

# The Dynamics of Affective Transitions in Simulation Problem-Solving Environments

Ryan S.J.d. Baker<sup>1</sup>, Ma. Mercedes T. Rodrigo<sup>2</sup>, Ulises E. Xolocotzin<sup>1</sup>

<sup>1</sup>Learning Sciences Research Institute, University of Nottingham, Nottingham, UK

<sup>2</sup>Department of Computer Science, Ateneo de Manila University, Quezon City, Philippines  
ryan@educationaldatamining.org, mrodrigo@ateneo.edu, lpxux@nottingham.ac.uk

**Abstract.** We analyze the antecedents of affective states in a simulation problem-solving environment, The Incredible Machine: Even More Contraptions, through quantitative field observations of high school students in the Philippines using that system. We investigate the transitions between affective states over time, finding that several affective states, including flow, boredom, and frustration, but not surprise, tend to persist over for relatively long periods of time. We also investigate how students' usage choices influence their later affect, finding that gaming the system leads to reduced confusion but increased boredom.

## 1 Introduction

In recent years, researchers of intelligent tutoring systems and interactive learning environments have investigated ways in which these systems can be made to recognize and respond to differences in students' affect. Researchers in these areas have developed systems which attempt to recognize differences in student affect, using a variety of types of data, from biometric sensors [1,6], to posture data [15], to interaction patterns in system logs [11]. Work towards developing systems that can reliably detect affect has sparked progress in developing systems that can respond to differences in affect. For instance, pedagogical agents have been designed with social intelligence that takes account of a learner's emotional state [13], in order to guide the learner to an optimal emotional state for learning [4]. The goal of learning environments that respond to differences in affect is to influence and improve each student's affect, and through doing so boost students' learning gains and enhance their overall learning experience.

However, until recently, there has been relatively little consideration of the natural dynamics and shifts in students' affect, in learning environments that do not explicitly attempt to monitor and alter affect. Understanding affective dynamics in non-affective learning environments will be useful to researchers in many fashions. First of all, it will help us set goals for the design of affective learning environments. For instance, the elimination of common affective transitions, where both the start and end affective states are considered negative (what D'Mello et al [10] refer to as vicious cycles), may be as important as creating positive affective transitions. In addition, if we know

which transitions between negative and positive affective states are reasonably common, we may more easily be able to improve students' affect by strengthening naturally occurring beneficial affective transitions, as opposed to attempting to create transitions which seldom occur naturally.

Second, understanding and modeling affective dynamics will provide a baseline to use in understanding the impact of systems that attempt to influence affect. If a common affective transition disappears or a previously unseen transition develops in an affective learning environment, it will be easier to infer how the environment has influenced students' affect.

Third, a model of affective dynamics will provide evidence on the base rate of an affective state in a given situation beyond just the overall frequency of that affective state. Understanding affective dynamics may therefore make it possible to develop more successful detectors of affect.

Hence, modeling and understanding affective dynamics may lead to the development of more accurate and useful affective learning environments.

However, there has only been limited research, thus far, into affective dynamics. In one early paper on this subject, Guastello and his colleagues [12] studied the dynamics of a single affective state, flow [8], over time. Two more recent papers have investigated specific aspects of affective dynamics. D'Mello, Taylor, and Graesser [10] studied the transition between affective states in an intelligent tutoring system, AutoTutor. They computed the likelihood that learners would transition between a set of affective states thought to be relevant to learning: boredom, flow, confusion, frustration, delight, and surprise. They found that learners experiencing negative affective states such as boredom and frustration were likely to remain in these states. They were not likely to transition out of them and into more positive states such as flow or delight, nor to potentially more positive states such as confusion or surprise.

A second recent paper, by Rodrigo et al [16], examined the relationship between these same affective states and a set of potential usage choices drawn from [3]: on-task, on-task conversation, off-task conversation, off-task solitary, inactive, and gaming the system, focusing on how affective states influence later usage choices. Baker, et. al [3] define gaming the system as "attempting to succeed in an educational environment by exploiting properties of the system rather than by learning the material and trying to use that knowledge to answer correctly." This study, conducted within a simulation problem-solving environment, The Incredible Machine, found that boredom, confusion, and the affective state coded as neutral were antecedents to gaming. Frustration, flow, and delight were not found to be antecedents to gaming the system.

