The Relationship between Carelessness and Affect in a Cognitive Tutor

Maria Ofelia Clarissa Z. San Pedro[1], Ma. Mercedes T. Rodrigo[1], Ryan S. J. d. Baker[2],

[1] Ateneo de Manila University, Loyola Heights, Quezon City, Philippines
[2] Worcester Polytechnic Institute, Worcester, MA, United States
sweetsp@gmail.com, mrodrigo@ateneo.edu, rsbaker@wpi.edu

Abstract. We study the relationship between student carelessness and affect among high-school students using a Cognitive Tutor for Scatterplots, using a machine-learned detector of carelessness and field observations of student affect. In line with previous research, we say a student is careless when he/she makes a mistake performing a task that he/she already knows. This construct is also known as slipping. Somewhat non-intuitively, we find that students exhibiting high levels of engaged concentration slip frequently. These findings imply that a student who is engaged in a task may be overconfident, impulsive or hurried, leading to more careless errors. On the other hand, students who display confusion or boredom make fewer careless errors. Further analysis over time suggests that confused and bored students have lower learning overall. Therefore, these students' mistakes stem from a genuine lack of knowledge rather than carelessness. The use of two versions of the tutor in this study, with and without an Embodied Conversational Agent (ECA), shows no significant difference in terms of the relationship between carelessness and affect.

Keywords: Carelessness, Slips, Engaged Concentration, Confusion, Boredom, Cognitive Tutors.

1 Introduction

Within learning, some students become careless, working unconscientiously [13] and making unintended errors [7]. This can happen when an individual is overconfident in carrying out a task [7], or carries out a task in an impulsive or in a hurried manner [13]. Carelessness is a common behavior among students, even among high-performing students [6, 7]. Carelessness not only reduces short-term performance, it can even lead to poorer overall academic performance among first year college students [17].

To study student carelessness in a fine-grained fashion within educational software, Baker, Corbett, and Aleven [2] created an automated detector of slips which operationalizes Clements' [7] definition of careless errors – errors committed on skills that a student knows (though slips may also occur for other reasons, such as shallow knowledge [3]). This model, which predicts carelessness from the features of student actions within the software (e.g. performance history on the associated skill, input type, if action was a help request, time taken, etc.), proved to be robust enough to successfully transfer without degradation in predictive accuracy to intelligent tutor software with different design features, and between schools in different countries [18]. Recent research used this detector to study individual differences in carelessness, finding that students with strong mastery or performance goals tend to be more careless than students who lack strong goal orientation [12].

It has been hypothesized that a major factor driving carelessness is student affect, which accords with results showing links between affect and achievement goals [14]. Student affect has been shown to be associated with other disengaged student behaviors. For instance, boredom [5] was found to be associated with gaming the system, a disengaged behavior associated with poorer learning, where students systematically take advantage of regularities in the software's feedback and help to obtain answers and advance within the tutoring curriculum [4]. Within this paper, we study the relationship between student affect and carelessness, using two versions of Cognitive Tutor software for Scatterplot generation and interpretation [4], differing in the presence or absence of an Embodied Conversational Agent (ECA) which responds to student gaming behavior [4]. We detect the incidence of careless errors by analyzing interaction logs from high school students, based on previous work at modeling this construct [e.g. 2, 18]. The detectors created and validated for this tutor in [18] are applied to student data to obtain average estimates of what proportions of each student's errors were careless errors [18]. We combine these estimates with the proportions of each affective state exhibited by each student, gathered using quantitative field observations [16]. We then assess which affective states are associated with careless errors through correlational analysis, also examining whether these relationships change over time. Studying this relationship between carelessness and affect may lead to a deeper understanding of carelessness as a student behavior, thus allowing tutor design to better respond when the student is careless.

