Button Theory: A Taxonomy of Student-Teacher Communication for Interface Design in Computer-Based Learning Environments*

Menachem Jona, Benjamin Bell, and Lawrence Birnbaum

Northwestern University
The Institute for the Learning Sciences
Evanston, IL 60201

Abstract

This paper introduces Button Theory whose two principle goals are, first, to provide a taxonomy of the ways that students might usefully interact with and control a computer-based teacher, and second, to provide a natural mechanism by which they may exercise that control. We have developed a small but comprehensive set of messages that students would find it useful to convey to a teacher during a tutorial interaction. and have associated each message with a button presented iconically on the computer screen We describe our experience with the use of Button Theory in a prototype computer-based teaching system, and demonstrate how, even with rather simple mechanisms, this framework enables surprisingly rich interactions.

One of the more serious problems facing education today is that students lack control over the instruction they receive (see, *e.g.*, Schank & Jona, 1991). When students feel that their learning needs are not being attended to, the result is boredom, frustration, and a dislike for learning in general. This is especially true in computer-based learning environments, where technological limitations, aggravated in some cases by poor design decisions, have historically put the student in an extremely passive role (see, *e.g.*, Papert, 1980). This paper introduces *Button Theory*, developed in order to allow students using computer-based learning environments to have as much control as possible over what they see, hear, and learn.

The two principle goals of Button Theory are, first, to provide a taxonomy of the ways that students might usefully interact with and control a computer-based teacher, and second, to provide a natural mechanism by which they may exercise that control. We have developed a small but comprehensive set of messages that students would find it useful to convey to a teacher during a tutorial interaction. These messages are organized into three groups, **feelings**, **questions**, and **control**, reflecting the three natural domains of discourse in a computer-based learning environment: The student may wish to talk about himself (**feelings**), about the domain or task of instruction (**questions**), or about the teaching itself (**control**). The **feelings** group contains messages about how the student is feeling about the lesson, allowing expressions of boredom, excitement, or disbelief. The **questions** group, as its name implies, contains various questions a student is likely to ask (*e.g.* Why?, What's the point?, etc.). The third group, **control**, allows the student to change the progression of the lesson to suit his or her needs, speeding up when the material is too easy, backing up when it is too difficult, or getting more detail about something interesting.

As the name of the theory indicates, rather than using natural language, the mechanism we provide to enable students to express these messages is a set of buttons presented iconically on the computer screen, each button corresponding to one message. The user can "press" any of these buttons using a mouse-pointer. This "Button Pad" can be thought of as analogous to the remote-control clicker of a TV or VCR, except that instead of controlling the volume or the channel, it allows the student to manage the instruction he or she is receiving.[1]

As described above, the types of communications that might take place between a student and a computer teacher can be divided into the three categories of **feelings**, **questions**, and **control**. These three categories are based on the three naturally-occurring domains of discourse that exist in any interaction between a student and a computer-based teacher, namely (a) the student, (b) the teacher/program, and (c) the domain and/or task being taught (see Figure 1).

The kinds of communications that fall into each of the three categories can be characterized as follows:

- **Feelings** – This category concerns the student's internal state. By "internal state" we mean the states of the student's emotional and belief systems. These are influenced by, among other things, various attributes of the domain and/or task, the teaching program, and the idiosyncrasies of the individual student. The student may feel bored or excited by the domain, may find his or her beliefs challenged, or may be frustrated by his or her inability to understand the material. Each student will also have different reactions to different teaching styles and subject matter.
- **Questions** –This category concerns the domain and/or task being taught. Communicative acts in this group involve efforts by the student to comprehend the subject

* This research was supported in part by the Defense Advanced Research Projects Agency, monitored by the Air Force Office of Scientific Research under contract F49620-88-C-0058 and the Office of Naval Research under contract N00014-90-J-4117, by the Office of Naval Research under contract N00014-J-1987, and by the Air Force Office of Scientific Research under contract AFOSR-89-0493. The Institute for the Learning Sciences was established in 1989 with the support of Andersen Consulting, part of The Arthur Andersen Worldwide Organization. The Institute receives additional support from Ameritech, an Institute Partner, and from IBM.