The findings from [10] and [16] lead to further questions: Are the affective state transitions found in D'Mello et al [10] generalizable? Are they particular to intelligent tutoring systems, or do they hold true for other types of interactive learning environments? And, do usage choices influence affective states to the same extent that affective states influence usage choices?

In this paper, we study these questions within the context of a simulation problem-solving environment, The Incredible Machine. We will consider data relevant to these questions, and discuss the implications for the design of affective learning environments.

## 2 Study Methods

We studied affective dynamics within a high school mathematics class in a private school in urban Manila, in the Philippines. Student ages ranged from 14 to 19, with an average age of 16. Thirty-six students participated in this study (17 female, 19 male).

Each student used *The Incredible Machine: Even More Contraptions* [18] (shown in Figure 1), a simulation environment where the user completes a series of logical “Rube Goldberg” puzzles. In each puzzle, the student has a pre-selected set of objects to use, such as scissors, ropes, and pulleys, electrical generators, and animals. The student must combine these objects in order to accomplish a pre-defined goal, such as lighting a candle or making a mouse run. If a student is stuck, he or she can ask for a hint; hint messages display where items should be located in a correct solution to the current problem (but do not show which item should be placed in each location).

Each student used *The Incredible Machine* for ten minutes, and each student’s behavior and affect was observed several times as he or she used *The Incredible Machine*. The observations were conducted using a method which incorporated aspects of Baker et al’s [3] quantitative field observations of student behavior categories, and Craig et al’s [7] laboratory observations of affect. The observations were carried out by a team of six observers, working in pairs. The observers were Masters students in Education or Computer Science, and all but one had prior teaching experience. Observations were conducted according to guide that gave examples of actions, utterances, facial expressions, or body language that would imply a behavior or an affective state, and practiced the coding categories in another school prior to this study.

As in Baker et al, each observation lasted twenty seconds, and was conducted using peripheral vision, i.e. observers stood diagonally behind or in front of the student being observed and avoided looking at the student directly, in order to make it less clear exactly when an observation was occurring. If two distinct behaviors were seen during an observation, only the first behavior observed was coded, and any behavior by a student other than the student currently being observed was not coded.

During the laboratory sessions in which the data was gathered, it was not possible for the entire class to use the software at the same time, due to the size of the school computer laboratory; hence, students used the software in groups of nine (one student per computer), during their class time. Each pair of observers was assigned to three students and alternated between them. Since each observation lasted twenty seconds, each student was observed once per minute.

Observing students more frequently than in [3,7] made it possible to directly analyze the relationship between a student’s affective state at a given time and their usage choices shortly thereafter.

Within an observation, each observer coded both the student’s behavior and affective state, using coding schemes developed in prior research. The observers trained for the task through a series of pre-observation discussions on the meaning of the usage and affective categories.

The usage categories coded were adapted from [3], and are as follows:

1. **On-task** – working within *The Incredible Machine*

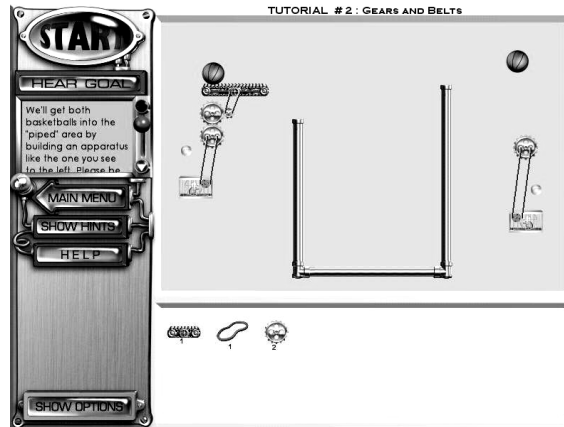


Figure 1. A screen shot from The Incredible Machine: Even More Contraptions.

