on the part of the tutor. Rather than overtly stating problem solving and study suggestions, the tutor may use TENTATIVE words and NONFLUENCIES to lessen the face threatening nature of these suggestions. In one tutoring session, the tutor tells the student that, although the teacher assigned the odd problems for homework, she should work additional problems to get more practice. While this can be portrayed as a mere suggestion, it is alluding to the student's deficient abilities and need for further practice. Suggesting additional work like this may induce the tutor to use more words like "maybe," "perhaps," and "umm", given that barking orders is unlikely to lead to the completion of this additional work.

Although the LIWC results do suggest the presence of work-related discussion, there is an undeniable socio-emotional factor involved in off topic conversation. Generally, the Off Topic mode contains more POSITIVE EMOTION words than does Scaffolding; this aligns with work by [16], which found that the emotion "happiness" was much more likely to occur with tutor and student off topic conversation than other portions of the tutoring session. These off topic conversations, then, may be used as a sort of short "break" from the tutoring material that restores positive emotion and builds rapport between the tutor and student. This positive emotion and rapport building may act as a buffer against some of the direct, negative feedback that expert tutors give [17]. However, other affective LIWC categories like AFFECTIVE PROCESSES, NEGATIVE EMOTION words, and ANXIETY are not used in significantly different amounts between Off Topic and Scaffolding. This may be indicative of students' greater comfort in discussing positive emotions, as their negative emotions are likely tied to past and current academic struggles and failures. However, it may instead reflect that the purpose of Off Topic is not to discuss the emotional state of the student during learning. Which may also mean that off topic conversation does not necessarily include "pep talks", as [8] suggest. In addition, tutors do not use a larger amount of ACHIEVEMENT words Off Topic, suggesting

that they are not trying to overtly bolster students' feelings of confidence. Instead, rapport seems to be built in more subtle ways, such as by using more SOCIAL PROCESSES words like "we" and "us", and perhaps by using higher-level strategies of rapport building like humor and solidarity statements.

4 Conclusion

These two methodologies seem to converge and support our initial casual observations that off topic conversation is more than simply social talk or irrelevant ramblings. Off topic does not seem to be simply an "other" category. We feel that the evidence supports claims that off topic dialogue may serve motivational uses, to discuss more global pedagogy or study skills [8], build rapport [11], and in certain cases serve as a much needed mental break from tutoring.

Exploratory studies often are limited by the use of a single methodology. This study benefits from the use of two distinct approaches to investigating off topic conversation in tutoring. One approach utilized pre-existing coding schemes to determine whether the activities generally assigned to pedagogical conversation are occurring to any degree during off topic conversation. The second approach brings in a new and different analysis of the dialogue. This approach allows for a more objective look at the data, whereas results from the pre-existing coding scheme could be critiqued as simply an artifact of our coding methodology. Given that both approaches showed off topic conversation to be complex and multidimensional, this convergence gives support to further exploration. These results have given an improved depiction of what occurs during off topic conversation; however, we are still only able to speculate on its anatomy. Future research will reveal the true role and importance of off topic conversation in tutoring.

These findings have afforded a framework to begin future, more directed investigations. First, they have allowed us to determine whether further exploration is even worthwhile. These exploratory findings suggest that further investigation is, in fact, warranted. Second, the findings here can be used as the basis for future coding schemes. Whether a manual coding scheme or an automated methodology, such as probabilistic topic models [19], is employed, either can be used to determine the proportion of off topic conversation that is dedicated to global pedagogy, building rapport, social topics, and possibly even irrelevant ramblings. Through the accurate depiction of off topic conversation, its most advantageous features can be applied to the building of ITSs.

While implementing off topic conversation into ITSs under the current conception of discussing social topics seems peculiar, incorporating those pedagogical and rapport-building dimensions of off topic conversation would be a useful addition. So while an ITS may never form a deep, meaningful social bond with a student, it could help to increase learning in a broader scope than simply the present topic. ITSs such as MetaTutor already incorporate strategies similar to our proposed global pedagogy during learning [20-22]. By incorporating the global pedagogy of expert tutors as well as more local pedagogical strategies, ITSs can give greater aid to struggling students in over many disciplines. Further analysis will be needed to know the exact nature of off topic conversation and its potential usefulness in building ITS systems.

Acknowledgements. This research was supported by a grant awarded by the U. S. Office of Naval Research (N00014-05-1-0241) and the Institute of Education Sciences (R305A080594). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Office of Naval Research or the Institute of Education Sciences.

References

[1] Bloom, B.S. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 1984, 13, p. 4-16.

[2] Cohen, P.A., Kulik, J.A., & Kulik, C.C. Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, 1982, 19, p. 237-248.

[3] Corbett, A., Anderson, J., Graesser, A., Koedinger, K., & van Lehn, K. Third generation computer tutors: Learn from or ignore human tutors? *Proceedings of the 1999 Conference of Computer-Human Interaction*, 1999, p. 85-86.

