From Calvin to Freud: Using an Artificial Intelligence Model to Investigate Cognitive Changes during Psychoanalysis*

Marianne Leuzinger-Bohleber and Horst Kächele*

> "There is the old joke about not really understanding something until you have taught it to somebody else. The new version now says you don't really understand it until you've programmed it into a computer." (Randall Davis)

1. Introduction

Can you imagine a greater difference than the one existing between a computer and a patient?

"You know for some reason I just thought about the bill and about payment again. (You shoudn't give me a bill.) <Uh->I was thinking that I (shoudn't be given a bill) of asking you whether it wouldn't be all right for you not to give me a bill. That is, I usually by (the end of the month know the amount of the bill), well, I immediately thought of the objections to this, but my idea was that I would simply count up the number of hours and give you a check at the end of the month." (Clippinger 1977, p. 146)

Perhaps you will find it hard to believe that this text was not taken from a verbatim transcript of a psychoanalytic session, but was produced by a computer program called ERMA, written by J. Clippinger (1977). Indeed, the text produced by his computer program is nearly the same as the actual discourse in the transcript of a real psychoanalytic session that Clippinger originally used to conceptualize the possible cognitive pro-

*To Ulrich Moser for his 60th birthday

*Supported by the German Research Foundation, Sonderforschungsbereich 129, project B6.

cesses of the patient, processes that may indeed have "run" inside her "black box" and led to this verbalization.

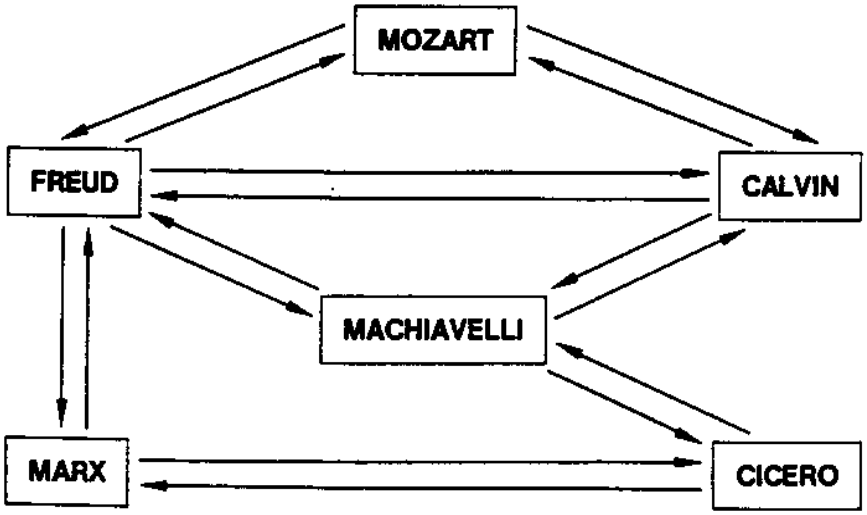
The program's production of this text may serve as evidence that the cognitive theory upon which the computer program was based is an adequate model of most of the cognitive processes involved. This technique of computer simulation to test theories (psychological in this case) is one application of the rapidly developing field of Artificial Intelligence (AI) (cf. Teller, this volume).

In our opinion this approach to testing complex psychological theories can also be fruitful for psychoanalysis. The models derived from AI theories, partly because of being tested by computer simulation, have a high degree of precision, consistency and clarity in the terminology used. Moreover they take into account, much more than psychoanalytic concepts do, modern knowledge from other scientific disciplines such as neurophysiology, brain research and research on memory structures. These advantages are one reason why we have used a theory of cognitive processes based on computer simulation models in our study of changes in these processes in psychoanalytic patients. Borrowing an AI model in an empirical study of psychoanalytic processes also helps to avoid the dangers of circularity, i.e., using a psychoanalytic model to justify itself.

Clippinger's theory of cognitive processes was convincing to us because it embodies the conception of conflicting processes taking place inside a black box, just as the structural theory in psychoanalysis does. That is, it conceptualizes cognitive processes as being determined by the interaction of separate cognitive modules. The processes (programs) running in one module can complete, modify or inhibit and interrupt those running in other modules. Among other things, this leads to characteristic structures in the interaction of the different modules and specific ways of perceiving and processing information.