2 Methods

The study was conducted in a large, urban high school in Quezon City, Philippines (PH). The school had 5,368 students and 216 teachers [1]. The school's community is relatively poor where about half of the students' parents was unemployed and 70% of the households earned PhP10,000 (US$230.00) per month or less. Data were gathered from 126 first year high school students who used a Cognitive Tutor unit on Scatterplots [4]. The students, aged 12 to 14, used the tutor to solve as many problems as they could within 80 minutes. Data on student carelessness were distilled from the logs generated from tutor usage, while data on affective states from quantitative field observations used in [5, 16]. Students had not explicitly covered these topics in class prior to the study, and students viewed conceptual instructions via a PowerPoint presentation with voiceover and some simple animations before using the tutor. Each student took a nearly isomorphic pre-test and post-test, counterbalanced across conditions. An analysis of learning gains themselves is outside the scope of this paper.

2.1 The Scatterplot Tutor

Within the Scatterplot Tutor, the learner is given a problem scenario. He/she is also provided with data that he/she needs to plot in order to arrive at the solution. He/she is asked to identify the variables that each axis will represent. He/she must then provide an appropriate scale for each axis. He/she has to label the values of each variable along the axis and plot each of the points of the data set. Finally, he/she interprets the resultant graphs. The Scatterplot tutor provides contextual hints to guide the learner, feedback on correctness, and messages for errors. The skills of the learner is monitored and displayed through skill bars that depict his/her mastery of skills.

Sixty-four of the participants (Scooter group) were randomly assigned to use a version of the tutor with an embodied conversational agent, "Scooter the Tutor". Scooter was designed to both reduce the incentive to prevent gaming the system and to help students learn the material that they were avoiding by gaming, while affecting non-gaming students as minimally as possible. Scooter displays happiness and gives positive message when students do not game (regardless of the correctness of their answers), but shows dissatisfaction when students game, and provides supplementary exercises to help them learn material bypassed by gaming. The remaining 62 participants (NoScooter group) used a version of the Scatterplot Tutor without the conversational agent. The number of students assigned to the conditions in this study was unbalanced because of data gathering schedule disruptions caused by inclement weather.

2.2 Affect Observations

Each student's prevalence of each affective state was assessed using quantitative field observations. Each student was observed 24 times by a pair of trained expert coders, with an interval of 180 seconds between observations lasting 20 seconds. Observations were conducted using peripheral vision where the observer appeared to be looking at another student, so that the student being observed would not know that he/she is the one being observed, in order to reduce the degree to which affect is altered by the observation process. As in past research using this method, only the first observed affect was recorded, to minimize bias. The coding scheme included: boredom, confusion, delight, engaged concentration (flow in [10]), frustration, surprise and neutral. The observers' inter-rater reliability was found to be acceptable at Cohen's kappa (κ) of 0.54, a moderate level of agreement between raters. Cohen's kappa measures the proportion of agreement between two observers, with adjustments for the proportion that would be expected to occur by chance [9]. More details on the quantitative field observation method, including examples from the coding manual, can be found in [5].

2.3 Carelessness Detection

The incidence of carelessness within the Cognitive Tutor was traced with a model designed to assess "slips" [cf. 2], treated as an operationalization of carelessness in accordance with prior theory discussed in Section 1. This model, termed the Contextual Slip model, contextually estimates the probability that a specific action indicates a slip/carelessness, where a student knows a skill but answers incorrectly on a problem step which requires that skill. In this model, the estimates of the slip probability is dynamic, and depends on contextual aspects of the action, such as speed of action and history of the student's help-seeking from the tutor. The Contextual Slip model has been shown to be a statistically significant predictor of post-test performance measuring learning from a Cognitive Tutor for Genetics, even after controlling for assessment of each student's knowledge within the software [3].

The Contextual Slip model is based on Bayesian Knowledge Tracing (BKT) [7], a model used to estimate a student's latent knowledge based on his/her observable performance. In its original articulation, BKT is used within Cognitive Tutors to infer student knowledge by continually updating the estimated probability a student knows a skill every time the student gives a first response to a problem step regardless whether the response is correct or not. It uses four parameters – two learning parameters L_O (initial probability of knowing each skill) and T (probability of learning the skill at each opportunity to make use of a skill), together with two performance parameters G (probability that the student will give a correct answer despite not knowing a skill) and S (probability that the student will give an incorrect answer despite knowing the skill) – for each skill (estimated from data information in each skill). These parameters are invariant across the entire context of using the tutor. Using Bayesian analysis, BKT re-calculates the probability that the student knew the skill before the response (at time n-1), using the information from the response, then accounts for the possibility that the student learned the skill during the problem step, such that [8]:

$$P(L_n \mid Action_n) = P(L_{n-1} \mid Action_n) + ((1 - P(L_{n-1} \mid Action_n)) * P(T)) \ . \qquad \textbf{(1)}$$