2. **On-task conversation** – talking to the teacher or another student about The Incredible Machine, or its puzzles
3. **Off-task conversation** – talking about any other topic
4. **Off-task solitary behavior** –behavior that did not involve The Incredible Machine or another person (such as reading a magazine or surfing the web)
5. **Inactivity** – instead of interacting with other students or the software, the student stares into space or puts his/her head down on the desk.
6. **Gaming the System** – sustained and/or systematic guessing, such arranging objects haphazardly or trying an object in every conceivable place. Also, repeatedly and rapidly requesting help in order to iterate to a solution.

The affective categories coded were drawn from [9,10]. Since many behaviors can correspond to an emotion, the observers looked for students' gestures, verbalizations, and other types of expressions rather than attempting to explicitly define each category. The categories coded were:

1. **Boredom** – behaviors such as slouching, and resting the chin on his/her palm; statements such as “Can we do something else?” and “This is boring!”
2. **Confusion** – behaviors such as scratching his/her head, repeatedly looking at the same interface elements; statements such as “Why didn’t it work?”
3. **Delight** – behaviors such as clapping hands or laughing with pleasure; statements such as “Yes!” or “I got it!”
4. **Surprise** – behaviors such as jerking back suddenly or gasping; statements such as “Huh?” or “Oh, no!”

5. **Frustration** – behaviors such as banging on the keyboard or pulling at his/her hair; statements such as “This is annoying!” or “What’s going on?!?”
6. **Flow** – complete immersion and focus upon the system [cf. 8]; behaviors such as leaning towards the computer or mouthing solutions to him/herself while solving a problem
7. The **Neutral** state, which was coded when the student did not appear to be displaying any of the affective states above, or the student’s affect could not be determined for certain.

Some of these affective categories may not be mutually exclusive (such as frustration and confusion), though others clearly are (delight and frustration). For tractability, however, the observers only coded one affective state per observation.

Past research has suggested that brief observations can be reliable indicators of a student’s affective state, whether carried out live [9] or by watching screen-capture videos [11]. 706 observations were collected, for an average of 19.6 observations per student. Inter-rater reliability was acceptably high across all observations — Cohen’s [4]  $\kappa=0.71$  for usage observations,  $\kappa=0.63$  for observations of affective state.

### 3 Prior Research with this Data Set

Data from this study was previously used in order to analyze how affect influences students’ behavior [16]. In this section, we will briefly recapitulate the results reported in that paper, in order to inform the analyses we present within the current paper.

Within this data, the most common affective state was flow, coded in 61% of the observations. The dominance of the flow state was similar to results seen in prior studies of affect in students using intelligent tutoring systems [7,10]. The second most common category was confusion, observed 11% of the time. Boredom (7%), frustration (7%), delight (6%), and the neutral state (5%) were each seen in a small but definite proportion of the observations. Boredom was less common than in previous work studying affect in intelligent tutoring systems [7,10], but frustration was more common. Surprise was the rarest category, but was still observed (3%).

The two most common behaviors observed were working on-task with the software (80% of observations), and talking on-task (9% of observations). Gaming the system was the third most common category of behavior, observed 8% of the time. Off-task conversation and off-task solitary behavior were quite rare, occurring 0.5% and 0.3% of the time — this frequency is much lower than the frequency of off-task behavior in intelligent-tutor classrooms [3] but comparable to the frequency of off-task behavior among students playing non-educational action games [20].

Gaming the system was strongly associated with boredom and confusion — a student who was bored or confused was significantly more likely to game the system both at the time he or she was bored/confused, and one minute afterwards. Frustration

significantly co-occurred with gaming, but was not associated with later gaming. The neutral state was associated with gaming one minute later, but did not co-occur with gaming. Delight never co-occurred with gaming or preceded gaming.

## 4 The Antecedents of Affective States

In this section, we will analyze how affect and usage choices influence a student's later affect, specifically looking at how given affective states and usage choices influence the probability of a student being in a specific affective state one minute later.