In this study we used Clippinger's model in a modified form by defining the six modules shown in Figure 1. These modules perform the following tasks:

MOZART selects what is attended to.
CALVIN represents the superego and the patient's values, and acts as censor.
MACHIAVELLI develops problem-solving strategies.
CICERO translates cognition into verbalizations.
MARX perceives and tests reality.
FREUD introspects and performs specific ego functions.



Arrows indicate paths of communication among the cognitive modules.

Figure 1 Interaction of Cognitive Modules

The psychoanalytic notion of the Id as a source of unconscious motivation is not explicitly represented in this model. Clippinger includes the concept of motivation in his sequential representation of cognitive processes, but its exposition here would take us too far afield. For a detailed understanding of the operation of the model the reader is referred to Clippinger (1977). Nonetheless it is obvious that unconscious motivations ultimately reveal themselves in cognitive processes, and it is the manifestation of these in the transcripts of what patients verbalize on the couch that we study.

2. Overview of the Study

In our empirical study we aimed at describing and analyzing changes in the problem-solving cognitive processes of five patients during their long-term psychoanalyses. We studied modifications of the way the

patients themselves handled their dreams during psychoanalytic sessions. We focused on this question because one general goal of a psychoanalysis is that unconscious conflicts should become conscious as a precondition for being able to reach other more external goals of a psychoanalytic treatment such as the ability to work, to love and to enjoy life. The patient should learn to recognize unconscious conflicts in order to avoid their interference with the satisfaction of his wishes and duties. In this special form of psychotherapy he is expected to develop specific problem-solving strategies for dealing with unconscious conflicts.

Therefore the changes in problem-solving cognitive processes as one important criteria for the success of a psychoanalytic treatment was the central question in our research project. We analyzed such cognitive processes in reference to dream interpretations because they can serve as an example of the way the patient deals with unconscious material, e.g. his dreams.

Here we shall give a highly condensed overview of a few of the results of our project. In the first phase of our study we generated hypotheses from a single-case study by exploring dream associations as recorded in a transvestite patient's (Leuzinger-Bohleber 1987) diary during the first and last hundred hours of his psychoanalysis. In the second phase we tested our hypotheses by using four additional psychoanalytic cases from the Ulm Textbank. Using two kinds of theory-directed content analyses, we compared each patient's dream reports taken from the first hundred with those from the last hundred psychoanalytic sessions.

In all that follows it should be understood that we use a very broad definition of "cognitive processes" as inner processes of perceiving and processing information that are always connected with physiological and emotional processes and cannot be studied separately (Pfeifer and Leuzinger-Bohleber 1986)

3. Methods: Theory-Directed Content Analysis by Complex Ratings and Computer-Aided Content Analyses

Our tools for testing the hypotheses derived from the Clippinger model consisted of (1) computer-aided content-analyses and (2) ratings by trained judges of transcribed sessions and written dream reports. We used two "naive" raters and two "expert" (psychoanalyst) raters in this study. The judges were intensively trained to understand our model of

cognitive processes. In several pretests they were acquainted with the kind of material to be rated. The training was very time-consuming, but we finally achieved reasonable interrater reliabilities. The raters judged randomly ordered sessions taken from early and late in the treatment in which there were dreams and dream associations, including the patients' own interpretations of their dreams.

The following examples of diary reports from our first patient are taken from one early and one late session:

Session 19

Let me first tell you a dream: There is a church, large, modern architecture, triangular walls with large windows. Daniela (a girlfriend) and me, probably also her relatives went with us into the church. While entering I already felt anxious that those relatives wouldn't sit nearby an exit door. To me it was clear in a war one would safeguard the possibility of rapid evacuation. In the church there was a staircase leading upward to a second story. Some people moved up there. To my satisfaction, we placed ourselves where I wanted to be seated. The pulpit was like a wooden cage. It was swinging on ropes fixed on the ceiling. Suddenly I heard the noise of aircraft; I hurried and left the church and hid myself behind a wall. I knew there was no absolute protection, but at least some shelter. Then I saw the airplanes dropping bombs, very clearly, especially the bombs. I felt like being in a nightmare so strong was my anxiety.

Associations: The church, its triangular shape reminds me of Lionel Feininger, the architect; maybe it's also influenced by the famous chapel of Ronchamps by Le Corbusier. As I myself am interested in architecture and do some drawings too from time to time, this church has attracted me very much. On specific questioning by my analyst, what interests me most are the curbs, the roof and the shadows, the uneven walls and the non-rectangular shapes.

I once went with Daniela into the church. Because she was used to going to church every Sunday, I thought that I would want to justify my prejudices. But it was no prejudice; I felt my judgment confirmed, especially with regard to the Catholic church.