[1] See Schank & Jona (1991) for a fuller discussion of this analogy and its implications.

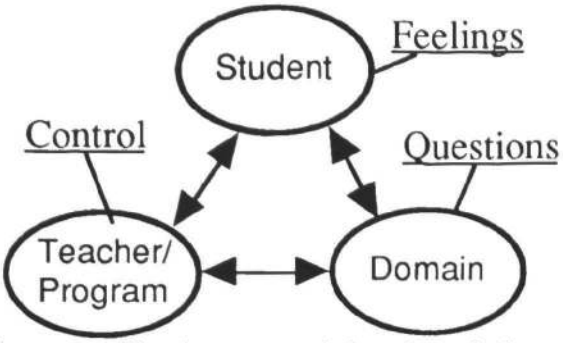

Figure 1. The three natural domains of discourse in a student-teacher interaction and their corresponding categories of communicative acts (underlined).

or task which he or she is engaged in learning. The student may be seeking to understand the reasons for doing or knowing something, or may want to know what to do next.

- **Control** – This third category contains expressions about the progress of the lesson. The student may wish to speed up or slow down the rate of instruction, or may wish to have material presented in greater detail. The structure of a lesson is often tightly linked with the domain being taught, and hence the two are often difficult to tease apart. Nevertheless, this category is intended to deal with the mechanics and style of teaching rather than with the domain.

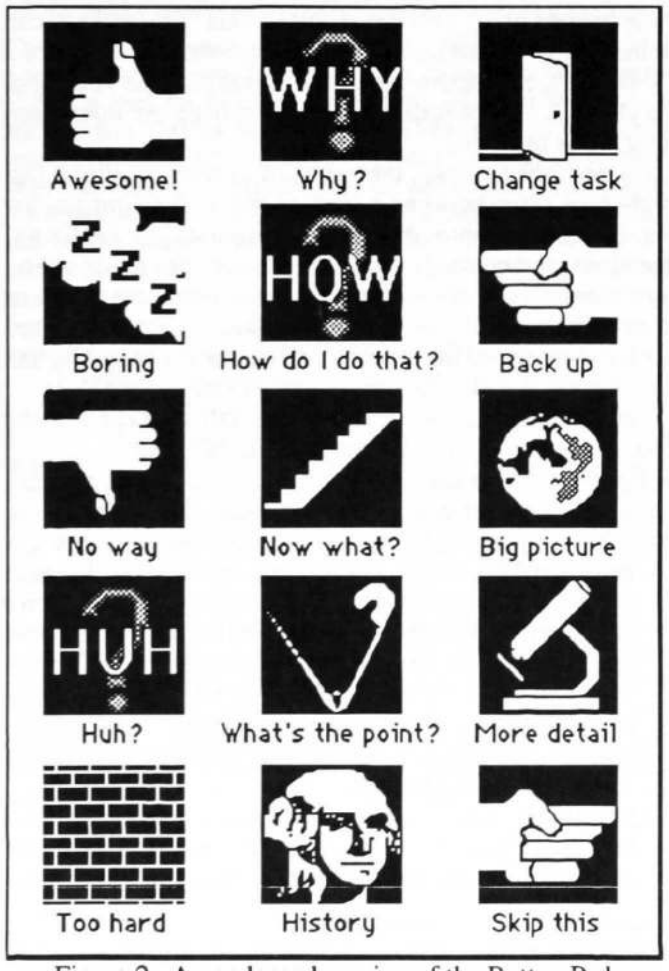

Figure 2. A condensed version of the Button Pad.