In conducting these analyses, it is important to take into account the base rates of each affective category. Flow was the dominant category within our observations; hence, flow is likely to be the most common affective state that follows *any* other affective state. In order to appropriately account for the base rate of each affective category in assessing how likely a transition is, we adopt D'Mello et al's [10] transition likelihood metric,  $L$ , which is statistically equivalent to Cohen's [5]  $\kappa$ . D'Mello et al's  $L$  gives the probability that a transition between two affective states will occur, given the base frequency of the destination state, and is computed:

$$L = \frac{\Pr(NEXT | PREV) - \Pr(NEXT)}{(1 - \Pr(Next))}$$

$L$  is scaled between 1 and  $-\infty$ . A value of 1 means that the transition will always occur; a value of 0 means that the transition's likelihood is exactly what it would be given only the base frequency of the destination state. Values above 0 signify that the transition is more likely than it could be expected to be given only the base frequency of the destination state, and values under 0 signify that the transition is less likely than it could be expected to be given only the base frequency of the destination state.

For a given transition, we calculate a value for  $L$  for each student, and then calculate the mean and standard error across students. We can then determine if a given transition is significantly more likely than chance (0), given the base frequency of the next state, using the two-tailed t-test for one sample.

### 4.1 Transitions Between Affective States

In this section, we will examine how a student's affective state at a given time influences their affective state at a later time, within the Incredible Machine. Within the data, four transitions were significantly or marginally significantly ( $p < 0.10$ ) more likely than chance, and five transitions were significantly or marginally significantly less likely than chance. Ten transitions did not occur with sufficient frequency for it to be possible to compute a standard error value, and thus will be excluded from formal statistical analysis (however, it can therefore be inferred that these transitions are quite rare). With seven affective states, 49 transitions are theoretically possible. Nine

of the 39 transitions with sufficient data were significant or marginally significant; by chance, one would expect 3.9 transitions to be significant or marginally significant. The probability of 9 of 39 transitions being significant or marginally significant is 0.014 (computed by Monte Carlo simulation [14], 100,000 runs), so it appears quite unlikely that the pattern of results is due to chance.

One very clear pattern that emerged from the data is that affective states are persistent — if a student is in a given affective state at a given time, he or she is likely to be in that state one minute later as well. This pattern is true of the transition boredom → boredom (Mean  $L$  = 0.26, SE = 0.11,  $t(7)$ =2.27,  $p$ =0.06), confusion → confusion (Mean  $L$  = 0.10, SE = 0.06,  $t(20)$ =1.68,  $p$ =0.11), delight → delight (Mean  $L$  = 0.10, SE = 0.06,  $t(14)$ =1.71,  $p$ =0.11), flow → flow (Mean  $L$  = 0.20, SE = 0.12,  $t(33)$ =1.75,  $p$ =0.09), and neutral → neutral (Mean  $L$  = 0.41, SE = 0.09,  $t(5)$ =4.59,  $p$ =0.01). It is not clearly true of the transition frustrated → frustrated (Mean  $L$  = 0.12, SE = 0.08,  $t(12)$ =1.56,  $p$ =0.14), and appears not to be true of the transition surprised → surprised (Mean  $L$  = 0.03, SE = 0.06,  $t(7)$ =0.54,  $p$ =0.61). Interestingly, D’Mello et al [10] found a similar degree of persistence and for more or less the same affective states, studying affective transitions within an intelligent tutoring system. Both our results and D’Mello et al’s results found persistent transitions for boredom → boredom, flow → flow, confusion → confusion, and delight → delight (at the edge of marginal significance in both studies), and did not find significant persistent transitions for frustrated → frustrated and surprised → surprised.