From this model, two detectors of Contextual Slip (one for the NoScooter group and one for the Scooter group) were produced from log files generated by the Cognitive Tutor software. A set of 26 transaction features identical to the set previously used in [2] to develop detectors of Contextual Slip for tutors in other domains was extracted and derived from the logs for each problem step. With the information from the logs, parameters needed for a baseline BKT model were fitted by employing brute-force search [cf. 3]. From this baseline model, estimates of whether the student knew the skill at each step were derived and used to label incorrect actions with the probability that the actions were slips, based on the student performance on successive opportunities to apply the rule [2]. As in [2], Bayesian equations were utilized in computing training labels for the Slip probabilities for each student action (A) at time N, using future information (two actions afterwards – N+1, N+2), in order to infer the true probability that a student's correctness at time N was due to knowing the skill, or due to a slip . The probability that the student knew the skill at time N can be calculated, given information about the actions at time N+1 and N+2 ($A_{N+1,N+2}$), and the other parameters of the Bayesian Knowledge Tracing model.

Models for Contextual Slip were then produced through Linear Regression to create models that could predict contextual slip without using data from the future. These models were used in the analyses in the remainder of the paper. The exact models used in these analyses (e.g. all parameters, weights and functional form) are given in [18].

3 Results

We assessed student carelessness using the slip detector from [18], as discussed in Section 2.3. For this analysis, we assessed carelessness of each student in each group (e.g. Scooter and NoScooter), by taking the average probability of carelessness (slip estimates) on each incorrect action the student made, as in [3]. The overall mean carelessness for students in the NoScooter environment was 0.09. The Scooter group had an overall mean carelessness of 0.14. The difference between the two conditions was significant, $t(124) = 8.38$, two-tailed $p < 0.001$. From the students' action logs, we also looked at each student's percentage of steps which were incorrect (careless and non-careless actions combined). For the NoScooter group, 46.55% of actions were incorrect. Although the Scooter group's percentage appeared slightly higher at 48.12%, the difference between groups was not significant, $t(124) = -0.49$, two-tailed $p = 0.63$. When we examined the incidence of affective states between the two conditions, we found no significant differences for any affective state; the largest difference found between conditions was for Engaged Concentration (Scooter students displayed this affect 37.17% of the time; NoScooter students displayed this affect 43.45% of the time; $t(124) = 1.52$, two-tailed $p = 0.13$).

We then studied the relationship between each student's proportion of carelessness and their proportion of each affective state, with correlational analyses conducted in SSPS, shown in Table 1. The results were somewhat surprising. Carelessness was negatively correlated with boredom in both interfaces, $r=-0.29$, $p=0.02$ for NoScooter group; $r=-0.41$, $p=0.001$ for Scooter group. Confusion was also negatively correlated with carelessness, $r=-0.31$, $p=0.01$ for NoScooter group; $r=-0.21$, $p=0.09$ for Scooter group. Carelessness was positively correlated with engaged concentration, $r=0.47$, $p<0.001$ for NoScooter group; $r=0.49$, $p<0.001$ for Scooter group. This indicated that more engaged students were careless more often, while less engaged students were careless less often. Carelessness and frustration were not significantly correlated in either condition, $r=-0.06$, $p=0.66$ for NoScooter group; $r=-0.04$, $p=0.76$ for Scooter group, which may be due to frustration's very rare occurrence among students. Hence, we focus the following analysis and discussion on carelessness's correlations with boredom, confusion and engaged concentration.

Table 1. Correlations of Carelessness and Affective State for Entire Tutor Usage. Statistically Significant Correlations in Bold.

