# The “DeMAND” coding scheme: A “common language” for representing and analyzing student discourse

Gwendolyn E. CAMPBELL[a], Natalie B. STEINHAUSER[a], Myroslava O. DZIKOVSKA[b], Johanna D. MOORE[b], Charles B. CALLAWAY[b] & Elaine FARROW[b]
*[a]Naval Air Warfare Center Training Systems Division*
*[b]Human Communication Research Centre, University of Edinburgh*

**Abstract.** We propose that a set of five dimensions forms a foundation underlying a number of prevalent theoretical perspectives on learning. We show how student contributions to instructional dialogue can be reliably annotated with these dimensions. Finally, we provide preliminary validation evidence for our coding scheme and illustrate the potential value of such an approach to analyzing student behavior in tutorial dialogue.

**Keywords.** Learning theories, intelligent tutoring systems, coding scheme, tutorial dialogue

## Introduction

It is difficult to compare results across learning studies, because they are often designed to investigate different learning theories and use difficult coding strategies. To alleviate this problem we want establish a "common language" for coding tutorial dialogues.

First, we identified five prominent learning theories and decomposed them into a set of underlying dimensions (see Table 1). Next, we transformed those dimensions (**de**pth, **m**otivation, **a**ccuracy, **n**ovelty and **d**oubt into a coding system, which we labelled DeMAND (see Table 2).

Table 1: Theories and Underlying Dimensions.

| The Theories: Learning results from… | Can be Decomposed into these Underlying Dimensions |
|---|---|
| Applying cognitive effort [1] | Statements that are DEEP |
| Engaging in accountable talk [2] | Statements that are DEEP and ACCURATE |
| Constructing new knowledge [3,4] | Statements with NOVEL information |
| Autonomous activity [5] | Statements that are internally MOTIVATED |
| Experiencing an impasse [6] | Statements that are not ACCURATE or show DOUBT |

In the rest of this paper, we show that this system can be reliably applied to a corpus, provide preliminary evidence of its validity and illustrate its potential utility.

Table 2: DeMAND Codes.

| Code | Definition |
|---|---|
| De(pth) | Credited if: contained explicit signs of deep processing, such as explaining, justifying, providing evidence, asking questions to gain a better understanding, etc. |
| M(otivation) | All statements were coded. A statement was externally motivated if it was made in reaction to a "demand" of the lesson or tutor. A statement was internally motivated if it could not be traced to any external demand. |
| A(ccuracy) | All statements were coded. Possible values include: Correct & complete, correct but incomplete, partially correct with some errors and incorrect. |
| N(ovelty) | Credited if: contained information that has not already been explicitly stated (or confirmed) by the tutor, slides, or student. |
| D(oubt) | Credited if: contained explicit indicators of a lack of confidence, such as hedges. |

## 2. Current Study

We created a curriculum and a computer-based learning environment to teach basic concepts in direct current circuits. The curriculum took students approximately 4 hours to complete. The learning environment contained lesson materials (including didactic text, exercises and discussion questions), a circuit simulator, and a message window where the participant and tutor interacted. All of the student-tutor interactions occurred through typing.

We created three versions of these lessons (open-ended, closed-ended, and middle), which varied only in the format of the questions that the students were asked. Our objective was to impact the type of language produced by the participants, with the open-ended condition eliciting more deep statements, more accountable talk and more internally motivated statements than the closed-ended condition. If our system reflects these differences, this provides some positive evidence for its construct validity [7].

The corpus includes dialogues from each of thirty participants distributed across three experienced tutors and conditions. The entire corpus includes 8,085 dialogue turns taken by the student and tutor, and 56,133 tokens (words and punctuation).

## 3. Results

Following [8], our raters achieved kappa values of 0.64 ("moderate"), 0.81, 0.89 and 0.88 ("substantial") for depth, motivation, novelty and doubt, respectively. Accuracy was coded in real time by the tutors.