Within each of the above categories, we have selected what we believe to be the five most important, most frequently used communicative acts, and assigned each to an iconic button on the Button Pad (see Figure 2). The five messages we have chosen for each category are as follows:

- **Feelings** — "Awesome!", "Boring", "No way", "Huh?", "Too hard"
- **Questions** — "Why?", "How do I do that?", "Now what?", "What's the point?", "History"
- **Control** — "Change task", Back up", Big picture", "More detail", "Skip this"

It should be clear that the first step in developing a theory of communicative interaction in tutorial situations must be to develop a taxonomy of kinds of communicative acts that a student might be expected to perform in such situations, as we have described above. This is true regardless of whether the ultimate application of the theory will be in the development of the sort of iconic interface we are pursuing, or in natural language processing. That is to say, the first step in developing any theory of this communicative process must be to characterize the class of expected inputs. Moreover, this characterization must be functional in nature, that is, it must facilitate a determination of the intended meaning of a communicative act, and of an appropriate response. In other words, a representation theory must be developed that enables us to categorize student inputs in terms of both content and intent. The taxonomy described above embodies one such theory.

It is certainly arguable that our choices within each category are, to a certain extent, arbitrary, or that we should have chosen more or fewer items in each category. Our response to such arguments is similar to that offered in defense of any knowledge representation theory.[2] We may not have the right number of buttons, nor do we claim that we have identified the only possible "correct set." Rather, we claim that this is one reasonable solution, and that any equally reasonable set of choices would be roughly equivalent in a functional sense. To put this another way, the choice of buttons in each category is not a question of right or wrong: The real question is, how well, and how easily, does this comparatively narrow channel of communication allow students to express themselves? Our experience using the Button Pad in several computer-based learning environments indicates that Button Theory provides a good balance of ease of use and expressivity.

The Phone Tutor: A simple case study of Button Theory in action

To get a better feel for the kinds of interactions supported by Button Theory, let's look at a concrete example of the theory in action. The example comes from a prototype system called the Phone Tutor which is designed to help phone company customers learn how to use special phone

[2] For example, see Schank (1975) for a discussion of this issue as it arises in the context of the Conceptual Dependency meaning representation for natural language.

services such as call waiting or call forwarding.[3] The idea behind the system is to provide users with an enhanced telephone keypad containing a few additional buttons that enable them to interact with an on-line tutor. The three additional buttons on the phone keypad are a small subset of the fifteen buttons discussed above, one from the **feelings** category ("Huh?"), and two from the **questions** category ("Why?" and "Now what?").

The system responds to button presses with recorded audio messages. A transcript of a sample interaction is shown below. The student's actions are shown in boldface.

Practice

Thanks for calling the Phone Tutor. The Phone Tutor can help you learn to use Custom Calling Services. Choose the service about which you are confused. Press 1 for call forwarding, press 2 for call waiting, press 3 for speed dialing, or press 4 for three-way calling.

1

With call forwarding, you can have your calls -- local or long distance -- transferred to another number. That means your calls can follow you when you leave home or work. Is this the first time you have practiced call forwarding?

Yes

OK, we will start by practicing the procedure for setting up call forwarding. The first step is to dial 72. Do that now.

Why?

Pressing 72 is the special code that tells our computers that you are about to set up call forwarding. Other special services use different codes. Now type 72.

72

Good. Now press the # key.

Huh?

The # key is just to the right of the 0 key. It is also labelled "No". Now press the # key.

Why?

Pressing the # key is needed to indicate that you want to access a special service instead of dialing a regular number. All special phone services, including call forwarding, require you to dial # after the special code. Now press the # key.

#

That's right. Now type the number you want to forward to.

Huh?

After dialing 72#, indicating that you want to set up call forwarding, the next step is to enter the number to which you wish to have your calls forwarded. Now type the number you want to forward to.

Now what?

Try it.

Now what?

Type the number you want to forward to.

5553500

So far you've practiced typing 72# and the number separately. Now try to do it all at once. Press 72, then the # key, then the number.

725553500

Let's try that again. Press 72, then #, then the number you want to forward to.