	NoScooter group	Scooter group
Careless – Boredom	**-0.29 (p = 0.02)**	**-0.41 (p = 0.001)**
Careless – Confusion	**-0.31 (p = 0.01)**	-0.21 (p = 0.09)
Careless – Engaged Concentration	**+0.47 (p < 0.001)**	**+0.49 (p < 0.001)**
Careless – Frustration	-0.06 (p = 0.66)	-0.04 (p = 0.76)

Next, we examined the changes in carelessness and affect over time to see if there were significant differences in the relationship and occurrence of carelessness and affect as the student used the tutor. We did this by separating the observed affect during the student's tutor usage into two halves: the first 12 affect observations for each student (across both raters), and the remaining 12 observations per student. To split the estimates of carelessness into halves, each student's actions within the tutor was split by overall time (e.g. the split was by total time elapsed, rather than number of actions). We computed the average carelessness for incorrect actions during each half of each student's tutor usage. Figure 1 shows the student affect percentages and average estimates of carelessness in each half of the students' tutor usage in both NoScooter and Scooter groups.

Across time, the changes in the incidence of different affective states were consistent across conditions. In both conditions, students exhibited less confusion in the second half of their usage than the first half, $F(1,124) = 61.76$, $p < 0.001$. They also exhibited more engaged concentration over time, $F(1,124) = 40.26$, $p < 0.001$, and exhibited more boredom over time, $F(1,124) = 11.89$, $p = 0.001$. The changes in the proportion of these affective states over time were not significantly different between the NoScooter and Scooter conditions – for boredom $F(1,124) = 0.01$, $p = 0.93$, for confusion $F(1,124) = 0.79$, $p = 0.38$, and for engaged concentration $F(1,124) = 1.82$, $p = 0.18$.

Carelessness, on the other hand, showed a non-significant reduction over time in the NoScooter condition, $t(61) = 1.35$, two-tailed $p = 0.18$, and an non-significant increase in Scooter condition, $t(63) = -0.98$, two-tailed $p = 0.33$. This difference between conditions was statistically significant, $F(1,124) = 81.84$, $p < 0.001$. Overall, student errors (whether careless or not) significantly decreased over time in both conditions, $F(1,124) = 30.47$, $p < 0.001$. The changes over time were also not significantly different between the NoScooter and Scooter conditions, $F(1,124) = 0.31$, $p = 0.58$.

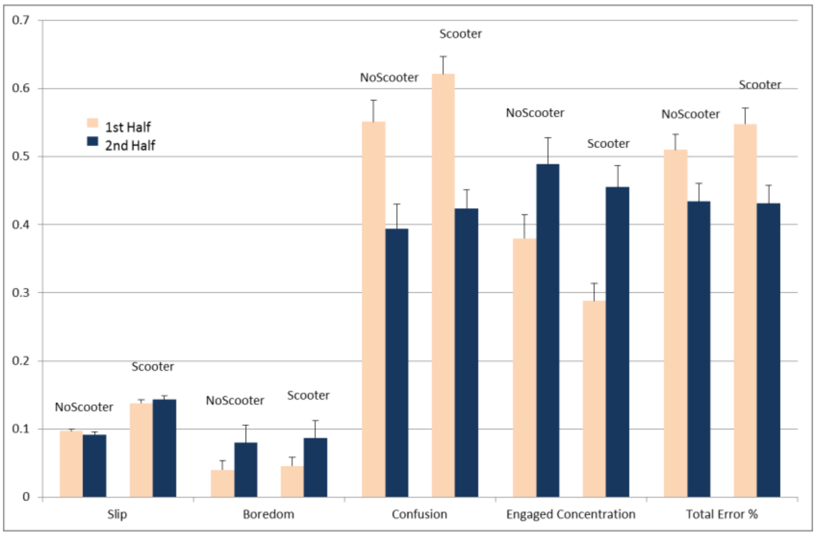

Fig. 1. Carelessness, Affect, Error, By Time of Tutor Usage

Table 2 shows the correlations between carelessness and the three most common affective states. During the first half of the tutor usage period, the correlations between carelessness and boredom trended negative in both groups, but were not significant, $r = -0.08$, $p = 0.53$ for NoScooter group; $r = -0.10$, $p = 0.43$ for Scooter group. During the second half, the correlations became significantly negative, $r = -0.26$, $p < 0.001$ for NoScooter group; $r = -0.44$, $p < 0.001$ for Scooter group. Steiger's Z-test [19], a standard test for comparing multiple correlations with no overlapping variables within a single population, was used to determine whether the correlation between carelessness and boredom was significantly different between the two time periods. The apparent difference in correlation was not significant for the NoScooter group, $Z = 1.28$, $p = 0.20$. However, the difference was statistically significant for the Scooter group, $Z = 2.31$, $p = 0.02$.