Beyond the state-to-itself transitions, six relationships were significant or marginally significant. The neutral state was significantly more likely than chance to transition to boredom (Mean  $L$  = 0.23, SE = 0.10,  $t(5)$ = 2.44,  $p$ =0.06), but significantly less likely than chance to transition to flow (Mean  $L$  = -1.22, SE = 0.37,

	BOR	CON	DEL	FLO	FRU	NEU	SUR
BOR	0.26 (0.11)	0.00 (0.13)	-0.03 (0.03)	-0.58 (0.42)	0.01 (0.06)	0.03 (0.08)	
CON	0.03 (0.05)	0.10 (0.06)	0.00 (0.03)	-0.23 (0.23)	-0.03 (0.02)	-0.03 (0.02)	0.03 (0.03)
DEL		-0.02 (0.08)	0.10 (0.06)	0.33 (0.23)			-0.01 (0.02)
FLO	-0.05 (0.02)	-0.04 (0.02)	-0.03 (0.01)	0.20 (0.12)	0.03 (0.03)	0.01 (0.03)	0.00 (0.01)
FRU	0.00 (0.05)	0.15 (0.11)	-0.05 (0.01)	-0.29 (0.32)	0.12 (0.08)		
NEU	0.23 (0.10)	-0.02 (0.07)		-1.22 (0.37)	-0.03 (0.04)	0.41 (0.09)	
SUR			-0.02 (0.04)	0.09 (0.38)	0.09 (0.12)	0.06 (0.12)	0.03 (0.06)

**Table 1.** The transitions between affective states. Horizontal rows represent previous affective states, and vertical columns represent affective states one minute later. The first number in each cell is the mean value of D’Mello’s  $L$  across students, the number in parentheses is the standard error. Cells with insufficient sample size are left blank (but can be inferred to be quite rare). Statistically significant relationships are in dark grey. Marginally significant relationships are in light grey.

$t(5) = -3.29, p = 0.02$ ). Flow was significantly less likely than chance to transition to boredom (Mean  $L = -0.05, SE = 0.02, t(33) = -3.10, p < 0.01$ ), confusion (Mean  $L = -0.04, SE = 0.02, t(33) = -1.86, p = 0.07$ ), or delight (Mean  $L = -0.03, SE = 0.01, t(33) = -2.18, p = 0.04$ ). Frustration was significantly less likely than chance to transition to delight (Mean  $L = -0.05, SE = 0.01, t(12) = -6.50, p < 0.001$ ). None of these relationships were significant in D’Mello et al’s study of affective transitions within an intelligent tutor. The complete pattern of transitions is shown in Table 1.

## 4.2 Usage Choices as an Antecedent to Affect

In this section, we will analyze how a student’s choice of how to use a learning environment at a given time influences their affective state at a later time. In our earlier work [16], we found that certain affective states are antecedents to certain usage choices. Specifically, boredom, confusion, and the neutral state served as antecedents to the choice to game the system, and that delight and flow are negative antecedents to the choice to game the system (i.e. a student in flow or delight is less likely to game one minute later). In the analysis presented here, we ask: Does a student’s current usage choice influence his or her affect, the next time it is observed?

Within the data from our study on students using The Incredible Machine, one antecedent relationship was significantly more likely than chance, and three antecedent relationships were significantly or marginally significantly less likely than chance. Sixteen antecedent relationships did not occur with sufficient frequency for it to be possible to compute a standard error value, and thus will be excluded from formal statistical analysis (though these relationships can be inferred to be quite rare). With 7 affective states and 6 usage categories, 42 antecedent relationships are theoretically possible. Four of the 26 antecedent relationships with sufficient data were significant or marginally significant; by chance, one would expect 2.6 antecedent relationships to be significant or marginally significant. The probability of 4 of 26 antecedent relationships being significant or marginally significant is 0.28 (computed by Monte Carlo simulation [14], 100,000 runs), so it appears that the pattern of results could be due to chance. However, it is still possible to trust that an individual antecedent relationship is significant if that antecedent relationship is still significant after applying a Bonferroni procedure [17], which takes the number of tests run into account. In this case, we find that  $p$  must be less than 0.0018 for the test to be truly significant. Antecedent relationships which are not significant, given a Bonferroni procedure, must be considered suggestive rather than statistically significant.