The bombs: I used to construct many models as a boy, airplanes and tanks. The airplanes had bombs as well, and I did it with special perfection. About eight years ago I used to be very interested in war materials. I was especially excited by their technical perfection. I was caught up by that in a way that I didn't realize how much human misery and suffering the technical world spreads in such ways. In the war movies I've seen, the technical aspects were exaggerated and the human aspects neglected. When I realize today

the way I felt then, it makes me shiver. I realize the way I used to be; that interest in technical aspects is basically there but still I have discovered these inconsistencies and now am trying (with some success) to mold this interest in more human concerns. The fear that I may be forced in the long run to earn my money in a technical job still remains. I also still feel anxious that I may relapse into this ideological thinking – I doubt it, but it is still there. Another aspect of this apprehension is combined with my sophistication for technical innovations; it might well be that I would invent something really new with a lot of controversial consequences. When I was nine years old I used to construct rockets, which really intrigued me.

How would I have developed if I had become an engineer in the war industry, terrible thought, isn't it? The interest in this stuff was not directly influenced by my parents; it seems to have developed out of my search for something elementary, out of my problems and circumstances surrounding me.

There is another event I remember: as a child I went Christmas shopping with my parents. I saw a tank and wanted to have it. My father said, "Look, what an ugly (maybe dangerous, but not inhuman) toy." With my construction kit I couldn't build a tank; it was late in school when I was able to build my first tank, which I was very proud of.

No, I haven't really worked it through. Whenever I daydream that I would stay alone in a house and a burglar would try to break in, I place such a little tank in front of him and would then be able to really shoot and hurt him. It looks as if I am preoccupied with defending myself, even in fantasy, against an invisible intruder and being invisible myself. All this is related directly or indirectly to the bombs, which I had a clear picture of in the dream.

Session 617

I had a dream: a shop, I stand in front of the shelf with stockings, but I cannot find the right ones. I leave the shop. I have about three quarters of an hour before my analytic session begins. Change of scenery. I wake up and find out that I have missed the analytic session. I should call my analyst and tell her, or else she might be worried. Something keeps me from doing it, but having missed the first session I should take care that I don't miss a second. This session will take place in X; taking a bike I should be able to make it just in time. While riding there I suddenly am confronted with a mountainous road, hard to follow and very difficult. The situation frightens me. In a small hut there is a young man with long hair; he's got to do something about the bad road. Time runs very short and though I see X on the horizon I am not able to make it in time. I am very anxious, feel stiff; I look at the freaky young man working and I think I should try to retrieve my bike which had been lost somewhere in the mountain.

Associations: It's not only my (female) analyst, but I can't reach any woman; I cannot open myself; I pretend to come to X. I am wondering, did I betray her? Could it be that the analysis has not been as successful? In the dream I have missed two session whereas in reality I never skipped one. Has it to do with the fear of not being loved any longer? The issue around men: do I want to observe them and imitate them; what could it be? It was yesterday that I realized what my analyst had in mind when she said there was a wish to imitate another man, but this is very hard. Most men I know have their difficulties so that there is very little to identify with, anyway at least with regard to my technical abilities. I always was a very efficient observer and looked carefully at how things work. However there are areas where I am not able to observe so keenly. What does X stand for? Well it looks as if I am on the verge of establishing a new part of my identity; recently I talked to my boss about a chance to share a social worker's position with someone else. I had this idea a few days ago. I had even disputed it with a friend of mine; however, I think this does not belong to the dream anymore.

Now, what hypotheses were being rated?

From CALVIN to FREUD – An Example

One of the central concepts in Clippinger's model is that cognitive processes are determined by the interaction of programs running both in and between cognitive modules. Accordingly, our first hypothesis was that in the diary notes (or verbatim protocols) from the early part of the psychoanalysis the programs of the CALVIN module would especially dominate those of the FREUD module. In other words, the introspective associations of the patient would be inhibited by interventions of a rigid forbidding superego structure. In contrast, we postulated that we would find the FREUD module dominating in the sessions at the end of psychoanalysis, corresponding to the possibility of "freer" associations as one result of psychoanalysis.

The raters were asked to judge whether, after having read the material of a session dealing with dream associations (as in the examples above), they had the impression that the cognitive processes operating in each of the different modules: (1) were controlled by those of other modules (1, on the scale); (2) could have "run" without being inhibited by those of other modules (2, on the scale); or (3) controlled the cognitive processes of other modules (3, on the scale).