72#5553500

OK, for practice, pretend you would like to forward a call to 555-8989.

[dialog continues...]

The program that produces these dialogs is rather simple in structure: It is essentially an augmented finite state machine, in which button presses and student responses are used to determine what sort of action the program should take, within the context provided by a network representing the hierarchical goal/sub-goal structure of the task, the student's current position in that task, and a few other features, *e.g.*, the last button they pressed. The relative richness and coherence of the dialog presented above is not really due to the algorithm employed. Rather, it reflects the utility of the three queries permitted by the available buttons in conjunction with the content of the task network. The message here is that with a good representation of what needs to be taught, and a sensible taxonomy of student queries, even rather simple mechanisms will suffice to produce relatively natural and flexible interactions with a teaching system.

Buttons as conceptual interface

An important goal of Button Theory is to shift the burden of mapping from the learner's needs to the system's responses from the learner to the system itself. For example, a great deal of research in human-computer interaction is concerned with how easily a user can associate a screen icon with a system action or component (see, *e.g.*, Pejtersen & Goodstein, 1988; Norman & Draper, 1986). In contrast, Button Theory is concerned with making it easy for the student to associate screen icons (*i.e.*, buttons) with conceptual categories that directly express his or her feelings, questions, or desires. Determining how to interpret and respond appropriately to a button press is up to the system itself; the user need not—indeed, often cannot—specify the particular action to be performed by the system.

As Norman (1986) points out, the chief task of any interface is "to bridge the gap between [user] goals and [the] system." He continues, "there are only two ways to do this: move the system closer to the user; move the user closer to the system. Moving from the system to the user means providing an interface that matches the user's needs, in a form that can be readily interpreted and manipulated" (p. 43). Herein lies the strength of the Button Pad interface. By incorporating the basic messages provided by our taxonomy

[3] The Phone Tutor system was built by the first author.

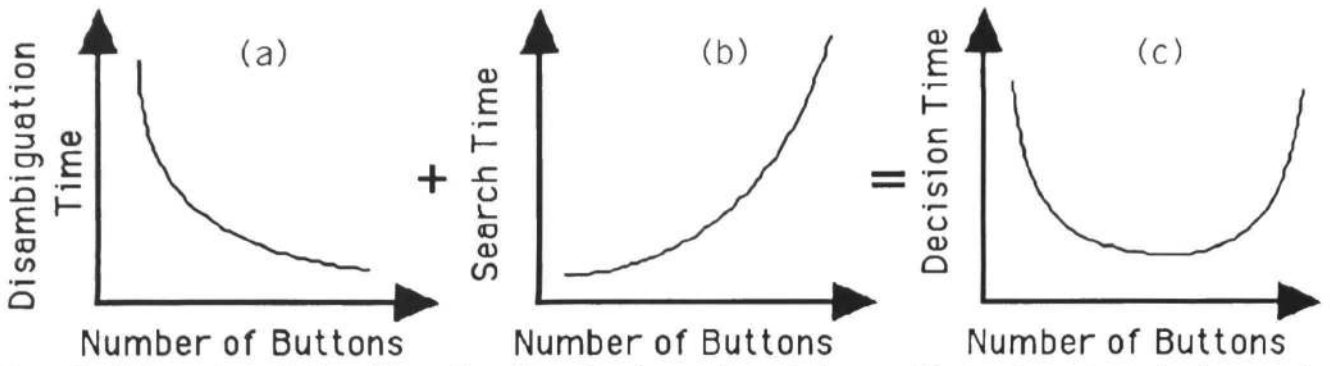

Figure 3. Schematic depiction of how disambiguation time and search time combine to shape the student's ease of use as a function of the number of buttons.

of student-teacher communication, the Button Pad interface enables the learner to express his or her needs directly, in a way that is nearly as simple and natural as natural language itself. This close match provides the Button Pad interface with what has been called *semantic directness*, the ability to easily express things of interest using the interface (Hutchins, Hollan, & Norman, 1986). Interfaces that provide users with these important characteristics we term *conceptual interfaces*.