As mentioned, four antecedent relationships are statistically significant or marginally statistically significant. Two involve gaming the system. First, a student who games the system is significantly less likely to be confused a moment later; this result is significant even after the Bonferroni procedure is applied (Mean  $L = -0.11, SE = 0.02, t(11) = -6.37, p < 0.001$ ). Since confusion is an antecedent to gaming behavior, this suggests that a student who is confused games the system to get past a confusing problem and that the gaming strategy generally succeeds in bypassing the confusing problem, enabling a transition out of the confused affective state.

A student who games the system also appears to be more likely to be bored a moment later, though this second result does not hold after applying the Bonferroni



procedure (Mean  $L = 0.26$ , SE = 0.12,  $t(11) = 2.18$ ,  $p = 0.05$ ). If this relationship is confirmed by later studies, it will have interesting implications. Since boredom is also an antecedent to gaming behavior, this result would suggest that a student who is bored games the system to alleviate their boredom, but that this strategy backfires, because the student's chance of being bored a minute after gaming is actually increased relative to the general incidence of the affective state.

It is worth noting that a student who games the system does not appear to be more likely to be frustrated a moment later (Mean  $L = 0.005$ , SE = 0.05,  $t(11) = 0.12$ ,  $p = 0.91$ ). In our earlier work, we found that frustration is also not an antecedent to gaming behavior, but that gaming behavior and frustration co-occur more than would be expected by chance. If frustration neither proceeds nor follows gaming, but co-occurs with gaming, it may be the case that a student who attempts to game the system but does not immediately succeed becomes frustrated until he or she succeeds in gaming the system, and that his or her frustration ends as soon as the attempt to game the system is successful.

Two other antecedent relationships are marginally statistically significant — both involve students who are on-task, working with the system. On-task behavior is marginally significantly negatively associated with a reduction in both frustration (Mean  $L = -0.026$ , SE = 0.014,  $t(34) = -1.85$ ,  $p = 0.07$ ) and delight (Mean  $L = -0.020$ , SE = 0.012,  $t(34) = -1.68$ ,  $p = 0.10$ ). Neither of these results hold after applying the Bonferroni procedure. The complete pattern of transitions is shown in Table 2.

	BOR	CON	DEL	FLO	FRU	NEU	SUR
GAMING THE SYSTEM	0.26 (0.12)	-0.11 (0.02)	-0.02 (0.04)	-0.47 (0.29)	0.01 (0.05)	0.09 (0.08)	
STARING INTO SPACE	0.03 (0.11)	-0.05 (0.08)	0.15 (0.21)	-0.90 (0.50)	0.21 (0.21)	0.05 (0.11)	
SOLITARY OFF-TASK							
TALKING OFF-TASK							
ON-TASK	0.00 (0.02)	0.00 (0.03)	-0.02 (0.01)	0.06 (0.12)	-0.03 (0.01)	0.01 (0.03)	0.00 (0.01)
ON-TASK CONVERS.	-0.03 (0.05)	-0.07 (0.05)	0.10 (0.07)	-0.14 (0.24)	0.04 (0.06)	-0.01 (0.03)	0.02 (0.04)

**Table 2.** Usage choices as antecedents to affective states. Horizontal rows represent usage choices, and vertical columns represent affective states one minute later. The first number in each cell is the mean value of D'Mello's  $L$  across students, the number in parentheses is the standard error. Cells with insufficient sample size are left blank (but can be inferred to be quite rare). Statistically significant relationships are in dark grey. Marginally significant relationships are in light grey.

## 5 Discussion and Conclusions

In this exploratory study, we have investigated the antecedents of affective states, both in terms of how affective states precede other affective states, and in terms of how usage choices influence later affective states, within the context of the use of a simulation-based learning environment, *The Incredible Machine: Even More Contraptions*. The first of these issues, how affective states influence later affective states, was previously studied by D’Mello and his colleagues [10], in the context of a very different type of learning environment, an intelligent tutoring system, and using quite different experimental methods. Their main finding was that if a student is in any affective state at a given time, they are likely to be in the same affective state a minute later. We replicated this finding over the majority of the affective states studied: flow, boredom, confusion, frustration, and delight (both studies found that  $\text{delight} \rightarrow \text{delight}$  was almost significant). At the same time, both our study and D’Mello et al’s study found that being surprised at a given time is *not* associated with being surprised a moment later. This result is not counter-intuitive — how long can someone genuinely be surprised? — but this result suggests that it is probably mistaken to conceptualize surprise as being part of the same category of affective states as flow, boredom, confusion, frustration, and delight. Instead, perhaps surprise should be conceptualized as a “transient affective state”, as distinguished from other “durable affective states”.