3. Results

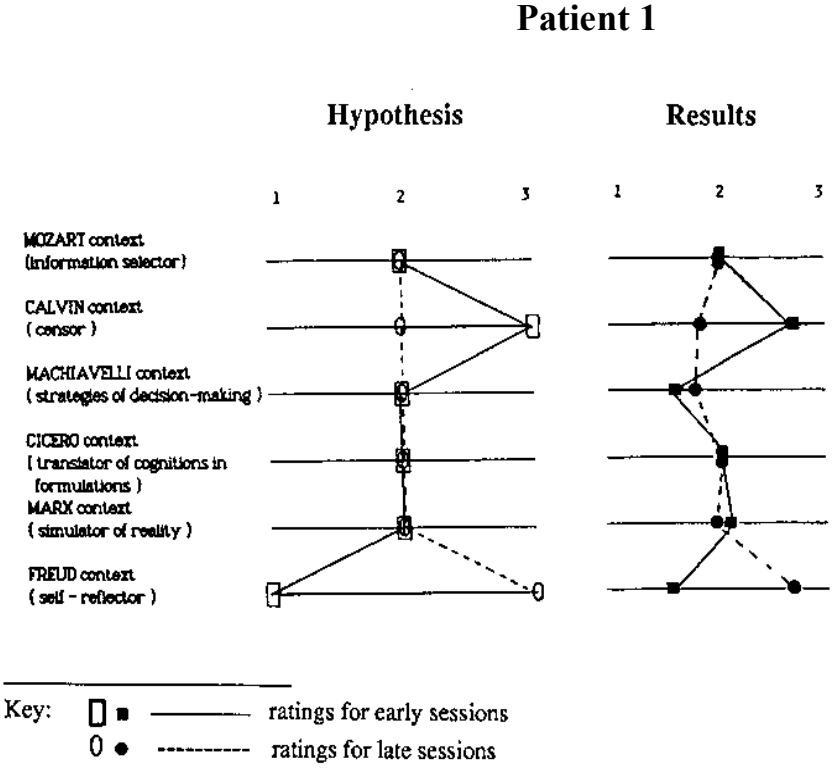


Figure 2 Hypotheses and Results: Rating of the Interactions of Cognitive Modules

The results, shown in Figure 2, supported our hypothesis. We were surprised at how easily the different clinical outcomes of the five psychoanalyses were distinguished by this pattern. Patients 1 and 5 supported our hypothesis in the most impressive way; patients 2 and 4 also supported it, but less strongly, and Patient 3 did not show the expected changes. The outcomes were assessed by: (1) global clinical impressions of the analysts of the patients, (2) the judgments of two independent analysts, and (3) the results of our theory-guided computer-aided content analyses. The outcomes of Patients 1 and 5 were judged to be "very successful," Patients 2 and 4 were judged "moderately successful," and Patient 3 was judged "unsuccessful," although this treatment had not yet been concluded at the time of our study. Indeed, the global

clinical impressions were sharpened, modified and corrected in some dimensions by our empirical study.

Other hypotheses focused more on changes of the content of the programs running inside the different modules. Figure 3 shows a few examples from the CICERO module for Patient 1.

1. How do you feel about the way the analysand expresses himself?
   Is his personal language

not striking, normal	.....x.....o.....	striking, special
simple	.....x.....o.....	complicated
idiosyncratic original	.....ox..	socially adapted
plain, simple	.....x.....o.....	sophisticated
hard to read	.....o....x.	easy to read
unpleasant	.....o.....x	pleasant
egocentric	.....o.....x..	communicative

2. How do you judge the range of the dream interpretation?

narrow	...o.....x	broad

3. Are there taboos in the themes in the sequences of dream- interpretation?

hardly observable	....x.....o....	clearly observable

4. Do you find indications in the text that the patient takes into account to whom his notes are directed (normally his analyst)?

hardly observable	...o.....x.....	clearly observable

5. How "psychoanalytic" do you feel the expressions of the patient are?
   psychoanalytic in a "technical" sense?

minimal	....o.....x.....	clearly

   psychoanalytic in an "integrated" sense?

minimal	.....o.....x ..	clearly

_____

Key: o = rating of early sessions
x = rating of late sessions

Figure 3 Questions for the Raters Concerning the CICERO Module

Please note that the different questions are not independent of each other. The aim of this part of our study was to generate hypotheses. The questions focus on the real or fictitious effect of the diary notes on the real or fictitious reader – the instrumental aspect of language as Bühler (1934) called it – which is part of the function of the CICERO module in our model.