During the first half of the tutor usage period, the correlations between carelessness and confusion trended negative in both groups, but were not significant, $r = -0.14$, $p = 0.28$ for NoScooter group; $r = -0.15$, $p = 0.24$ for Scooter group. During the second half, the correlations became significantly negative, $r = -0.28$, $p = 0.03$ for NoScooter group; $r = -0.36$, $p=0.004$ for Scooter group. However, the difference in correlation between the time periods was not statistically significant for either group, $Z = 0.89$, $p = 0.37$ for NoScooter group; $Z = 1.36$, $p = 0.17$ for Scooter group.

During the first half of the tutor usage period, the correlations between carelessness and engaged concentration trended positive in both groups, but were not significant, $r = 0.17$, $p = 0.20$ for NoScooter group; $r = 0.23$, $p = 0.07$ for Scooter group. During the second half, the correlations became significantly positive, $r = 0.46$, $p < 0.001$ for NoScooter group; $r = 0.64$, $p < 0.001$ for Scooter group. The correlation between carelessness and engaged concentration was significantly different between the first half of tutor usage and the second half of tutor usage, for both groups, $Z = -2.00$, $p = 0.05$ for NoScooter group; $Z = -3.23$, $p = 0.001$ for Scooter group.

Finally, the correlations between carelessness and students' total error percentage were statistically significantly negative during the first half of the tutor usage period in both groups, $r = -0.71$, $p < 0.001$ for NoScooter group; $r = -0.78$, $p < 0.001$ for Scooter group. During the second half, the correlations remained significantly negative, $r = -0.85$, $p < 0.001$ for NoScooter group; $r = -0.85$, $p < 0.001$ for Scooter group. These correlations between the time periods were statistically significant for the NoScooter group, $Z = 2.19$, $p = 0.03$ but not statistically significant for the Scooter group, $Z = 1.31$, $p = 0.19$.

Table 2. Correlations Between Carelessness and Affect By Time Within Tutor Usage. Significant Correlations in Bold.

	Control	Experimental
Careless-Boredom (1st Half)	-0.08 (p = 0. 53)	-0.10 (p = 0.43)
Careless-Confusion (1st Half)	-0.14 (p = 0.28)	-0.15 (p = 0.24)
Careless-Engaged Concentration (1st Half)	+0.17 (p = 0.20)	+0.23 (p = 0.07)
Careless-Total Error Percentage (1st Half)	**-0.71 (p < 0.001)**	**-0.78 (p < 0.001)**
Careless-Boredom (2nd Half)	**-0.26 (p = 0.04)**	**-0.44 (p < 0.001)**
Careless-Confusion (2nd Half)	**-0.28 (p = 0.03)**	**-0.36 (p = 0.004)**
Careless-Engaged Concentration (2nd Half)	**+0.46 (p < 0.001)**	**+0.64 (p < 0.001)**
Careless-Total Error Percentage (2nd Half)	**-0.85 (p < 0.001)**	**-0.85 (p < 0.001)**

4 Discussion and Conclusion

In this paper, we studied the relationship between student carelessness and affective states within a Cognitive Tutor for Scatterplots, building off prior work in carelessness detection [2, 18] and observations of affective states [16]. Student carelessness, as represented by the presence of slips or careless errors, was estimated and detected per student using the Contextual Slip and Guess model by Baker et al. [2], implemented for this specific tutor by [18]. The detector inferred carelessness from the features of the individual student action [cf. 18]. The detector's assessment of each student's carelessness was then studied in conjunction with that student's proportion of each affective state.