An interesting area for future research will be to determine if the other five affective states are all persistent to the same degree — if some affective states predominate for 20 minutes in a row whereas others tend to dissipate after 3 or 4 minutes, it will suggest that some affective states may actually be “moods” [cf. 17; personal communication, Sidney D’Mello], which last for considerably longer than other types of affective states. The data used in this paper was neither lengthy enough nor did it have a large enough sample to answer this question; addressing this question will be an important focus of future research in the field. It will be interesting, as well, to determine if the distinctions between moods and affective states apply across multiple types of learning environments, as the distinction between surprise and other affective states did.

The second major issue investigated in this paper, how usage choices within interactive learning environments influence later affect, has not to our knowledge been explicitly addressed in prior research, though the converse question, how affective states influence usage choices, was studied in [16]. Our results in this paper found that relatively few usage choices appeared to influence later affect. This forms an interesting contrast to the pattern of results in [16], where almost every affective state significantly influenced whether a student would later game the system. The combined pattern of results suggests that affect and usage choices may form a mostly one-way relationship — affect appears to influence usage much more than usage influences affect. However, we should obtain data similar to the data obtained in this paper for a system where off-task behavior is a more common choice (such as intelligent tutors [cf. 3]) before concluding that this pattern is generally true; off-task behavior may have stronger implications for future affect than the usage choices common in the *Incredible Machine*.

However, the primary relationships between usage choice and affect found in this study provide insight as to the implications of students' usage choices. Given existing knowledge that both boredom and confusion lead to the choice to game the system [16], it is interesting that gaming the system is positively associated with future boredom but negatively associated with future confusion. This result suggests that gaming the system can be an effective strategy for alleviating confusion (albeit a strategy that carries with it the cost of significantly poorer learning [3]), but not an effective strategy for alleviating boredom. To the extent that students game the system for different reasons, and with different results, a system's responses to gaming behavior should likely vary depending on why the student is gaming.

If we acknowledge that gaming the system does relieve students' confusion, perhaps we can develop systems to take advantage of the engagement with the task that confusion implies, and that attempt to respond to gaming by helping students to find more constructive ways to alleviate their confusion (such as, for instance, supplementary exercises [2]). On the other hand, since gaming the system does not successfully alleviate boredom, perhaps a system faced with a bored student should attempt to prevent gaming behavior entirely by taking immediate measures to alleviate the student's boredom. Since such a system may address boredom more effectively than the student could by gaming, this approach could completely eliminate bored students' incentive to game the system.

Overall, by studying how affect changes over time, and how usage choices and affect influence one another, we will be able to come to a full and rich picture of learners' experiences as they use interactive learning environments. This knowledge will be essential to the project of developing intelligent affective environments which can assess differences in students' affect and behavior and can respond to those differences in effective and constructive fashions.

**Acknowledgements.** We would like to thank Maria Carminda Lagud, Sheryl Ann Lim, Alexis Macapanpan, Sheila Pascua, Jerry Santillano, Leima Sevilla, Jessica Sugay, Sinath Tep, and Norma Jean Viehland for their assistance in organizing and conducting the studies reported here. We would also like to thank Sidney D'Mello for a very valuable conversation and suggestions. The work presented in this paper was funded in part by research fellowships from the Learning Sciences Research Institute, at the University of Nottingham.

## References

1. Amershi, S., Conati, C., Maclaren, H. Using feature selection and unsupervised clustering to identify affective expressions in educational games. *Proceedings of the Workshop "Motivational and Affective Issues in ITS"*. In conjunction with ITS2006, 8th International Conference on Intelligent Tutoring Systems (2006), Jhongli, Taiwan.
2. Baker, R.S.J.d., Corbett, A.T., Koedinger, K.R., Evenson, E., Roll, I., Wagner, A.Z., Naim, M., Raspat, J., Baker, D.J., Beck, J. Adapting to When Students Game an Intelligent Tutoring System. *Proceedings of the 8th International Conference on Intelligent Tutoring Systems* (2006), 392-401.