We interpret the many different results as indications that the expression of language in the second part of the diary came closer to the purpose of such a diary, i.e., the language was simpler, plainer, and less exaggerated. The language also more closely served the function of communication with a real or imagined partner, one reason why the notes were judged to be more readable and desirable by the raters. We also think that we have found indications that several aspects of the language performance and the knowledge of social relationship of the patient were expanded by psychoanalysis, i.e., broader range of themes, fewer taboos in the notes, greater range of performance and of psychoanalytic knowledge, and taking into account the effect on the possible partner in communication.

Two examples will illustrate how the results of these ratings were supported by our computer-aided content analyses. First, the judgment of fewer taboos late in the analysis was supported by the fact that the late diary entries contained more direct sexual terms than early ones. Second, the impression of the raters that the patient took the reaction of a possible reader of the diary more into account at the end of the psychoanalysis was supported by the fact that interventions of the analyst were mentioned in only 5 percent of the early entries, whereas in the late ones 16 percent contained reflections on the analyst's statements. Furthermore, only 2 percent of the early references were to the analyst's interpretations, but this increased later to 10 percent; and the early references tended to be the analyst's questions, confrontations and clarifications, compared with later reflections on his interpretations.

Summary of all Five Patients

I. Changes in Problem-solving Cognitive Processes: Interactions among Cognitive Modules

The problem-solving cognitive processes of the two successful patients at the end of their psychoanalyses can be characterized by a high degree of flexibility, by an enlarged cognitive range, an associative and

"gestaltlike" way of thinking, and by a capacity for a functional and realistic style of problem-solving. Different information could be perceived and worked on at the same time and led to a process of generating and testing hypotheses that could compete with, modify, or contradict each other. Cognitive dissonances were recognized, reflected, and influenced, among other things, the decision-making process.

Unpleasant affects had an important function as signals indicating cognitive processes to be taken into account in the problem-solving process. In terms of our model, we found: (1) increased cognitive and affective knowledge used in a functional way in different modules, (2) interrupt programs that functioned well and corresponded better to reality, and (3) an uninhibited interaction of cognitive processes in the different modules. The two moderately successful patients showed results tending in the same direction, but not as impressive as the two very successful patients. Up to this point in the treatment, the unsuccessful patient had not yet changed his cognitive processes in the ways described.

Changes Within the Cognitive Modules

*1. **MOZART**: Changes in What was Attended To*

In the sessions at the end of the psychoanalyses of the successful patients the following changes were observable, according to our raters:

-More of the text of the dreams was attended to and worked over cognitively.
-The context of the dreams was taken into account.
-The analyst's interventions were part of the patient's dream associations.
-The patients pursued hypotheses about their dreams more systematically.
-The process of generating hypotheses took place easily, without much hesitation.
-The patients considered more than one hypothesis about the meaning of a dream.

In a separate assessment we observed the following systematic changes in three dimensions of the manifest dream content, based on the model by Moser et al. (1980): *Expressed relationships, dream atmo*

sphere and problem solving. In the late dreams of the successful patients we observed:

Expressed Relationships:

-The dreamer expressed better relationships with both his objects and himself.
-The range of interactions in these relationships was increased e.g. in the late dreams he was more often alone, as well as interacting with one or more partners.
-Although the relationships were more often tender and friendly than in early dreams, to our surprise, they were also seldom neutral, and included conflictual relations – an indication, to us, that the range had been increased.

Dream Atmosphere:

-The variety and intensity of affects in the manifest dream content was increased.
-The atmosphere was more positive with less anxiety, but aggressive, sad and frightened moods were also expressed. This contradicted our original hypothesis that a single positive mood would prevail.

Problem solving:

-More problem-solving strategies were recognizable.
-Problem solving was more successful than not and the dreamer was more active in doing it, and seldom avoided it.
-The range of problem solving was greater than in early dreams.

In the content analyses we found less concern with the major psychopathological symptoms in the successful patients. In the late dreams the content was more personal, with a greater variety of expressed activities. Moreover, the patients' dream interpretations were more "dialogue oriented," more convincing and more directed at understanding the unconscious meanings of the dream. The associations were more constricted early and more varied in the late sessions. These are hints that the range of attention of the successful patients was enlarged.