Overall tutor usage showed that the more confused or bored a student is, the less likely errors are to be careless, likely due to the higher occurrence of non-careless errors stemming from poorer learning. The significant negative correlation between carelessness and boredom may be explained, at least in part, by earlier findings showing a negative relationship between boredom and strategies that lead to learning, constructs such as self-regulation and elaboration strategies, as well as effort [15]. In addition, boredom has previously been found to be correlated with poorer learning [10, 15]. As such, errors committed by students who were frequently bored may have stemmed from a lack of knowledge, which came in turn from a lack of interest in pursuing knowledge. This implies that these errors were not committed out of carelessness, where the student already knows the skill. The significant negative relationship between confusion and carelessness may be similarly explained by the students' lack of knowledge and his/her awareness of what he/she does not understand. Lack of knowledge is explicitly an implication that an error is not careless, within the model of carelessness used here [e.g. 2, 18]. On the other hand, the more a student displays the affective state of engaged concentration, the more likely he/she is to display carelessness, possibly due to overconfidence.

Examination of slip and affective state behavior over time and their respective correlations supported these assertions. The correlation between confusion and carelessness became significantly negative as students used the tutor more. However, confusion (as well as total errors – whether careless or not) decreased over time. Hence, the students who were struggling most and remained confused even after using the tutor for a substantial amount of time were less likely to become careless. The correlation between carelessness and engaged concentration became positive and significant during the second half of tutor usage, in both conditions. One possible interpretation for this is that a student who is engaged most of the time may succeed, become over-confident, and then commit careless errors. It is interesting to note that the connections between affect and carelessness were so much stronger in the second half of tutor usage than the first, across affective states. It is possible in this case that errors made during initial tutor usage in the Scooter condition were "honest" mistakes, born out of a genuine lack of knowledge and not out of carelessness. In the early part of their tutor usage, all students may have been learning the task, making it harder for an error to be careless.

Finally, these findings relate in interesting ways to recent work about carelessness on goal orientation [12]. In this work, carelessness was found to be positively correlated with academic efficacy and negatively correlated with disruptive behavior and self-presentation of low achievement [12]. In addition, carelessness was found to be higher among students with mastery or performance goals than among students manifesting neither type of goal orientation. These goal orientations were found to be related to learning and academic performance, such that mastery and performance goals lead to academic competence [11]. Hence, it unexpectedly seems that carelessness is more frequent among students displaying goal orientation and affect that are associated with positive learning.

The findings presented here also accord with results in Clements' [7] classic work on careless errors. Clements found that mathematically competent and confident children tend to make a greater proportion of careless errors than other children. In this current study, carelessness has a positive relationship with affect that has positive impacts on learning (engaged concentration [e.g. 10]) and a negative relationship to affect associated with negative effects on learning (boredom [e.g. 10, 15]) as well as affect stemming from difficulty learning the material (confusion).

In summary, these results illustrate the key role of affect in student carelessness, and suggest that adaptive responses to carelessness should take probable student affect into account.

Acknowledgments. This research was supported by the Philippines Department of Science and Technology Philippine Council for Advanced Science and Technology Research and Development under the project "Development of Affect-Sensitive Interfaces", and by the Pittsburgh Science of Learning Center (National Science Foundation) via grant "Toward a Decade of PSLC Research", award number SBE-0836012. We thank Mrs. Carmela Oracion, Jenilyn Agapito, Ivan Jacob Pesigan, Ma. Concepcion Repalam, Salvador Reyes, Ramon Rodriguez, the Ateneo Center for Educational Development, the Department of Information Systems and Computer Science of the Ateneo de Manila University and the faculty, staff, and students of Ramon Magsaysay Cubao High School for their support in this project.