3. Baker, R.S., Corbett, A.T., Koedinger, K.R., and Wagner, A.Z. Off-Task Behavior in the Cognitive Tutor Classroom: When Students “Game the System”. *Proceedings of ACM CHI 2004: Computer-Human Interaction*, 383-390.
4. Chaffar, S. Frasson, C. Using and Emotional Intelligent Agent to Improve the Learner’s Performance. In “*Proceedings of the Workshop on Social and Emotional Intelligence in Learning Environments*”. In conjunction with Intelligent Tutoring Systems (2004), Maceio, Brazil.
5. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20 (1960), 37-46.
6. Conati, C., Chabbal, R. and Maclaren, H. A study on using biometric sensors for detecting user emotions in educational games. In *Proceedings of the Workshop “Assessing and Adapting to User Attitude and Affects: Why, When and How?”*. In conjunction with UM ’03, 9<sup>th</sup> International Conference on User Modeling (2003), Pittsburgh, PA, U.S.A.
7. Craig, S.D., Graesser, A.C., Sullins, J., Gholsen, B. Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media*, 29 (3) (2004), 241-250
8. Csikszentmihalyi, M. *Flow: The Psychology of Optimal Experience* (1990). New York: Harper and Row.
9. D’Mello, S. K., Craig, S. D., Witherspoon, A., McDaniel, B., Graesser, A. Integrating affect sensors in an intelligent tutoring system. In “Affective Interactions: The Computer in the Affective Loop Workshop” In conjunction with *International conference on Intelligent User Interfaces* (2005), 7-13.
10. D’Mello, S., Taylor, R.S., Graesser, A. Monitoring Affective Trajectories during Complex Learning. *Proceedings of the 29<sup>th</sup> Annual Meeting of the Cognitive Science Society* (in press).
11. de Vicente, A., Pain, H. Informing the detection of the students’ motivational state: an empirical study. *Proceedings of the 6th International Conference on Intelligent Tutoring Systems* (2002), 933-943.
12. Guastello, S.J., Johnson, E.A., Rieke, M.L. (1999) Nonlinear Dynamics of Motivational Flow. *Nonlinear dynamics, psychology, and life sciences*, 3 (3), 259-274.
13. Johnson, W.L., Kole, S., Shaw, E., Pain, H. Socially Intelligent Learner-Agent Interaction Tactics, *Proceedings of Artificial Intelligence in Education: Shaping the Future of Learning through Intelligent Technologies* (2003), 431-433.
14. Metropolis, N., Ulam, S. The Monte Carlo Method. *Journal of the American Statistical Association*, 44 (1949), 335.
15. Mota, S. & Picard, R. W. Automated posture analysis for detecting learner's interest level. Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction, CVPR HCI, (2003).
16. Rodrigo, M., Baker, R., Lagud, M., Lim, S., Macapanpan, A., Pascua, S., Santillano, J., Sevilla, R., Sugay, J., Tep, S., Viehland, N. Affect and Usage Choices in Simulation Problem-Solving Environments. *Proceedings of the 13<sup>th</sup> International Conference on Artificial Intelligence in Education* (in press). Available online via Google Scholar.
17. Rosenthal, R., Rosnow, R.L. *Essentials of Behavioral Research: Methods and Data Analysis: Second Edition* (1991). Boston: McGraw-Hill.
18. Sierra Online, Inc. (2001) *The Incredible Machine: Even More Contraptions*.
19. Thayer, R.E. *The biopsychology of mood and arousal* (1989). New York, NY: Oxford University Press.
20. Ziemek, T. Two-D or not Two-D: gender implications of visual cognition in electronic games. *Proceedings of the 2006 Symposium on Interactive 3D graphics and games* (2006), 183-190.