The Button Pad (Figure 2) always displays the complete range of expressions possible. The visual presence of the Button Pad on the screen is a subtle but important aspect of the efficacy of this type of interface. The sequence of mappings that the user must construct from the thought he or she wants to express, to a screen icon, to a program command, is made simpler and more direct. The availability of the buttons offers a palette of general utterances from which the student need only select the one closest to what he or she wants to say.

A related question, given the restricted vocabulary provided by Button Theory, is the effect that the interface has on how the learner thinks about the domain, "because the interaction language demanded by the physical system comes to color the thoughts of the person" (Norman, 1986, p. 39). The manner in which the way we think about something is shaped by the the vocabulary we have for talking about it is known as *linguistic determinism* (Hutchins et al., 1986).[4] This is problematic in that the restrictions imposed by the interface may come to restrict the user's ability to think flexibly about the domain (Hutchins et al., 1986). Our expectation is that, because the Button Pad provides the learner with an interface that embodies natural categories of expression, it is unlikely to force an overly limited perspective on the learner. In fact, because it always displays the entire range of queries possible, the Button Pad interface turns linguistic determinism to its advantage, by assisting learners in the often difficult process of question formation. Indeed, it may well lead to better, more creative thinking, because it encourages students to ask more and better questions than they might ordinarily ask (Schank, 1986, 1988).

[4] This is, of course, closely related to the Whorf-Sapir hypothesis.

Tradeoffs in designing a set of buttons

There exist a number of underlying tradeoffs in designing a set of buttons. From the perspective of building a program that can respond intelligently to button presses, clearly the more buttons the better. More buttons, in general, means that there will be less ambiguity to resolve for each individual button press. The range of possible meanings associated with each button press will decrease as the number of buttons increases, since more buttons are being used to cover the same range of possible meanings. Unfortunately, an indefinitely large set of buttons is not a plausible option when building a user interface.

Moreover, from the user's perspective the situation is not so clear-cut. With a small number of buttons, the student is faced with the task of fitting his intended message to the button that he thinks best subsumes that message. Whether or not this is a severe bottleneck will depend upon the details of the decision procedure the student employs in this task. If the student first formulates an intended message and then attempts to determine the button that best expresses it, as seems plausible (especially in the case of novice users), then his task will be the inverse of that facing the interpreter of the button press, and he will face a similar problem of ambiguity (Figure 3a). With a large number of buttons, the student is faced with another difficult task, that of finding the button that best expresses his or her question (Figure 3b). Presumably there is a middle ground (analogous to the basic level for categories described by Rosch et al., 1976) which closely matches the kinds of things students typically wish to express, minimizing the cognitive effort the student must devote to expressing his question by reducing the amount of searching and fitting that is required (as reflected by decision time; see Figure 3c). We have selected a set of buttons which we believe corresponds closely to the needs of the student while at the same time providing sufficient clarity to facilitate implementing the appropriate responses in a computer-based learning environment.

Interpreting button presses

Another prerequisite for the success of Button Theory is that computer-based teachers employing the buttons be capable of interpreting what the learner means by a given button

press in the context of a tutorial dialog. The tractability of this interpretive task is dependent upon the availability of contextual information that can significantly constrain the meaning of button presses. We believe that this contextual information can be made available in tutorial dialogs.

In computer-human interactions, context acts to reduce ambiguity in two distinct ways. First, it can make the meaning of the buttons in a specific situation clearer to the learner (Jones, 1987; Pejtersen & Goodstein, 1988), thus reducing the likelihood that a user pushes the "wrong" button. Second, the context naturally constrains the range of possible meanings the learner may be expected to have in mind when posing a question to the program. For example, Moore (1989) has investigated appropriate responses to "huh" questions, Moore & Swartout (1990) present an analysis of "why" and "how" questions, and Kahneman & Miller (1986) discuss "why" queries. This work, in conjunction with our own experience to date, supports the claim that the disambiguation of button presses by a computer-based teacher is tractable.