References

1. Ateneo Center for Educational Development. "Ramon Magsaysay Cubao High School: School Profile Report," Available from the Ateneo Center for Educational Development, Ateneo de Manila University, Loyola Heights, Quezon City, Philippines (2009)
2. Baker, R.S.J.d., Corbett, A.T., Aleven, V.: More Accurate Student Modeling through Contextual Estimation of Slip and Guess Probabilities in Bayesian Knowledge Tracing. In: 9th International Conference on Intelligent Tutoring Systems, 406–415 (2008)
3. Baker, R.S.J.d., Corbett, A.T., Gowda, S.M., Wagner, A.Z., MacLaren, B.M., Kauffman, L.R., Mitchell, A.P., Giguere, S.: Contextual Slip and Prediction of Student Performance after Use of an Intelligent Tutor. In: 18th Annual Conference on User Modeling, Adaptation, and Personalization (2008)
4. Baker, R. S. d. J., Corbett, A. T., Koedinger, K. R., Evenson, S. E., Roll, I., Wagner, A.Z., Naim, M., Raspat, J., Baker, D. J., Beck, J.: Adapting to When Students Game an Intelligent Tutoring System. In: 8th International Conference on Intelligent Tutoring Systems, 392--401 (2006)
5. Baker, R.S.J.d., D'Mello, S.K., Rodrigo, M.M.T., Graesser, A.C.: Better to Be Frustrated than Bored: The Incidence, Persistence, and Impact of Learners' Cognitive-Affective States

during Interactions with Three Different Computer-Based Learning Environments. International Journal of Human-Computer Studies, 68 (4), 223–241 (2010)

6. Clements, M.A.: Analysing Children's Errors on Written Mathematical Tasks. Educational Studies in Mathematics, 1--21 (1982)
7. Clements, M.A.: Careless Errors Made by Sixth-Grade Children on Written Mathematical Tasks. Journal for Research in Mathematics Education , 13(2), 136–144 (1982)
8. Corbett, A.T., Anderson, J.R.: Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge. In: User Modeling and User-Adapted Interaction, 4, 253–278 (1995)
9. Cohen, Jacob. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement, 20, 37–46 (1960).
10. Craig, S. D., Graesser, A. C., Sullins, J., & Gholson, B.: Affect and Learning: An Exploratory Look into the Role of Affect in Learning with AutoTutor. Journal of Educational Media, 29(3), 241–250 (2004)
11. Harackiewicz, J.M., Barron, K.E., Tauer, J.M., Elliot, A.J. (2002) Predicting Success in College: A Longitudinal Study of Achievement Goals and Ability Measures as Predictors of Interest and Performance from Freshman Year through Graduation. Journal of Educational Psychology, 94 (3), 562–575.
12. Hershkovitz, A., Wixon, M., Baker, R.S.J.d., Gober, J., Sao Pedro, M.: Carelessness and Goal Orientation in a Science Microworld. In: 15th International Conference on Artificial Intelligence in Education, 6738, 462–465 (2011)
13. Maydeu-Olivares, A., D'Zurilla, T.J.: A Factor Analysis of the Social Problem-Solving Inventory Using Polychoric Correlations. European Journal of Psychological Assessment, 11 (2), 98–107 (1995)
14. Pekrun, R., Elliot, A.J., Maier, M.A.: Achievement Goals and Discrete Achievement Emotions: A Theoretical Model and Prospective Test. Journal of Educational Psychology, 98 (3), 583–597 (2006)
15. Pekrun, R., Goetz, T., Daniels, L.M., Stupnisky, R.H., & Perry, R.P.: Boredom in Achievement Settings: Exploring Control-Value Antecedents and Performance Outcomes of a Neglected Emotion. Journal of Educational Psychology, 102(2), 531–549 (2010)
16. Rodrigo, M.M.T., Baker, R.S.J.d., Nabos, J.Q.: The Relationships between Sequences of Affective States and Learner Achievements. In: 18th International Conference on Computers in Education (2010)
17. Rodriguez-Fornells, A., Maydeu-Olivares, A.: Impulsive/Careless Problem Solving Style as Predictor of Subsequent Academic Achievement. Personality and Individual Differences, 28, 639–645 (2000)
18. San Pedro, M.O.C.Z., Baker, R.S.J.d., Rodrigo, M.M.T.: Detecting Carelessness through Contextual Estimation of Slip Probabilities among Students Using an Intelligent Tutor for Mathematics. In: 15th International Conference on Artificial Intelligence in Education, 6738, 304–311 (2011)
19. Steiger, J. H.: Tests for Comparing Elements of a Correlation Matrix. Psychological Bulletin, 87, 245–251 (1980)