[4] Cade, W.L., Copeland, J.L., Person, N.K., and D'Mello, S.K. Dialogue modes in expert tutoring. *Proceedings of the 9th International Conference on Intelligent Tutoring Systems*, 2008, p. 470-479.

[5] D'Mello, S.K., Person, N., & Lehman, B.A. Antecedent-consequent relationships and cyclical patterns between affective states and problem solving outcomes. *Proceedings of 14th International Conference on Artificial Intelligence In Education*, 2009, p. 57-64.

[6] Lepper, M.R., Woolverton, M., Mumme, D.L., & Gurtner, J. Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In Lajoie, S.P., & S.J. Derry, S.J. (Eds.), *Computers as cognitive tools*, 1993, p. 75-105. Hillsdale, NJ: Lawrence Erlbaum Associates.

[7] Lepper, M.R., & Chabay, R.W. *Socializing the intelligent tutor: Bringing empathy to computer tutors*, 1988. New York, NY: Springer-Verlag.

[8] Rosé, C.P., Kumar, R., Aleven, V., Robinson, A., & Wu, C. CycleTalk: Data driven design of support for simulation based learning. International *Journal of Artificial Intelligence in Education*, 2006, 16, 195-223

[9] Baker, R.S., Corbett, A.T., Koedinger, K.R., & Wagner, A.Z. Off-task behavior in the cognitive tutor classroom: When students "game the system". *In Proceedings of ACM CHI 2004: Computer-Human Interaction*, 2004, p. 383-390.

[10] Rowe, J., McQuiggan, S. Robison, J., & Lester, J. Off-task behavior in narrative-centered learning environments. *In Proceedings of the 14th International Conference on Artificial Intelligence and Education*, 2009, p. 99-106.

[11] Catt, S., Miller, D. & Schallenkamp, K. You are the key: Communicate for learning effectiveness. *Education*, 2007, 127(3), 369-377.

[12] Ramsden, P. Student learning and perceptions of the academic environment. *Higher Education*, 1979, 8, 411-428.

[13] Cromley, J.G., & Azevedo, R. What do reading tutors do? A naturalistic study of more and less experienced tutors in reading. *Discourse Processes*, 2005, 40(2), 83-113.

[14] Graesser, A.C., D'Mello, S.K., & Person, N.K. Meta-cognition, meta-communication, and meta-affect in tutoring. *Presented at the annual meeting of the American Educational Research Association*, 2009, San Diego, CA.

[15] Pennebaker, J.W., Francis, M.E., & Booth, R.J. *Linguistic Inquiry and Word Count (LIWC): LIWC2001*, 2001. Mahwah, NJ: Lawrence Erlbaum Associates.

[16] Lehman, B.A., Matthews, M., D'Mello, S.K., and Person, N. What are you feeling? Investigating student affective states during expert human tutoring sessions. *In Proceedings of the Ninth International Conference in Intelligent Tutoring Systems*, 2008, p. 50-59.

[17] Person, N.K., Lehman, B.A., & Ozbun, R. Pedagogical and motivational dialogue moves used by expert tutors. *Presented at the 17th Annual Meeting of the Society for Text and Discourse*, 2007, Glasgow, Scotland.

[18] Cohen, J. A power primer. *Psychological Bulletin*, 1992, 112(1), 155-159.

[19] Steyvers, M., & Griffiths, T. Probabilistic topic models. In Landauer, T., McNamara, D., Dennis, S., & Kintsch, W. (Eds.), Latent semantic analysis: A road to meaning, 2007. Mahwah: Erlbaum.

[20] Azevedo, R., Witherspoon, A.M., Graesser, A., McNamara, D., Rus, V., Cai, Z., et al. MetaTutor: An adaptive hypermedia system for training and fostering self-regulated learning about complex science topics. *Paper to be presented at a Symposium on ITSs with Agents at the Annual Meeting of the Society for Computers in Psychology*, 2008, Chicago, IL.

[21] Azevedo, R., Witherspoon, A., Graesser, A. McNamara, D., Chauncey, A., Siler, E., et al. (2009). MetaTutor: Analyzing self-regulated learning in a tutoring system for biology. In Dimitrova, V., Mizoguchi, R., du Boulay, B., & Graesser, A. (Eds.), *Building learning systems that care: From knowledge representation to affective modeling*, 2009, p. 635-637. Amsterdam: IOS Press.

[22] Azevedo, R., Witherspoon, A., Chauncey, A., Burkett, C., & Fike, A. MetaTutor: A MetaCognitive tool for enhancing self-regulated learning. *Paper presented at the Annual Meeting of the American Association for Artificial Intelligence, Symposium on Metacognitive and Cognitive Educational Systems*, 2009, Washington, DC.