Future work and conclusions

To investigate the usefulness of Button Theory we have incorporated the Button Pad interface in several prototype computer-based learning environments. In addition, we are currently investigating a number of empirical questions raised by Button Theory, including whether all of the buttons we have devised are indeed useful, and what kind of response a teacher should make to each one. We have begun to design a suite of experiments to investigate these questions, which are intended to shed light on such issues as the optimal number of buttons, which messages are most necessary to include, what kind of response students expect when they press a given button, what kind of response human teachers actually give, and how well these latter two correlate. We would also like to investigate the extent to which the buttons affect both the quantity and quality of questions asked by students. The hypothesis to be investigated is whether, rather than constraining the kinds of questions asked, the Button Pad will actually encourage not only more question asking, but also better question asking (Schank, 1986, 1988). We hope to report the findings of these experiments in the near future.

It should be apparent that the problem of correctly interpreting and responding to a given button press by a given student in a given context is, in its full extent, as difficult a problem as natural language understanding. In fact, it is the same problem. Our hope is that the framework provided by Button Theory will permit us to tackle this problem incrementally, starting with simple responses for a subset of the buttons, and gradually increasing the sophistication of the interpretive mechanisms underlying each button as necessary. The key point is that even with rather simple mechanisms, this framework enables surprisingly rich interactions. These interactions are only as successful as the taxonomy upon which they are based; but even the rather straightforward analysis of student communicative needs offered by Button Theory provides a great deal of leverage.

References

Hutchins, E.L., Hollan, J.D., & Norman, D.A. (1986). Direct manipulation interfaces. In D.A. Norman & S.W. Draper (Eds.), *User Centered System Design: New Perspectives on Human-Computer Interaction.* Hillsdale, NJ: Lawrence Erlbaum.

Jones, W.P. (1987). How do we distinguish the hyper from the hype in non-linear text? In H.J. Bullinger & B. Schackel (Eds.), *Human-Computer Interaction: INTERACT '87.* Amsterdam: North-Holland.

Kahneman, D., & Miller, D.T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93,* 136-153 .

Moore, J.D. (1989). Responding to "huh?": Answering vaguely articulated follow-up questions. In *Proc. of the Conference on Human Factors in Computing Systems* , Austin, TX, April 30-May 4 1989.

Moore, J.D., & Swartout, W.R. (1990). Pointing: A way toward explanation dialogue. *Proc. of Eighth National Conference on Artificial Intelligence* (pp. 457-464). Cambridge: MIT Press.

Norman, D.A. (1986). Cognitive engineering. In D.A. Norman & S.W. Draper (Eds.), *User Centered System Design: New Perspectives on Human-Computer Interaction.* Hillsdale, NJ: Lawrence Erlbaum.

Norman, D.A. & Draper, S.W. (Eds.) (1986). *User Centered System Design: New Perspectives on Human-Computer Interaction.* Hillsdale, NJ: Lawrence Erlbaum.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas.* New York: Basic Books.

Pejtersen, A.M., & Goodstein, L.P. (1988). Beyond the desktop metaphor: Information retrieval with an icon-based interface. In P. Gorny & M.J. Tauber (Eds.), *Visualization in Human-Computer Interaction.* Berlin:Springer-Verlag.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382-439.

Schank, R.C., & Jona, M.Y. (1991). Empowering the student: New perspectives on the design of teaching systems. *The Journal of the Learning Sciences, 1,* 7-35.

Schank, R.C. (1988). *The Creative Attitude: Learning to ask and answer the right questions.* New York: Macmillan.

Schank, R.C. (1986). *Explanation Patterns: Understanding Mechanically and Creatively.* Hillsdale, NJ: Lawrence Erlbaum.

Schank, R.C. (1975). *Conceptual Information Processing.* New York: American Elsevier.