2. ***CALVIN:*** *Changes in the Superego and Internal Values*

The raters observed significant changes in the internal values of the successful patients. Their criticisms and judgments grew more mature, milder, more flexible, more adequate, and more encouraging, but remained consistent with inner ethical values already present at the

beginning of their treatments. This included how patients judged themselves as well as how other relevant other persons judged them. The content analyses supported the raters' findings.

3. ***MACHIAVELLI:*** *Changes in Problem-Solving Strategies for Dreams*

The raters judged that the successful patients increased their ability to interpret their dreams. In late sessions they displayed more strategies for interpreting their dreams, such as working with dream symbols, integrating different themes in the dream and associations, and dealing with contradictory information.

Since free associations are an essential part of dream interpretation, we defined five dimensions for assessing the strategies employed: *quantity* (the amount of free association), *variability* (the range of strategies employed), *quality* (the usefulness of the associations), *introspection* (the ability to reflect on the interpretations), and *incapacity* (helplessness in the face of the task). All five dimensions discriminated the very successful from the moderately successful and from the unsuccessful patients.

4. ***CICERO:*** *Changes in Language Expression*

We found a variety of changes in both the ratings and the content analyses of the patients' language. Three of these changes were: (1) the language became more socially communicative rather than egocentric; (2) affects were more integrated into expressions rather than remaining isolated; and (3) the vocabulary became more varied in later sessions.

5. ***MARX:*** *Changes in Reality Perception and Reality Testing*

We observed fewer changes in these functions than we had expected. We had significant results only in changes in the patient's *self-descriptions.* The four "very" and "moderately" successful patients all described themselves more realistically and less conflictually in the late sessions. And by the end the two very successful patients became more "empathic" in their self-descriptions.

6. ***FREUD:*** *Changes in the Capacity for Introspection*

Here the differences in clinical outcome were clearly observable. The raters found that the very successful patients showed the most

increase from the beginning to the end in their ability to introspect. Moreover, their introspections were "productive" (complex, profound, more intensely experienced) rather than "intellectualized" or "rationalized," and led to new insights and thorough working-through of conflicts. The raters also found more "good" late analytic sessions in the successful patients. There were also some instances in which the successful patients reflected on their own dream interpretation strategies. Finally, although raters could not find support for our hypothesis that these capacities would be based on demonstrable identifications with the analyst, the analysts and patients were emotionally "closer" to each other at the end. Thus we concluded that the introspective capacities of the successful patients were less inhibited and analytically more fruitful.

7. Changes in Patients' Motivation to Understand Dreams

The raters found an increase in the motivation to understand dreams in one of the very successful patients (P1) and the two moderately successful patients (P2 and P4). Patient 5, the other very successful patient, was already highly motivated to understand his dreams at the beginning of analysis and continued so. The unsuccessful patient showed no increase in this motivation.

4. Summary

We have presented a brief and sketchy overview of a study designed both to generate and to empirically test hypotheses about changes in problem-solving cognitive processes of five patients during their psychoanalyses. In Phase 1 the hypotheses were generated and empirically supported in a project in which we analyzed a diary kept by a transvestite patient on his dreams while in psychoanalysis. In Phase 2 the hypotheses were further tested using verbatim transcripts of four other psychoanalyses stored in the Ulm Textbank (see also Leuzinger-Bohleber in preparation).

The five psychoanalyses were independently and reliably assessed by each of the treating analysts and by two independent analysts, resulting in three outcome categories: "very successful" (Patients 1 and 5), "moderately successful" (Patients 2 and 4), and "unsuccessful" (Patient 3). The most striking result of the study was that the ratings of four judges of a substantial number of theory-derived variables, as well as our computer-aided content analyses, discriminated among the three outcome categories. Of course even our five combined single-case studies are still a small empirical base, but the many results differen-

tiating the three outcomes can serve as specific predictors of success to be tested in future studies of other cases.

Finally, methodologically, our use of an Artificial Intelligence model as a guide for our theory-centered ratings and our computer-aided content analyses has proved quite fruitful. The variety and face validity of many of our results should lead to a better understanding of changes in cognitive processes during psychoanalysis. We believe that this innovative trial of the model was worthwhile and provided us with more consistent, precise, differentiated and practical foci for further empirical studies than existing psychoanalytic theories. We do not go as far as Teller (this volume) when she suggests that Artificial Intelligence can be a basic science for psychoanalysis, but we believe these approaches can be quite useful.