We conducted three one-way, between-subjects analyses of variance (ANOVAs) with lesson condition (open, middle and closed) as the independent variable for each, and the percentage of student utterances that were coded as containing evidence of (a) cognitive effort, (b) accountable talk or (c) internally motivated as the three dependent variables (DVs), respectively. In each case, the overall ANOVA was significant, and a Fischer's least significant difference (LSD) test indicated that the DV was larger in the open condition than in the closed condition (see Table 3).

We conducted a hierarchical multiple linear regression, with post-test score as the dependent variable and the pre-test score and the five learning theories (cognitive effort, internal motivation, constructivism, accountable talk, and experiencing impasses) as independent variables. The result was significant, $F(1,28) = 8.1$, $p=.008$, with $R^2$ = 0.224. The only significant predictor was impasses, which were negatively related to post-test score, $\beta$=-0.47, $t$=-2.85, $p$=.008.

Table 3: Statistics.

| ANOVA | LSD | Descriptive Statistics for Open Condition | Descriptive Statistics for Closed Condition |
|---|---|---|---|
| (a) DV = percentage of student utterances coded as containing evidence of cognitive effort | | | |
| $F(2,29) = 66.13, p<.01$ | $t(19) = 10.67, p<.01$ | $M = .45, SD = .07$ | $M = .13, SD = .07$ |
| (b) DV = percentage of student utterances coded as accountable talk | | | |
| $F(2,29) = 59.82, p<.01$ | $t(19) = 9.67, p<.01$ | $M = .41, SD = .07$ | $M = .11, SD = .07$ |
| (c) DV = percentage of student utterances coded as internally motivated | | | |
| $F(2,29) = 6.22, p = .006$ | $t(19) = 4.00, p = .003$ | $M = .07, SD = .04$ | $M = .02, SD = .02$ |

## 4. Summary

Without a common language, it can be difficult to compare results across different learning studies. We proposed that five prominent theories of learning could be represented by a series of underlying dimensions (some singly, others with combinations). We showed that these dimensions could be reliably assessed and that they respond as predicted to an instructional manipulation (an accepted approach for collecting evidence of construct validity [7]). Finally, we showed that this coding scheme allowed us to simultaneously evaluate the relative explanatory power of the different learning theories for our students' learning gains.

While additional work is required to refine and evaluate this approach, we believe that decomposing learning theories into a set of "least common denominators" as we have done here will facilitate a deeper understanding of the relationships between theories, allow us to more readily compare results across studies and support a systematic investigation into the relative impact on learning gains of different types of events, experiences and pedagogical methods.

## References


[1] J. W. Thomas & W. D. Rohwer Jr., Proficient autonomous learning: Problems and prospects. In M. Rabinowitz (Ed.), *Cognitive science foundations of instruction* (pp. 1-32). Lawrence Erlbaum Associates, Hillsdale, NJ, 1993.

[2] M. K. Wolf, A. C. Crosson & L. B. Resnick, Accountable talk in reading comprehension instruction. CSE Technical Report 670. University of Pittsburgh, January 2006.

[3] R. J. Osborne & M. C. Wittrock, Learning science: A generative process. *Science Education*, **67** (1983), 489-508.

[4] T. M. Duffy, & D. H. Jonassen, *Constructivism and the technology of instruction: A conversation*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1992.

[5] S. B. Nolen, Reasons for studying: Motivational orientations and study strategies, *Cognition and Instruction,* **5** (1988), 269-287.

[6] K. VanLehn, S. Siler, C. Murray, T. Yamauchi & W. B. Baggett, Why do only some events cause learning during human tutoring? *Cognition and Instruction,* **21** (2003), 209-249.

[7] D. T. Campbell & D. Fiske, Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin,* **56** (1959), 81-105.

[8] J. Carletta, A. Isard, S. Isard, J. Kowtko, G. Doherty-Sneddon & A. Anderson, The reliability of a dialogue structure coding scheme. *Computational Linguistics,* **23** (1997), 13-32.